Technische Universität Berlin

Zentrum Technik und Gesellschaft

Discussion paper Nr. 31/11

Dezember 2011

Jochen Gläser und Grit Laudel

Life With and Without Coding

Two methods of early-stage data analysis in theory-guided qualitative research

Nr. 31/**11**y Jochen Gläser und Grit Laudel

Life With and Without Coding

Two methods of early-stage data analysis in theory-guided qualitative research

Zentrum Technik und Gesellschaft
Sekretariat: HBS1
Hardenbergstraße 16-18
10623 Berlin

Tel: (030) 314-23 665
Fax: (030) 314-26 917

www.ztg.tu-berlin.de

Zusammenfassung

Qualitative Forschung, die ‚mechanismische' Erklärungen konstruieren will, muss die bereits existierende Theorie mit in den Daten gefundenen Mustern integrieren, was eine besondere Herausforderung für qualitative Datenanalyse darstellt. In diesem Paper diskutieren wir die Nutzung zweier Methoden – Kodieren und qualitative Inhaltsanalyse – als erste Schritte im Prozess der Datenanalyse. Beide Methoden produzieren durch Kategorien strukturierte Datenbasen, die für die Suche nach Mustern in den Daten und deren Integration in systematische, theoretisch eingebettete Erklärungen genutzt werden kann. Wenn es als separate Methode außerhalb der Grounded Theory genutzt wird, führt Kodieren auf einen indizierten Text, d.h. der Text selbst und die seinen Index bildenden Codes werden in den folgenden Schritten weiter verarbeitet. Qualitative Inhaltsanalyse extrahiert die relevanten Informationen (separiert sie vom Text) und verarbeitet sie getrennt weiter. Die qualitative Inhaltsanalyse scheint dem Kodieren überlegen, wenn Forschungsfragen aus der Theorie abgeleitet werden und ihre Beantwortung auf Informationen über die relative Position von Aussagen in Texten verzichten kann, was bei der Suche nach ‚mechanismischen' Erklärungen häufig der Fall ist. Coding scheint überlegen, Forschungsfragen stärker explorativer Natur oder die Position von Informationen im Text für die Beantwortung der Forschungsfragen relevant ist.

Abstract

Qualitative research aimed at ‚mechanismic' explanations poses specific challenges to qualitative data analysis because it must integrate existing theory with patterns identified in the data. We explore the utilisation of two methods – coding and qualitative content analysis – for the first steps in the data analysis process, namely 'cleaning' and structuring qualitative data. Both methods produce an information base that is structured by categories and can be used in the subsequent search for patterns in the data and integration of these patterns into a systematic, theoretically embedded explanation. Used as a stand-alone method outside the grounded theory approach, coding leads to an indexed text, i.e. both the original text and the index (the system of codes describing the content of text segments) are subjected to further analysis. Qualitative content analysis extracts the relevant information, i.e. separates it from the original text, and processes only this information. We suggest that qualitative content analysis has advantages compared to coding whenever the research question is embedded in prior theory and can be answered without processing knowledge about the form of statements and their position in the text, which usually is the case in the search for 'mechanismic' explanations. Coding outperforms qualitative content analysis in research that needs this information in later stages of the analysis, e.g. the exploration of meaning or the study of the construction of narratives.

Contents

1. Positioning methods in qualitative research processes[1]

The field of qualitative research encompasses a wide range of aims and methods which are only weakly related to each other. This makes choosing the best method for a specific aim a difficult and risky decision. The aim of this paper is to position two of the most widespread methods of qualitative data analysis, coding and qualitative content analysis, in the qualitative data analysis process leading to causal explanations, and to compare the role both methods can play in the analysis. By doing this, we would like to contribute to a qualitative research methodology that systematically links types of research problems to methods.

Such a methodology is currently almost nonexistent for at least three reasons. First, there is a great variety of types of research goals of qualitative research. Its empirical objects include phenomena from all levels of aggregation and include individual constructions of meaning as well as societal processes, and its ambitions range from describing empirical phenomena to providing theoretical explanations. Second, protagonists of methods are often reluctant to link their methods to specific types of goals, which would include characterising not only the input but also the output of the methods. The enormous variation between the approaches, their partial overlap, and the breadth of legitimate research goals in qualitative research make it impossible to construct a framework in which all methods can be located. Methods are usually described without any reference to other methods.

Third, many qualitative methods claim to lead to an answer to the research question but do not specify all steps between the text and the answer. This is not surprising because qualitative research places heavy emphasis on interpretation. Interpretation is an ill-structured activity for which no algorithm can be provided. At the same time, the widespread reluctance to define intermediary steps and their outputs makes it often difficult to assess the contribution of a specific method along the way from texts to answers to research questions and the quality of that contribution.

By comparatively positioning qualitative content analysis and coding in the data analysis process, we intend to achieve three clarifications. First, the comparison will lead to a better understanding of the logic underlying the first steps of qualitative data analysis for causal explanations. Second, we present an alternative to coding that achieves a stronger reduction and structuration of the data material during these first steps, and therefore might be more appropriate for some qualitative research processes. Thirdly, our comparison provides at least some ideas about the range of applicability of methods, thereby contributing to the badly needed methodological background that tells us what to use which qualitative method for.

[1] We would like to thank Martina Schäfer, Melanie Jaeger-Erben and the participants of the Lugano 2011 data analysis workshop for helpful comments.

In order to clarify the role of specific methods in qualitative data analysis, it appears necessary to work backwards from the research goals by distinguishing types of goals and asking what steps must be gone to reach these goals. This is an enormous task that has not even been formulated, let alone been worked at. We will address this problem only briefly to the extent to which clarification is necessary for the purposes of this paper (1). Thus, we will establish theoretical explanations that use the identification of causal mechanisms as one type of research goals of qualitative research (2), discuss the steps that need to be gone in order to reach this goal (3), position both coding (4) and qualitative content analysis (5) in this sequence of steps. Our comparison demonstrates that while coding is often focused on the earliest stage of data analysis, both coding and qualitative data analysis can be used equivalently because they both lead to a database that is used in the subsequent search for patterns in the data (6).

2. Theoretical explanation as an aim of qualitative research

If one compares current discussions on qualitative data analysis in sociology and political science, a remarkable bifurcation becomes apparent. Political scientists appear to be primarily concerned with the question how theory can be advanced by using comparative case studies that employ qualitative methods without spending too much thought on how the data should be created and processed in the first stages of the analysis. The current discussion on "qualitative methods" and "comparative case studies" just assumes that the data are there, i.e. can be produced in the form necessary for theoretical analysis. No requirements concerning data collection or data analysis are formulated in the various suggestions for producing theories with case studies (see e.g Ragin and Rihoux 2004; George and Bennett 2005; Mahoney 2008).

Sociologists, on the other hand, are (still) focused on the ways qualitative data (texts and pictures) can and should be analysed but are rather vague about what such an analysis is supposed to achieve. When research goals of qualitative sociological research are mentioned, they remain highly abstract and sufficiently vague to suggest that methods for the analysis of qualitative data are only weakly associated with different types of research goals. For example, Boeije (2010: 11) writes that the purpose of qualitative research is "to describe and understand social phenomena in terms of the meaning people bring to them". Bernard and Ryan (2010: 8-10) list exploration, description, comparison, and testing models as the four "research goals" of qualitative research. Theory building does not occur in this list and generally appears to play a minor role. The only methodological context in which it is systematically treated is that of the "Grounded Theory" approach that builds new theory from empirical data (e.g. Glaser and Strauss 1967, Strauss and Corbin 1990, Glaser 1992).

Assessing the analytical strategies proposed by both camps would require linking data requirements of the theory building proposed by political scientists to the properties of data that emerge from the qualitative data analysis methods proposed by sociologists. This cannot be done here. Instead, we take what we consider to be the most ambitious goal of social science research – causal explanation - as an assumed outcome of qualitative data analysis, and work our way backwards from this outcome towards the first step of data analysis. This way we can chart one way from start to finish without having to take into account all other possible outcomes and ways towards them. For our purposes, it is sufficient to assume that at least some strategies of qualitative research will be covered by the frame suggested here and can make use of it, and leave the classification of research goals and the systematic comparison between types of goals and sequences of steps from data to answers to research questions to further work.

We suggest that the most ambitious goal is to provide a generalised causal explanation of a class of social phenomena, i.e. an explanation that links

types of conditions, types of causal social mechanisms, and types of effects (Maxwell 2004a, b).[2] Rather than providing cause-effect relationships, a full explanation must describe how the effects are produced. A 'mechanismic' explanation is the only one that does not leave any black boxes and therefore can be considered satisfactory.

Such a 'mechanismic' explanation is inevitably embedded in a theory (of a middle range, see Merton 1968a). The possibility of developing causal explanations with qualitative methods is still contested by both quantitative and qualitative researchers (Maxwell 2004b). Furthermore, it is important to note that the specific understanding of mechanisms and explanations provided here is not the only one and possibly not even the dominant view of the qualitative methodology community.[3] However, it is one of the few approaches that systematically link qualitative methodology and theory by supporting both theory-driven qualitative research and generalisations from qualitative data.

The idea of social mechanisms (we use 'causal' and 'social' mechanism as interchangeable throughout this paper) has been introduced by Merton (1968a). Interest rekindled in the 90ies, and many different understandings of causal mechanisms have been suggested in the literature since then (see Mahoney 2003 for an overview of definitions). Following Mayntz (2004: 41), we define a social mechanism as a sequence of causally linked events that occur repeatedly in reality if certain conditions are given and link specified initial conditions to a specific outcome (for a similar but less precise definition see Merton 1968a: 42-43; for many other versions see Mahoney 2003: 13-15). Being ambitious, we can state that a full explanation of a social phenomenon has been achieved if we can provide the social mechanism by which the phenomenon was produced and the conditions that initiated and upheld the mechanism. A list of 'causal factors' that produce a phenomenon with a certain probability is only a contribution to an explanation.[4]

Merton's description of the mechanism of the self-fulfilling prophecy illustrates the understanding of social mechanisms applied in this paper: If an incorrect definition of a situation is made part of the situation and affects human behaviour, then people might act in a way that makes this definition of the situation accurate. The classic example is a perfectly healthy bank that rumours describe as being in financial trouble. The rumour affects belief formation, which makes an increasing number of customers seek to

---

[2] Thus, we would argue that the first three of Bernard's and Ryan's "research goals" (exploration, description, and comparison) are just means to an end, and that the fourth (model testing) is an unduly narrow representation of theory development.

[3] For a very extensive and interesting treatment of causal explanation through case study research see George and Bennett (2005).

[4] Ragin's "Qualitative Comparative Analysis" (QCA) is similar to the statistical approach in its attempt to causally link conditions to effects without providing the mechanisms (Ragin 1987, 2000; Ragin and Rihoux 2004). The major difference is that QCA is applied to cases that are described by dichotomous or fuzzy values of variables, and uses Boolean logic rather than statistical reasoning. QCA has interesting problems of its own, which we cannot discuss here (see e.g. Kelle 2003, Lieberson 2004).

withdraw their money from the bank, thereby creating the liquidity problems reported by the rumour. This spreads a new definition of the situation which makes even more people withdraw their money (Merton 1968b). This cycle of events - belief formation, acting on that belief, creating the situation that was believed to be true – is a very general and powerful mechanism. We can also see some of the conditions that trigger and maintain the mechanism: The definition of the situation must be public and communicable, it must be strong enough to lead to the actions through which the mechanism operates, and the results of these actions must change the definition of the situation.

The mechanisms we look for in our empirical investigation are likely to be much more specific. However, the social explanation we aim at contains the following:

a) The generalised description of one or several mechanisms (the events, the ways in which they are linked, and the outcomes of the sequence), and

b) The conditions that are necessary to trigger and to sustain the mechanisms (as well as promoting and hindering conditions).

Taken together, this amounts to a system of statements that links varying conditions, varying mechanisms and varying outcomes, i.e. a contribution to a theory of middle range.

If we assume 'mechanismic' explanations to be the ultimate goal of social sciences, quantitative and qualitative methods can be seen as being 'epistemologically equal' and complementary: Quantitative methods can establish causal relationships and their range of validity (the conditions or population for which they hold) but are unable to identify causal mechanisms while qualitative methods can identify causal mechanisms but are unable to empirically delineate the social realm in which they operate.

The crucial and, unfortunately, so far unanswered question is how to identify social mechanisms from descriptions of social reality provided in the texts we analyse. We can tentatively state that since we are looking for generalised descriptions of mechanisms and conditions, it is necessary to identify *patterns in our data*. These patterns need to be *integrated*, e.g. by developing *typologies.* These typologies must satisfy both the *theory* that was used to structure data collection and the data, i.e. the *empirical variation of phenomena* described by our data. Theory may be revised or even substituted in that process.

Searching for 'mechanismic' explanations is thus best achieved by constructing and linking two kinds of typologies, namely typologies of conditions and typologies of mechanisms. Typologies for conditions can be created on the basis of any property of theoretical interest but would include as one of their dimensions the impact of conditions on mechanisms (promo-

ting, hindering, necessary, and sufficient).[5] Typologies of mechanisms are difficult to further categorise because we just don't know enough about mechanisms. However, we would expect a typology of (sequences of) events which includes conditions that set this sequence in motion and outcomes of that sequence.

The discussion of 'mechanismic' explanations enables one further conclusion: In order to identify social causal mechanisms and the conditions under which they operate we need rich descriptions of empirical phenomena (because we know only partially what we are looking for) and variance. The latter is crucial for assessing the conditions under which mechanisms operate or don't operate. If we have just one case, it is difficult if not impossible to abstract from the empirical circumstances of that case and to identify conditions as necessary or sufficient or the sequence of empirically observed events as an instance of a more general mechanism. This is why single case studies often use "in-case comparisons", which effectively means that they construct several cases within one. A well-known example is 'before-after' designs that divide a single longitudinal case into two sub-cases (George and Bennett 2005: 166-167).

In qualitative research, the best way to obtain variance is to conduct comparative case studies. With regard to conditions, Ragin (2000) stated that necessary conditions can be found by identifying causal conditions shared by cases with the same outcome, while sufficient conditions can be found by examining cases with the same causal conditions to see whether they also share the same outcome. So far we have not found suggestions how to exploit variance in the search for causal mechanisms.

[5] For an interesting discussion of the role of necessary and sufficient conditions as well as INUS (insufficient necessary parts of a condition that itself is unnecessary but sufficient) and SUIN (sufficient but unnecessary parts of a factor that itself is insufficient but necessary), see Mahoney (2008). It is important to note that the analysis of conditions will always start with all conditions observed. The subsequent analysis should not be reduced to causal conditions in the sense of conditions that make or don't make something happen. An analysis aiming at identifying causal mechanisms would always be interested in conditions that shape the process linking causes and effects (i.e. slowing them down or speeding them up).

3. From texts to explanations

We can now specify the sequence of steps between texts and answers to research questions: The last two steps in our analysis can be described as "search for patterns" and "integration of patterns". These steps are obviously highly creative, and it remains to be seen how much support for them can be provided by qualitative data analysis methods. But whatever their support for the more creative steps is, the qualitative data analysis methods must at least prepare the data for them.

In order to arrive at explanations of social situations or processes, we need to systematically reduce the complexity of the data we generated in the qualitative data collection. While it is absolutely central to qualitative research to create this complexity in the first place, it is nevertheless essential to reduce the complexity of the data in order to arrive at generalised explanations. Qualitative analysts only have the choice between reducing complexity stepwise and systematically or – if they treat whole texts as data during the whole course of their analysis – to reduce complexity spontaneously und subconsciously. While the latter 'approach' may be retreating from qualitative research (not least thanks to the spread of coding as a data analysis technique!), it still exists. The two methods discussed in this text support the former approach.

This link between a specific research goal and the two methods discussed here is neither exclusive nor exhaustive. Our argument is that if we want to find explanations consisting of causal conditions, other conditions, and possibly mechanisms, we need to systematically reduce complexity and bring our data in a form that supports pattern recognition. However, the two methods can be (and are) applied in the context of other approaches, too. At the same time, there are also qualitative data collection and analysis procedures for which there is no irrelevant information because all content as well as the form of a text contributes to the answer to the research question. For example, narrative interviews are often conducted and analysed with the aim of identifying structures of whole texts. In investigations of this kind it is very important what was said in what order and in which context, which makes every utterance in the text extremely important (e.g. Cortazzi 1999)

Having identified "search for patterns" and "integration of patterns" as the last steps, we can turn the sequence around and start from the text. Diagram 1 shows the steps and includes a point we will not further discuss, namely that any practical analysis will move back and forth between steps.

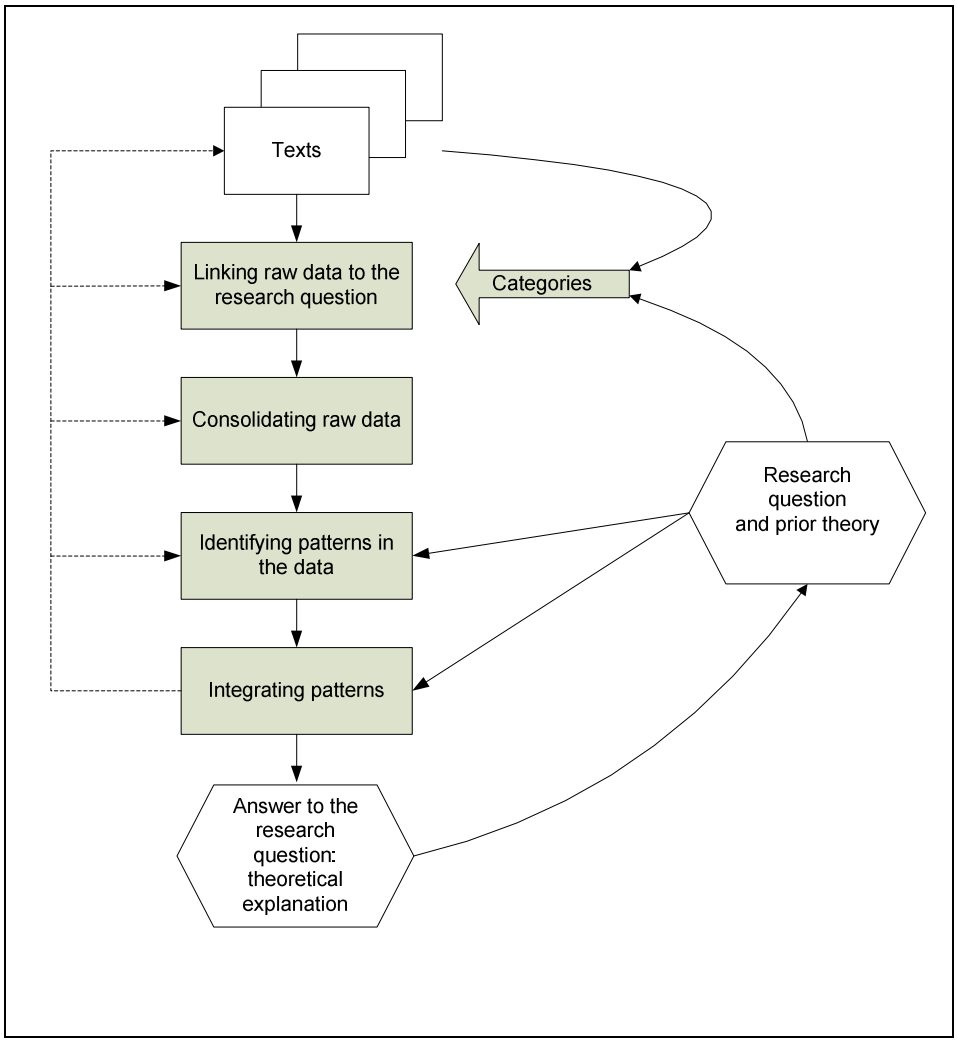


Diagram 1: Steps between texts and explanation

3.1. Linking raw data to the research question

The texts we work with are linked to the research question because their production or collection were guided by information requirements derived from that question. However, at the beginning of the analysis the information contained in the texts is not systematically linked to the research questions, nor is it structured according to the theoretical background against which it has been formulated. This is why qualitative data analysis begins with linking raw data in the texts to the research question. This first step includes identifying, locating, and structuring raw data. These operations are only analytically separable, in most cases they are conducted simultaneously when texts are processed for the first time.

3.1.1. Identifying and locating raw data

Raw data are those containing information relevant to our research question. A data analysis that moves from texts to theoretical explanations assumes that not all that is said in a text is relevant to a specific research question. In many research processes, collecting qualitative data inevitably includes the creation of large amounts of unnecessary information. The

texts we create or collect during fieldwork contain redundant information – things that were repeatedly said or written by respondents or repeatedly recorded by observers – and irrelevant information – things that were said, written, or recorded but have nothing to do with the research question. This 'dilution' of relevant information – the data – by irrelevant information is a necessary corollary of qualitative data collection because one of the latter's tenets is to assign more control of data generation to respondents. Since the frames of reference and frames of relevance of respondents are different from ours, they will also communicate information we don't need. This is why the first step of qualitative data analysis often is to identify and locate the relevant raw data in the abundance of information created in the data collection process.

Identifying and locating raw data includes two tasks. First, the analyst needs to recognize parts of the text as containing information that is relevant for answering the research question. Second, a decision needs to be made which variable/category the information belongs to. These tasks are practically inseparable because the decision about relevance includes a reason ("Why is this information relevant?"), an argument which cannot be made without reference to specific aspects of the research questions and the variables describing these aspects.

This first step is already based on interpretation by the analyst. Depending on their interpretation, analysts will regard different parts of a text as relevant, and will link parts of texts to different variables/categories.

3.1.2. Structuring raw data

Structuring raw data means detailing the link between the data and the research questions and identifying links between data. The major tool for this step is a system of categories. The nature and role of these categories often remain unclear in the literature. They are best imagined as containers for empirical information that can be either created ad hoc (based on the information in the text) or derived from theoretical considerations. Categories are thus similar to variables because they, too, are constructs that can assume different states depending on the empirical phenomena they describe. In order to simultaneously satisfy both the demand that research needs to be linked to theory and the demand that research needs to be open to unexpected information in texts, qualitative data analysis methods are usually based on a combination of both strategies of category creation, which can mean:

- some categories are derived from empirical information in the text, others from theory,
- categories are derived from theory and changed or supplemented according to empirical information in the text, or

- categories are derived from empirical information in the text and later changed in the light of applicable theories.

There are three ways in which a link between empirical information and categories can be achieved, which vary in the degree to which the form of the data is changed.

a) Indexing themes: Data are indexed by attaching codes (short strings that represent categories and that briefly state what the information is about) to the part of the text containing the information. The outcome is an indexed text, i.e. a text with attached codes that list the relevant themes addressed in each part of the text.

b) Indexing content: Data are translated into the analytic language of the investigation, i.e. into statements that describe states of categories. These descriptions are attached to the text. The resulting index contains not only labels indicating what was talked about in the text but also labels representing what was actually said.

c) Extracting content: Raw data are translated into the analytic language of states of categories and moved into a database that collects raw data according to the categories for which they are relevant. Thus, a database is created that contains only raw data and supplants the original text in the further analysis. The link to the source (the part of the original text) is kept with the data throughout the analysis, but the original text is only rarely used after the extraction.

Indexing and extracting raw data are two different approaches to the same problem, namely the mix of data and noise in the original text. Indexing keeps the text as it was and adds information as to where the raw data belonging to the various categories are located (indexing themes) or what states or values of categories are located at a specific position (indexing content). Thus, they define all information that has not been coded as noise. Extracting content means separating the relevant information from the text, subsuming it to categories and storing it separately for further processing. The noise remains with the text that is not analysed anymore. All three variants of structuring the raw data are based on the analyst's interpretations of the relationships between information in the text and the system of categories. It is important to note here that the three ways of preparing raw data or the two methods do not differ with regard to their openness for unexpected information. All three provide opportunities for changing the system of categories during the processing of data. This is achieved by adding or changing categories.

3.2. Consolidating raw data

While the first step has separated raw data from noise, it has left the raw data as they were in the text. This means that they are still likely to contain errors and redundancies. Cleaning the data means collating redundant information and correcting obvious errors. This further reduces the amount

and the complexity of the data we have to deal with. Again, this step depends on interpretations (of data as redundant, of data as erroneous).

3.3. Searching for patterns in the data

We now arrive at the steps that, while still supported by methods, are crucially dependent on the analyst's creativity and ability to recognize patterns in the data. The contribution of methods to the search for patterns in the data is their support of ordering the data according to various principles (time, actors, actions, locations and so on). The patterns in the data we search for include:

- More-than-once-occurring sequences of events,
- More-than-once-occurring combinations of conditions/ processes/ outcomes,
- Conflicting accounts of events or processes.

Since patterns can be recognised only if the idiosyncratic description of information in the original text has been standardised somehow, the procedures that provide data that are already translated into the analytic language are more suitable for this step than those that just indicate where the information is but leave the information itself in its idiosyncratic form. But even when reformulated in an analytic language, the amount and the complexity of data are usually far too high for patterns to become easily visible.

A useful starting point for the search for mechanisms is a 'thin description' of sequences of events within each case, which can then be compared between cases with the aim to detect repeating patterns or important differences that can be linked to variations in conditions.

More generally speaking, the most powerful tool for doing this is building typologies. Building typologies means selecting very few (up to three) variables, identifying distinct (qualitatively different) states of these variables and defining the combinations of these states as types. This can be done with only one variable (chain smokers – occasional smokers – non-smokers) but is much more interesting when done with two or three variables because the complexity can be further reduced. For the number of types to be low enough to be handled in the search for patterns, the maximum number of dimensions is probably three. With three dimensions and two states in each dimension one arrives at eight types, for three states we have already 27 types. Diagram 2 provides an example of a typology.

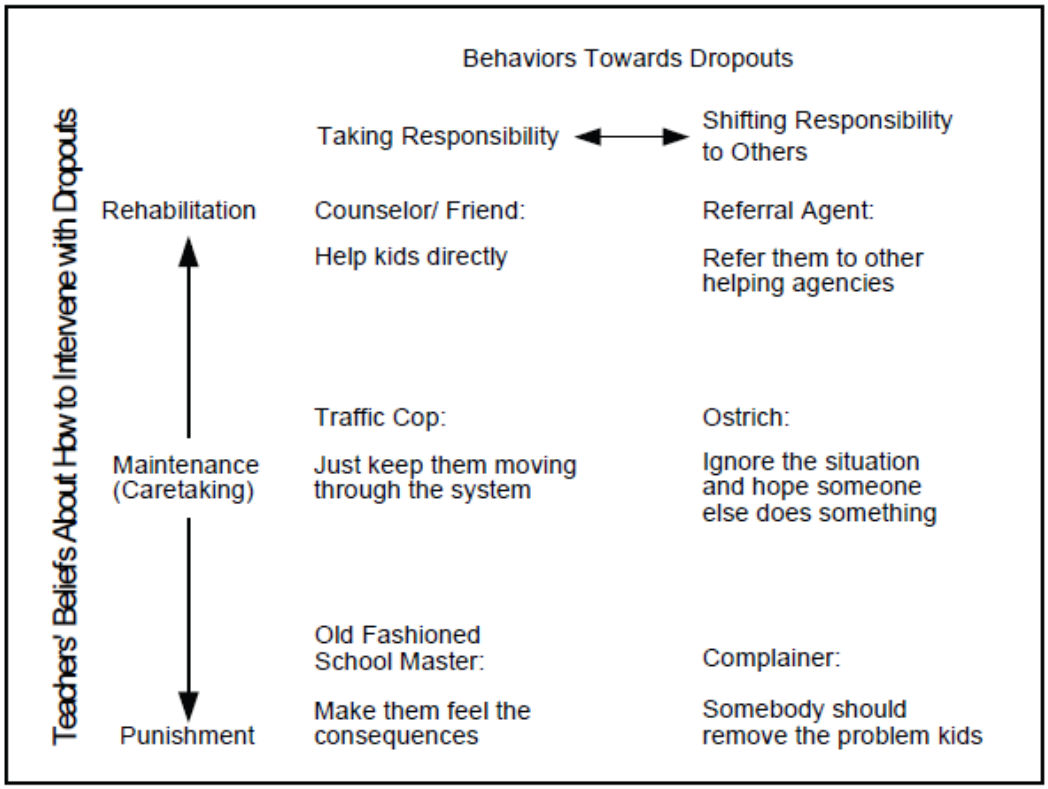


Diagram 2 Example of an empirical typology: „An Empirical Typology of Teacher Roles in Dealing with High School Dropouts" (Patton 1990: 413)

If we have types that describe important patterns in our data, they can be used as a starting point for a further search by combining them with other categories. Another strategy that might lead to pattern recognition is the search for extreme or counterintuitive examples. These examples usually shed light on necessary conditions for a sequence of events or on rarely occurring sequences of events.

These techniques are usually first applied to cases, i.e. used in the analysis of comparative case studies. However, while cases are important because knowledge about important variations of variables went into their selection, it is dangerous to simply assume that 'we got it right' when we selected cases, i.e. that cases represent different mechanisms or distinct patterns of data. Within-case variation may be more important than between-case variation. In many investigations, analysts deal with nested cases, i.e. definitions of cases at more than one level of aggregation. Thus, while cases are the first port of call for the search for patterns, it is always useful to compare data patterns across cases.

The major technical requirement for a search for patterns is that the data base constructed in the previous steps enables an easy rearrangement of data. After all, pattern recognition is recognising characteristic combinations of data, which is most easily achieved when we try groupings of data and look at them. Thus, manipulations of data such as the sorting of data, the selection and re-arranging of variables must be easy. In principle, both

coded texts and the data bases produced by qualitative content analysis provide these opportunities. However, the already reduced complexity and the removal of noise from the data makes the products of qualitative content analysis more suitable.

3.4. Integrating patterns

Once we have found patterns, it is important to integrate them. The first question is usually whether all patterns are in fact different or whether some of them can be merged into one. Having done this, integration means linking all data that had no part in the identification of patterns to these patterns. This usually means that more conditions for mechanisms or outcomes will be added. Furthermore, it is very important that all cases and data that don't fit the patterns are fully explored. While statistical reasoning is happy to flag the proportion of explained variance, unexplained variance is simply inadmissible for a 'mechanismic' explanation. Therefore, we have to provide explanations for any combination of data we have found. This explanation may be idiosyncratic. However, it must be provided.

If we have included all information and explained the variance, we can attempt generalisations. In the case of a 'mechanismic' explanation, a generalisation usually takes the form "Whenever conditions A exist, mechanism X is most likely to operate and to produce outcome Ω". This has been called "contingent generalisation" in the literature (George and Bennett 2005, Blatter and Blume 2008). As was mentioned above, a qualitative investigation is unable to empirically establish in which population conditions A exist. A generalisation that specifies the conditions under which a mechanism produces an outcome is theoretically precise but empirically contingent. However, we want to leave something for the colleagues from quantitative research: Empirically establishing the scope of a mechanism (the occurrence of conditions that trigger it and let it operate) can only be achieved by drawing a representative sample of a population and statistically generalising from the sample to the population.

4. Coding

4.1. History and variations

Coding is a very old technique that has been widely used for a long time to structure text (Kelle 1997: 2.1). It became popular as basic technique of the Grounded Theory methodology (Glaser and Strauss 1967), and is today probably the most popular technique of data analysis. This idea has migrated from the grounded theory approach into general qualitative data analysis. Today coding is a recommended technique of qualitative data analysis in other approaches that do not explicitly subscribe to Grounded Theory. Among the authors who recommend coding outside the grounded theory approach are Miles and Huberman (1994, Chap. 4), Patton (1990: 381-384), and Boie (2010: 93-121). Most commercial and freeware software packages for the support of qualitative data analysis (e.g. ATLAS/ti, MAXQDA, and NVIVO) support coding and only coding, thereby contributing to its transformation into a disembedded standard technique of qualitative data analysis (Coffey et al. 1996).

Our discussion of coding refers to its 'decontextualised' application as a set of procedures that is used both within and outside the Grounded Theory approach. This distinction is important because our discussion of coding must not be seen as a discussion of Grounded Theory. We discuss coding procedures developed and used in the Grounded Theory approach (with a special emphasis on the suggestions in Strauss and Corbin 1990) as one version of coding besides others that are recommended in the literature.

It comes as no surprise that the coding-based approaches to qualitative data analysis vary in their underlying methodologies and in the use they make of coding. Two important variations, which we briefly discuss below, concern the extent to which preexisting theory is used in the coding process and the distinction between indexing themes and indexing content.

Thus, grouping all coding techniques under the one heading seems to be an inappropriate simplification. However, there are some basic aspects shared by all coding techniques, and we can focus on these aspects in our demonstration of the difference between coding and qualitative content analysis.

4.2. Codes as categories

The core idea of coding is that the texts containing the raw data are *indexed*. Codes – keywords, phrases, mnemonics, or numbers – that signal the occurrence of specific information are assigned to segments of the text. In the list of codes, each code is linked to all text segments to which the code has been assigned. See, for example, the description by Miles and Huberman:

> "Codes are categories. They usually derive from research questions, hypotheses, key concepts, or important themes. They are retrieval and organizing devices that allow the analyst to spot quickly, pull out, then

> cluster all the segments relating to the particular question, hypothesis, concept, or theme. Clustering sets the stage for analysis" (Miles and Huberman 1984: 56)

This quote describes both the function of codes and a particular way in which they can be constructed. The function of codes is to support the retrieval of text segments, which in turn can be used to group them according to thematic aspects of the data they contain. This function of a code, which is essentially that of an index of a book, stood at the beginning of coding.

Codes can be derived either from theoretical considerations prior to the data analysis or from the text itself. The place of theory in the development of codes is one of the crucial differences between the various coding-based methods. The 'grounded theory approach' originally demanded

> "literally to ignore the literature of theory and fact on the area under study, in order to assure that the emergence of categories will not be contaminated ..." (Glaser and Strauss 1967: 37).

This position is epistemically naïve because it ignores the theory-ladenness of observations (which, curiously enough, is acknowledged by Glaser and Strauss in the same book on page 3). The subsequent development of the grounded theory approach by its two proponents has done little to remedy this problem. According to Kelle's (2005) excellent discussion of this problem, "[m]uch of GLASER's and STRAUSS' later methodological writings can be understood as attempts to account for the "*theoryladenness*" of empirical observation and to bridge the gap between "emergence" and "theoretical sensitivity" (ibid: [10]). However, the original limitations remain:

- "Theoretical sensitivity" is developed by engaging with existing theory but is understood as the ability to "see relevant data" (Glaser and Strauss 1967: 46, see also Strauss and Corbin 1990: 76) or "an awareness of the subtleties of meaning of data" (Strauss and Corbin 1990: 41). This utilisation of existing theory is indirect and unsystematic. It relies on developing personal qualities of researchers, which are then brought to bear on the data spontaneously rather than in methodologically controlled way.
- After the bifurcation of Grounded Theory, both Glaser and Strauss have included procedures for the integration of codes that give the impression of being theoretical. In his book on 'Theoretical Sensitivity' Glaser (1978) suggested to supplement 'substantive coding' (developing codes during the analysis) with 'theoretical coding'. The 'theoretical codes' are "from highly diverse theoretical backgrounds, debates and schools of philosophy or the social sciences" (Kelle 2005: [13]). They are very general and include concepts like 'limit', 'extent', 'goal', or 'social norms', and are grouped in 14 'coding families' which are not disjunct. Strauss and Corbin (1990) chose a similar approach by supplementing 'open coding' (developing codes while reading the texts) with 'axial coding'. During 'axial coding', the categories and concepts that have been developed during 'open coding' are integrated and empirically generalised by orga-

nising them into a 'coding paradigm' that links phenomena to causal conditions, context of the investigated phenomena, additional intervening conditions, action and interactional strategies, and consequences (ibid: 96-115). While their 'coding paradigm' is better structured and more consistent than Glaser's 'coding families', the two approaches have in common that they are inspired rather than informed by social theory. Strauss and Corbin refer to Grounded Theory as an "action-oriented model" (ibid: 123) without specifying theoretical sources for their paradigm, while Glaser's (1978) list is obviously eclectic. Both approaches to including 'theory' are applied to codes that have been developed in a supposedly theory-free step.

These procedures still ban specific theory from coding and relegate some general theory to its fringes. They thereby deprive researchers of the most important function of prior theory, namely that of a source of (comprised, abstracted) information about the empirical object of their research.[6]

In contrast, Miles and Huberman state very clearly that *all* theory is an important source for constructing codes:

> „One method – the one we prefer – is that of creating a ‚start list' of codes prior to fieldwork. That list comes from the conceptual framework, list of research questions, hypotheses, problem areas, and key variables that the researcher brings into the study." (Miles and Huberman 1994: 58)

An apparent additional advantage of coding – to start 'from scratch', i.e. without variables/categories that are defined ex ante – is in fact a disadvantage because it hides the fact that it is impossible to conduct an analysis without prior assumptions. Defining at least some codes and categories ex ante just forces analysts to make their assumptions explicit.

---

[6] Even though it is not our topic here, we would like to note that the version of Grounded Theory presented by Strauss and Corbin is still averse to existing theory. Consider this: Strauss and Corbin (1990: 29) state that the difference between theory and description is that "theory uses concepts" and "the concepts are related by means of statements of relationship", which fits the outcomes of the Grounded Theory approach but is a rather weak definition of theory. In their chapter on "The Uses of Literature" they reiterate the original theory-phobic stance of grounded theory: "So, if you begin with a list of already identified variables (categories), they may – and are indeed very likely to – get in the way of discovery. Also, in grounded theory studies, you want to explain phenomena in light of the theoretical framework that evolves during the research itself; thus, you do not want to be constrained by having to adhere to a previously developed theory that may or may not apply to the area under investigation." (ibid: 49). This is not different from the 1967 statement quoted above. When discussing the naming of categories, Strauss and Corbin state that is possible to use existing theory (the "technical literature") as a source but effectively discourage the reader to do so (ibid.: 68-69). Glaser's version, at least in the form recently presented by Charmaz (2006), is very similar in this respect. Charmaz applies 'theoretical sampling', 'theorizing' and 'theoretical sensitivity' exclusively to the construction of theory from data (ibid.: 99-108, 133-140). Theory constructed by other researchers enters her version of grounded theory when the 'literature review' and 'theoretical framework' are written (ibid.: 163-172), which Charmaz recommends to delay until the empirical research is finished (ibid.: 166).

Codes can be hierarchical or a network of equally-ranked terms. Miles and Huberman state that it is no problem to start the analysis with a list of 80 to 90 codes if the list has a clear structure. Indeed, their example list contains less than 10 main codes with subcodes and sub-subcodes (Miles and Huberman 1994: 58-59). Diagram 3 contains a section of the list of codes provided by Miles and Huberman. Codes can be numbers, mnemonics, single words, or short phrases.

External Context	EC (PRE) (DUR)
EC: Demographics	EC-DEM
In county, school personnel	ECCO-DEM
Out county, nonschool personnel	ECEXT-DEM
EC: Endorsement	EC-End
In county, school personnel	ECCO-END
Out county, nonschool personnel	ECEXT-END
EC: Climate	EC-CLIM
In county, school personnel	ECCO-CLIM
Out county, nonschool personnel	ECEXT-CLIM
Internal Context	IC (PRE) (DUR)
IC: Characteristics	IC-CHAR
IC: Norms and Authority	IC-NORM
IC: Innovation History	IC-HIST
IC: Organization Procedures	IC-PROC
IC: Innovation-Organization Congruence	IC-FIT
…..	

Diagram 3: Example for a start list of codes (extract), source: Miles and Huberman (1994: 58-59)

Miles and Huberman (1994: 58-66) list several strategies for creating structured lists of codes. All these suggestions refer to coding as indexing themes, i.e. adding a hyperlinked index to the text that provides information about what was talked about where – just as the index of a book does.[7] The second possibility, which is not often mentioned in the literature, is to go one step further by indexing content – i.e. not only what was talked about but what was actually said. This can be easily achieved by adding another level of hierarchy to the 'code tree' – to each code, a short description of the content of the information found in the text segment is added as a new subcode. Thus, the code 'receives' as many content subcodes as there are text segments to which it is applied, or fewer if the same content occurs in more than one text segment. This appears to have been the approach of Turner (1981) who – in the pre-software period of coding – constructed

[7] This also applies to the "pattern coding" suggested by Miles and Huberman (1994: 69-72). The pattern code "RULE-INF" would index "rules about informant behavior" – without specifying what the rule says (ibid: 72).

coding cards on which he collected all content that was reported for one code in the text (diagram 4).

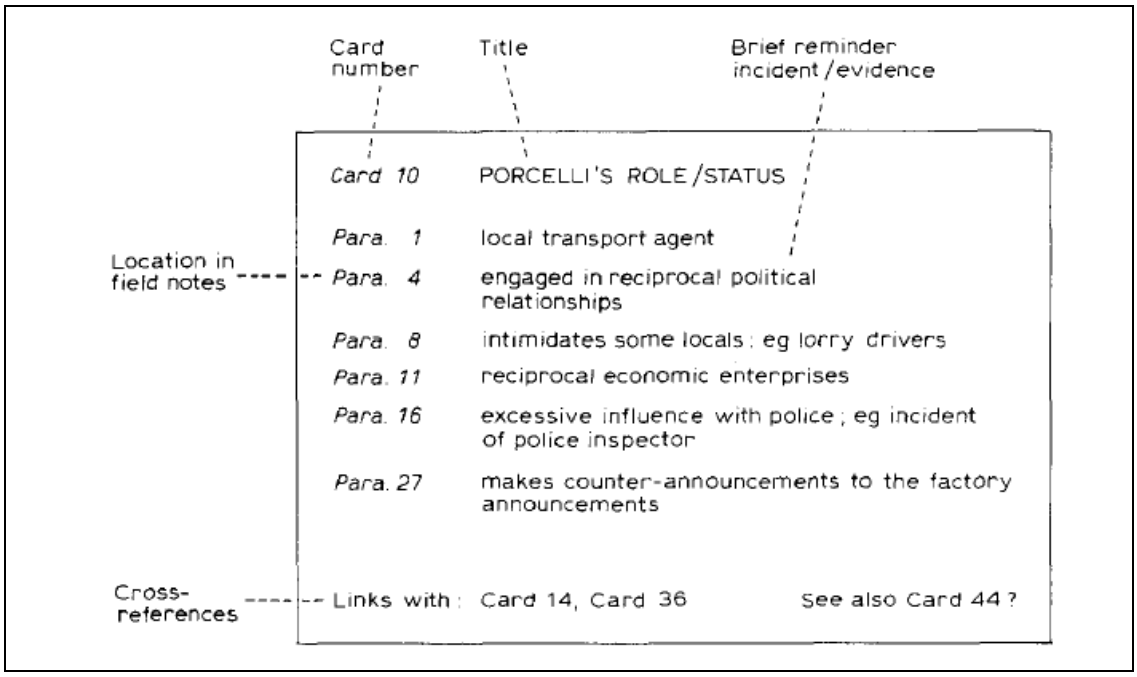


Diagram 4: Example of a qualitative data category card (taken from Turner 1981: 234)

Indexing content is also the outcome of the process of developing a category during the 'open coding' proposed by Strauss and Corbin (1990: 69-72). Building categories represent the first empirical generalisation because they group concepts under a higher order, more abstract concept (the category). The empirical phenomena represented by a category have certain properties, which can vary. 'Dimensionalizing' a property means establishing the dimension in which they vary. Both properties and dimensions are discovered in the data (ibid: 69).[8] 'Dimensionalizing' the properties of a category makes it possible to locate each instance of a category that occurs in the data somewhere along the dimensional continua (ibid: 70), which amounts to indexing content.

Indexing content comes close to extracting content (translating the information contained in the text in an analytic language and storing it elsewhere) because the codes would contain all necessary information and could be used in the subsequent analysis separately from the text. This version of coding appears to be rarely used. As the description of codes by Miles and Huberman (above, section 4.2) and the section on further processing (below, 4.4) suggest, the most widespread version appears to be indexing texts for further processing. A possible explanation for the predominance of indexing texts is that indexing content is not too well supported by

[8] An interesting disadvantage of not using prior theory becomes apparent here. If the category is related to a theoretical concept, theory might help find properties that are never mentioned in the material but should be, thereby offering an additional perspective on the data.

the commercial software packages for coding. These software packages enable only two kinds of writing activities: Attaching a code, which often is limited to few characters or writing memos, which can be long but cannot be structured and are difficult to handle or analyse.

4.3. Coding texts

Coding means applying codes to text. The text needs to be structured. It will usually consist of paragraphs that characterise narrative units. However, codes can be applied to text segments of all lengths – from a single word to a whole text. Depending on the strategy chosen, the analyst reads a paragraph and decides whether there is relevant information in the paragraph. If this is the case, the analyst will attach a code to the text segment containing the information. The code can be either an existing one (one that was defined prior to the analysis or one that was derived from the text during the analysis) or a new one, which the analyst defines in order to represent the information. The code will also include a link to the text segment in order to make it retrievable.

4.4. Further analysis

The typical result of coding as indexing themes is a coded text (see diagram 2 for an example) and a structured index that represents the structure of raw data in the text(s) and supports the handling of text segments according to the occurrence or co-occurrence of codes. It is important to note that codes as the one depicted in diagrams 2 and 4 do not contain the information – they just indicate what kind of information can be found in a segment of text thus coded. In the example provided in diagram 5, IC-HIST indicates that a coded segment contains information about an innovation history. This is why codes can be worked at and analysed but cannot be used separately from the text. The analysis of codes (e.g. of frequencies of occurrences and co-occurrences of codes, of the networks of codes resulting from co-occurrences) is a useful step but cannot be the last word. The coded text, which is the relevant section of the original text, i.e. the selected raw data – must be interpreted at some stage.

The steps following the coding depend on the overall approach in which they are embedded. Within the context of the Grounded Theory approach, the analysis that follows the initial coding consists of comparing them, (empirically) generalising them and finding connections between them (see e.g. Turner 1981, Charmaz 2006). To that purpose, codes can be reorganized and reformulated by 'axial coding' and 'selective coding' which basically means that structures in the systems of codes are explored, codes are subsumed under more general codes, and the text is re-coded accordingly. In the version advocated by Strauss and Corbin (1990), these two steps are called 'open coding' and 'axial coding', and are intertwined first steps. They are followed by the so-called 'selective coding (ibid: 116-142), which further integrates the linked categories into a "story" (a "descriptive narrative

about the central phenomenon of the study"), which has a "story line" (the "conceptualization of the story", ibid: 116). These versions of Grounded Theory have in common that they are empirical generalisations, and that the only explanations that can be achieved this way are explanations of the empirically investigated phenomena.

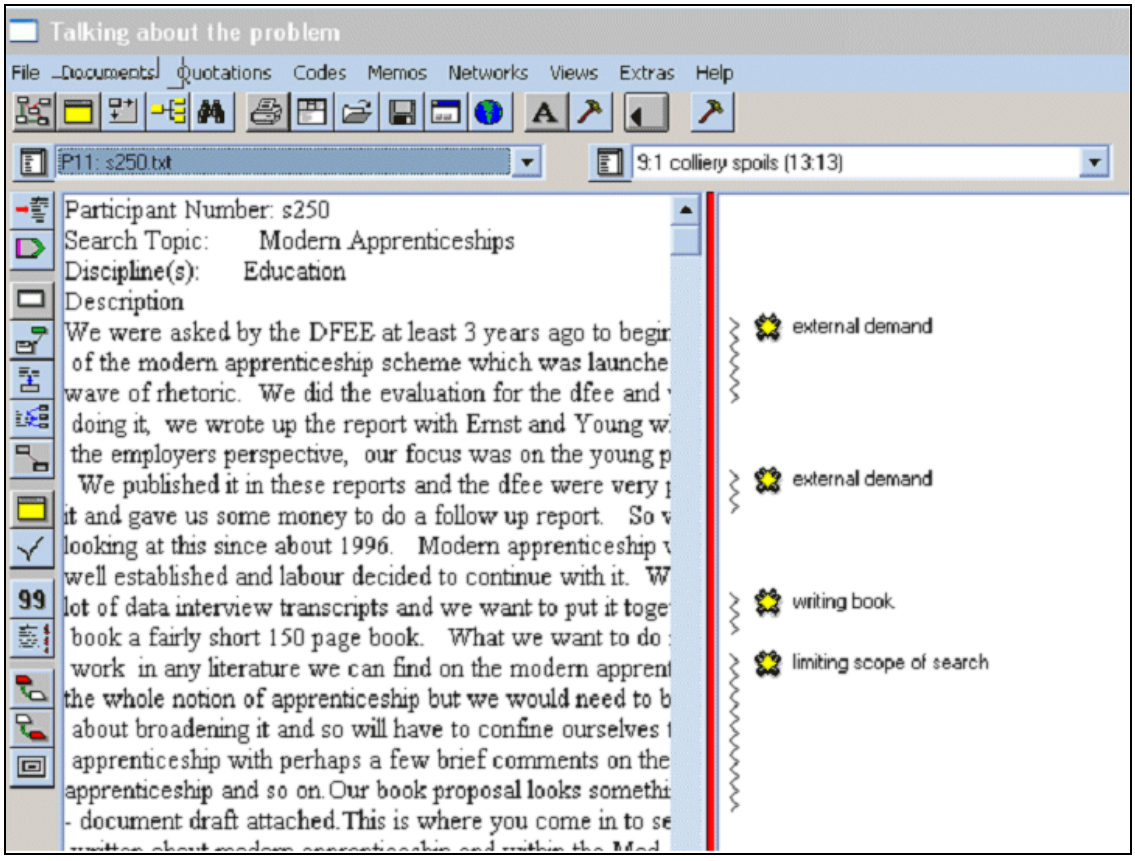


Diagram 5: Example of a coded text in the software package Atlas/ti (source: Wilson 2004)

There are only few suggestions for steps following the coding when it is used outside Grounded Theory. One strategy, which expresses some limitations of the coding technique, is to use the codes just as an index for a subsequent analysis by other means. This is what Miles and Huberman suggest:

> Codes are used to retrieve and organize the chunks [of text] mentioned earlier. The organizing part will entail some system for categorizing the various chunks, so the researcher can quickly find, pull out, and cluster the segments relating to a particular research question, hypothesis, construct, or theme. Clustering, and, as we will see, display of condensed chunks, then sets the stage for drawing conclusions. (Miles and Huberman 1994: 57)

Similarly, Kelle (1997, 5.9) states:

> Coding is the necessary prerequisite for a systematic comparison of text passages: text segments are retrieved and analyzed in order to

discover 'dimensions' which can be used as a basis for comparing different cases.

Thus, the most common strategy is apparently to selectively retrieve text, i.e. to collect all text segments that are tagged with the same code, and to compare these text segments. Miles and Huberman (1994) refer to this phase as "display data", which they consider to be an essential step in the search for patterns. They distinguish between "within-case displays" and "cross-case displays", and distinguish two families of data displays, namely matrices and networks (drawing lines between cods). Both approaches have in common that they essentially suggest combining different bits and pieces of information in various arrangements until patterns become visible. To use matrices for that purpose, data have to be cross-tabulated according to whatever criteria make sense for the study. The crucial steps are "creating the format" (which Miles and Huberman consider to be easy) and "data entry", which they describe as follows:

> "What takes time is data entry itself. Coded data segments have to be located in the transcribed field notes; they have to be extracted, condensed, and summarised. Some further data transformation and selection may also be involved: doing ratings, making judgements, picking representative quotes." (Miles and Huberman 1994: 98)

This is remarkable because Miles and Huberman suggest using codes as an index for the text and then extracting content from retrieved text segments for further analysis. Apart from the inefficiency of this approach – the extraction from text segments might have to be repeated with each new attempt at displaying data – it is also very similar to what we defined above as extracting content. This suggests the possibility to live without coding, which we will explore now.

5. Qualitative Content Analysis

Regardless of its popularity, coding procedures often lead to two problems that are not easily solved, namely an overload of codes and an overload of texts. The first problem often results from the step called 'open coding'. In this step, the researcher goes through all texts and indexes them, i.e. adds codes to text segments that signify the existence of important information in this segment. This bottom up – process of coding may easily lead to large numbers of codes that cannot anymore be memorised and handled (the "code swamp", Friese 2011: [19], see also Ezzy 2002: 125-126). The second problem is that this procedure reduces the amount of information only to a limited extent. While it enables the exclusion of irrelevant text segments (those that don't receive any code), coding leaves the relevant text segments unchanged, which means that they still contain irrelevant parts and are inefficiently worded. These text segments can thus add up to huge amounts of text. Coding content would avoid that problem. It is in fact very similar to the extraction of content and thus to qualitative content analysis. However, it does not seem to be widely used.

Since these negative aspects appear to be 'hardwired' in the coding procedure, it is worthwhile to look for alternatives. In this section, we present such an alternative, namely the extraction of information from the original text and its separate processing. Since the core idea of this method is to consciously leave the original text behind and to analyse the information extracted from it, it is best termed qualitative content analysis. The main difference between qualitative content analysis and other methods is that the former stepwise reduces the data material (Flick 2002: 190).

5.1. History and variations

Among the qualitative methods of data analysis, the qualitative content analysis is the only one that has direct roots in a quantitative method. Quantitative content analysis was originally developed as a tool for the analysis of the large numbers of similar texts in the emerging mass media. The quantitative content analysis (which is often just referred to as 'content analysis'[9]) subsumed the texts to a system of categories, counted the occurrences of categories and subjected the resulting numbers to a statistical analysis (Berelson 1952). This approach was soon criticised because it purposefully ignored the varying meanings of text elements (Kracauer 1952). Each categorisation of a text element reduces its meaning to the meaning that the category is supposed to measure. "The meat of the problem with content analysis (and its relatives) is ... how analysts usually simply trade off their tacit members' knowledge in coining and applying whatever cate-

[9] See for example Franzosi (2004), Silverman (2001: 123-124), and descriptions in dictionaries of sociology under the keyword "Content Analysis" (e.g. Abercrombie, Hill, and Turner 2000: 72; Jary and Jary 2000: 111; Roberts 2001; Scott and Marshall 2005: 110). Unfortunately, the term 'qualitative content analysis' is very ambiguous. Sometimes it includes every kind of qualitative text analysis, sometimes even coding, and sometimes quantitative (statistic-based) content analysis (Hsieh and Shannon 2005).

gories they do use." (Silverman 2001: 124). Therefore, attempts were made to produce more 'qualitative' variants of quantitative content analysis (e.g. Mostyn 1985; Morgan 1993).

The most consistent attempt of adopting content analysis to the requirements of qualitative data analysis has been undertaken by Mayring (1993; 2000). Mayring argues for combining the strengths of quantitative content analysis, namely its theory guided and rule guided approach to data, with the qualitative tradition of letting the actual content of data structure the analysis. He proposes a whole range of techniques that can be used to create a system of categories that is adapted to the content of the texts. Most of these techniques use a system of categories that are designed ex ante (derived from theory) and then test these categories and adapt them to the material in a pre-analysis of 30% to 50% of the material. The most extreme technique is fully inductive. Similar to the open coding, it derives the system of categories exclusively from the text that is to be analysed.

While Mayring's proposal seems to open up content analysis and to set up a qualitative procedure, his attempts to keep major steps of quantitative content analysis create important limitations. Regardless of the procedure that is applied to create categories, Mayring's qualitative content analysis leads to a step in which a 'final' system of categories is applied to the text, and occurrences of categories can be counted. He considers this as a strength because it enables the combination of qualitative and quantitative approaches, and because quality criteria of quantitative methods such as inter-coder-reliability can be applied. However, using such a fixed and rigid system of categories has some serious problems. Techniques that adapt a system of categories by applying it to part of the material in a 'pilot run' rest on the assumption that the other half of the material does not contain any surprises, which is unlikely. In particular, the assumption that we could know all possible values of a category, which is inscribed in the use of closed lists of categories' possible values, forces the analyst to subsume all empirical information to one of the categories and to one of their values in very much the same way as the quantitative content analysis does.

Thus, while Mayring uses the whole arsenal of qualitative methodology to refine the categories of his content analysis, he does so only to prepare a final step that is not open to the data anymore. We thought it necessary to overcome this last restraint produced by the quantitative tradition, and to use the basic idea of content analysis – to *extract* information from a text and to process this information independently of the text – in an approach that is both more open and does not contain zero-sum games between theory and data.

5.2. Categories for extraction

The core idea of our qualitative content analysis is to work with a system of categories that is derived from theoretical consideration but can be changed and extended during the whole process of data analysis in a way that en-

ables the preservation of theoretical considerations (Gläser and Laudel 2010). We think that any contradiction between theoretical considerations and data should not be resolved 'on the spot', i.e. when it occurs in one segment of the text, but at a later stage when a decision can be grounded in all data and theory.

Similar to Mayring, we prepare for our data analysis by deriving categories from a theoretical framework (which already has guided our data collection). The framework contains variables and assumptions about processes (systems of actions) that mediate the impact of variables (see the appendix for an example). The use of variables is a bit unusual in qualitative analyses and therefore merits a short discussion. Variables are commonly associated with quantitative social research, where they are understood as monodimensional, and where most variables have 'higher-level' scales that support mathematical operations beyond counting and comparing (ordinal, interval, ratio scales) The crucial steps for the appropriation of the variable concept by qualitative research include (a) emphasis on nominal scales and (b) emphasis on complex (multidimensional) variables. We see one major advantage of qualitative research in its ability to use the complex variables that are part of sociological theory without having to translate them into the one-dimensional indicators that can be processed by statistics. Many of the variables we use differ from variables of quantitative research in that they are multidimensional, i.e. have attributes that vary along different dimensions. For example, we can describe institutions by variables. If we regard institutions as systems of informal and formal rules (North 1990, Scharpf 1997), then they can be described by variables that contain the following dimensions:

- the subject of the rule (which actions should be influenced),
- the content of the rule (in which situation which action should occur),
- the scope of the rule (i.e. the actors whose actions should be regulated), and
- the character of the rule (whether it is a formal or an informal rule).

Each dimension of the variable 'institution' can assume different 'values' that can be described verbally (i.e. are not quantifiable). The variable cannot be disassembled into a set of one-dimensional variables because the character refers to a certain content, which in turn refers to a certain subject of the rule, and so on. Since a qualitative analysis is interested in the social mechanisms that mediate between variables, one important dimension of all variables is a time dimension which records the time or period for which values in the other dimensions have been found. Thus, while not all variables we deal with are multidimensional in the sense institutions are (for example we still may need basic information on actors such as gender, age, position in an organisation and so on.), each variable has at least one material and one time dimension.

The variables and mediation processes of our theoretical model inform the construction of categories for the qualitative content analysis. If the state of theory is poor and makes it difficult to derive variables then at least influencing factors (conditions of actions) can be used as a basis for the construction of categories. These factors can be derived from a general theory of action. This approach is similar to Strauss' and Corbin's 'coding paradigm' (Strauss and Corbin 1990: 96-113). However, contrary to Strauss and Corbin we see this approach as a last resort and recommend using the most specific theoretical offers available.

Categories for qualitative content analysis are multi-dimensional too and contain:

- the material dimensions of the variable/ properties of the process, which take in the 'values' reported in the text;
- a time dimension that picks up the point in time or period of time for which 'values' were reported; and
- causal 'dimensions' that are not dimensions of the variable or properties of the process but are used to take in instances of causation reported in the text (regardless of their coverage by the initial theoretical model, all reported instances of causation are collected).

Indicators for each category describe how statements belonging to the category are likely to look like and thus help to find the relevant information in the text. Diagram 6 provides an example of a category that is based on an institutional variable.

Since qualitative content analysis extracts information from a text for further use independently of the text, the extracted 'values' of categories must represent the data contained in the text as precisely and completely as possible. This is why the system of categories, their dimensions, possible 'values' cannot be exhaustively defined in advance. Whenever we encounter information that does not fit the categories, we construct new categories or new dimensions of categories. The only rule for this adaptation is not to abandon the original categories. The original system of categories can be supplemented by new categories and dimensions but should not be 'cleaned' by removing variables or dimensions. If theoretically derived categories do not fit a specific part of the data we keep them for extracting the data that fit them, and deal with the resulting contradictions after the extraction of information has been finished.

Similarly, while we might have a pretty good idea about some values of some dimensions prior to the data analysis (for example that some of the rules in a university will be formal and will apply to all academics), there are many others we cannot predict. This is why our categories use open lists of values, to which we can add new ones during the whole analysis.

Category: University rules of fund allocation

Definition of the underlying variable = Rules governing the allocation of funds and are not tied to evaluations or refer to specific sources of funding

Indicators:

- rules of fund allocation to faculties, institutes, centres and specific categories of staff such as early career researchers

- internal rules governing applications for external funding

Dimension	Some empirical instances that are already known
Time	Point in time or time-span for which the rule was reported
Subject of the rule	e.g. - internal distribution of the money allocated to the university through a funding formula - buildup and use of strategic funds, - distribution of research student scholarships,
Scope of the rule	e.g. university/ faculty/ school
Content of the rule	e.g. - Specific internal formulae used - Sources of strategic funds
Reported Causes	e.g. decisions at various levels of the university hierarchy
Reported Effects	e.g. sanctions, perceptions, other actions

Diagram 6: Example of a category in Qualitative Content analysis

5.3. Extracting information from a text

Before we begin the data analysis, we use our knowledge from the collection of the empirical data collection and check the model. Are the constructions of the variables and their dimensions appropriate? Are they appropriately defined? Which additional indicators exist that can help us to find facts of interest in our empirical material? This revision of our model is not done systematically on the empirical material, but using our impressions from the empirical analysis. Therefore, it is essential to expand the model rather than to narrow it down. Otherwise, important factors could be omitted from the analysis.

Extraction essentially means to identify relevant information, to identify the category to which the information belongs, to rephrase the information contained in the text as short concise statements about the value of each dimension, to assign these statements to the relevant dimensions of the category and to collect them separately from the text. A link to the original text is kept in order to enable clarifications if later necessary.

We usually apply qualitative content analysis to transcripts of semi-structured interviews in which the unit of analysis is a paragraph. Each paragraph gets a specific identifier that enables the identification of the source of information. We read the paragraph and decide whether it contains relevant information, and if so, to which category the information belongs. Then we extract the relevant information by formulating short de-

scriptive statements about the values in the dimensions (for an example see the appendix). As a result, each dimension contains either a single word or a phrase, in exceptional cases even a whole sentence. If a paragraph contains information about different values at different times, the extraction is repeated with the same category. If it contains information about more than one category, other categories are used for the extraction of information from the same paragraph. This way we extract information from each paragraph of each text and store it in one or more categories. Comments (about interpretations, marking contradictions etc.) can be added. Thus, the extraction is a process of constant interpretation. We must read a paragraph, interpret it and decide to which variable and to which dimension the information should be assigned and how we best summarise the information.

During this extraction process, unanticipated information – information that we didn't consider in our theoretical preparation of the study – is likely to be found. It is important that the categories neither suppress nor distort this information. We have four opportunities to include such unanticipated information:

- We can add new dimensions for already used variables, or
- We can include whole new variables.
- The values of the variables are not fixed before the extraction. Although some pre-defined values might exist, most of the values will only emerge during the analysis.
- In the causal dimensions, all influences appearing in the empirical material can be stored if they seem to be relevant for answering the research question. This way, influences which we did not foresee in the theoretical considerations are included.

The outcome of the extraction is an extensive structured raw material which contains all information of the empirical material about the values of variables and causal relationships of variables. All subsequent analyses use this material. We only go back to the original text if errors or doubts occur in the subsequent analysis.

5.4. Processing the extracted data

The extracted raw data can now be processed in order to further consolidate our information base. The aim of this step is to improve the quality of the data by summarising scattered information, remove redundancies and correct errors. This again reduces the amount of the material, and enables a further structuration of the data.

To make this step reversible, we archive the original extraction files and use copies. The concrete techniques of the data processing depend on the aim of the investigation and the type of variable. For example, information can be structured by chronological order or by subject matter. Beyond these specific procedures, the following general steps can be described:

1) Scattered information is summarised. Information about the same values of a variable at the same time are often scattered over different interviews and extraction tables. Sorting brings such information together and allows summarising them in the extraction table.
2) Information with the same meaning is aggregated.
3) Obvious errors are corrected. Contradictory information can sometimes be corrected, using the interview text. In other cases, the contradiction must remain and be marked as such.
4) Different information is kept.

The summarising can be done in several steps and each step must be documented. Variables can also be summarised more than once in different ways. While summarising information, we always keep the identifier. This allows us to jump back to the original text and to reproduce single steps.

The outcome of this step is an information base that is structured by both theoretical considerations and structures of empirical information, is largely free of redundancies, and contains the relevant empirical information in its shortest possible form.

5.5. Further analysis

Qualitative content analysis is similar to coding in that it does not contain any techniques for pattern recognition or pattern integration. Both coding and qualitative content analysis produce an information base, which must be further analysed in order to answer the research question. We have described the general strategies in sections 2 and 3. The techniques suggested by Miles and Huberman (1994) can be applied to the data base produced by qualitative content analysis, too. Since this data base consists of tables containing the information about all categories, it is better suited to building matrices than networks, and lends itself more easily to building typologies. Three techniques that have turned out to be very valuable in the search for patterns by creating typologies are a) just sorting the tables according to different criteria, b) reducing the tables by either omitting all columns that contain information that is not to be used for building a typology or by further standardising the descriptions of empirical information, and c) reorganizing tables by combining rows or columns from several categories.

6. Comparing the two approaches

Coding and qualitative content analysis have roughly the same function in qualitative data analysis (diagram 7). Both methods help us to locate relevant information in the texts that contain our data, i.e. to distinguish raw data from 'noise'. Both methods produce an information base for the analysis and interpretation of the data. They are also similar in that they both fulfil an important requirement of qualitative data analysis, namely that equal weight is given to all information in a text. Since both methods require the

researcher to read and interpret all texts that might contain data, they avoid the fallacy of 'subconscious interpretation', where researchers form an opinion on their data by reading part of it and then interpret the rest of their material selectively with a frame that has formed during the first reading (Hopf 1982: 315-316).

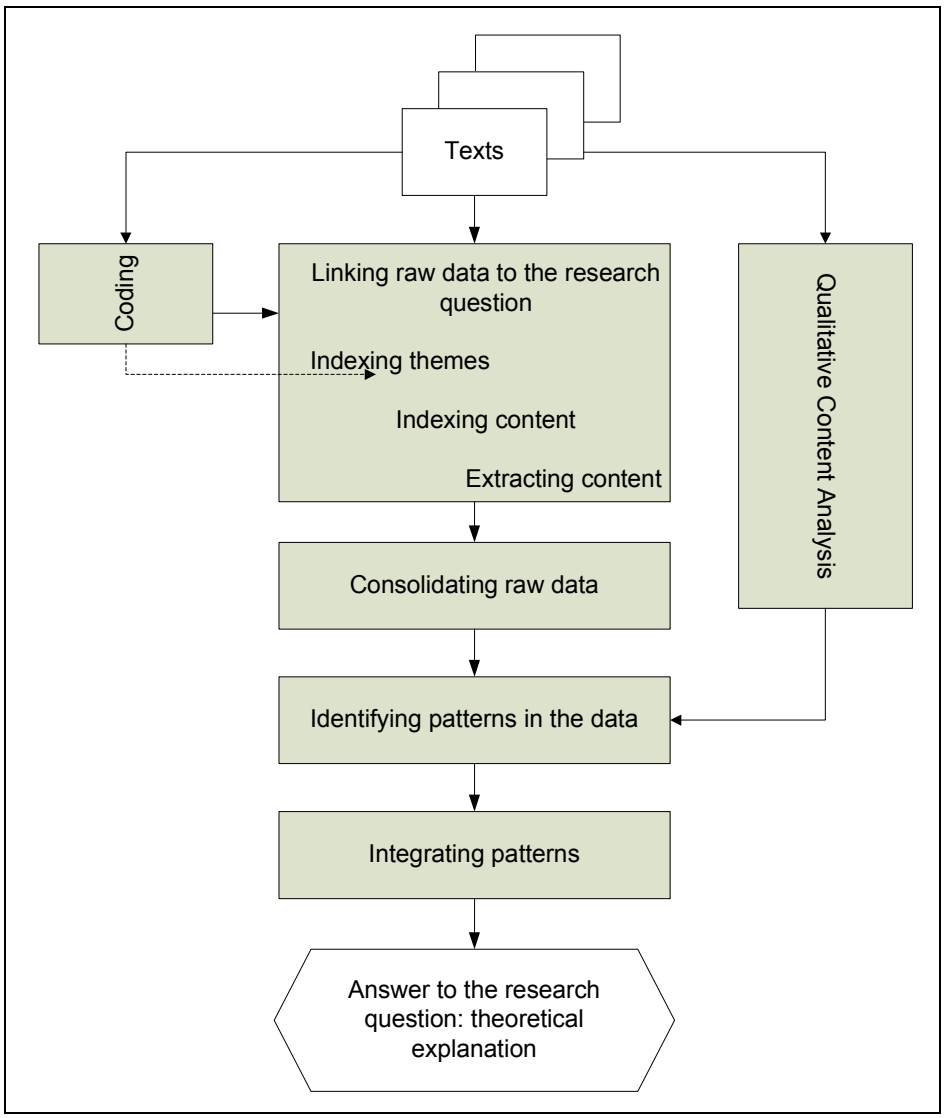


Diagram 7: Positioning coding and qualitative content analysis in the steps between texts and explanation

The two methods are also similar in their adherence to the main principles of qualitative social research. They both enable a theory-guided and rule-based approach, and are both open to unanticipated information. However, there is also a difference between the methods in this respect. While a theory-guided approach to coding is in principal possible, it is still actively discouraged within Grounded Theory, poorly supported by non-Grounded Theory methodological frameworks, and by no means enforced by any of those. As a result, coding can also be conducted 'theory-free', which is an advantage in some types of investigations (see below) but also a temptation to abandon theory where it could be used. This is different for qualitative content analysis, which cannot start without categories derived *ex ante* from

theoretical considerations. Qualitative content analysis is thus liable to the opposite problems: It is less suitable for investigations that cannot build on prior theory and offers the temptation to force concepts on the data. We would like to emphasize, however, that the temptations offered by the methods are temptations to do sub-standard research.

The major difference between the two methods is the kind of information base that is produced by coding respectively qualitative content analysis. The information bases significantly differ from each other and thus provide different opportunities for data analysis. Coding leads to an indexed text which can be reorganised in many ways in the process of data analysis. Since the codes signify the existence of information rather than containing all the information, data analysis in coding-based procedures means working with the codes and the text. This is very similar to searching through a book. While the book's index provides interesting information about the book's content and its structure, one has to go to the page that is recorded in the index in order to use the information whose existence is indexed.

Qualitative content analysis avoids the problems that accompany this indexing strategy – a possibly large and overly complicated system of codes, an only slight reduction of the initial material – by extracting all relevant information and processing it separately.[10] In the first step – the extraction – we replace the respondents' descriptions of relevant information by our reformulation in an analytic language, which is shorter, more concise, and better adapted to our research interest. We thus achieve a significant reduction of the material in the first step, and more reductions in the subsequent steps.

This comparison indicates the different areas of applicability of the two methods. Qualitative content analysis is rather demanding in that it requires a clear understanding of the data we need from our texts prior to the analysis. This means that investigations that are purely descriptive or highly explorative might fare better with a coding procedure, particularly with open coding. Another difference between the applicability of the two methods is linked to the role of the original text. The application of qualitative content analysis presupposes that it is only important *what* was said, not *how* it was said. If at any later stage of the investigation the actual phrasing of a significant part of the information is important, coding and retrieval procedures are better suited because they offer better access to the original text. Similarly, when the relative position of information in the text (the co-occurrences of certain kinds of information) is important, indexed texts offers easier access to this aspect of the data and coding should be preferred.

Even though qualitative content analysis supports the analysis of the phrasing and the position of information in principle by keeping links to the origi-

[10] Qualitative content analysis is actually the only method of qualitative data analysis which begins by separating the data from the original text, systematically reduces the amount of information, and structures it according to the aim of the investigation. While a link is kept that enables the return to the source of the information, in qualitative content analysis this is considered an exception rather than a strategy.

nal texts, accessing this information is more awkward than indexed texts. However, whenever only the content of data matters, and particularly when large amounts of texts need to be processed, the systematic and drastic reduction of the material that can be achieved by qualitative content analysis makes the latter the method of choice.

References

Abercrombie, Nicholas, Stephen Hill, and Bryan Turner, 2000. *The Penguin Dictionary of Sociology*. London: Penguin.

Berelson, Bernard, 1952. *Content Analysis in Communication Research*. Glencoe, Ill.: Free Press.

Bernard, Russel H. and Gery W. Ryan, 2010. *Analyzing Qualitative Data: Systematic Approaches*. London: Sage.

Blatter, Joachim and Till Blume, 2008. Co-Variation and Causal Process Tracing Revisited: Clarifying New Directions for Causal Inference and Generalization in Case Study Methodology. *Qualitative Methods*, Spring 2008, 29-34.

Boeije, Hennie, 2010. *Analysis in Qualitative Research*. London: Sage.

Charmaz, Kathy, 2006. *Constructing Grounded Theory*. Los Angeles: Sage.

Coffey, Amanda, Beverly Holbrook, and Paul Atkinson, 1996. Qualitative Data Analysis: Technologies and Representations. *Sociological Research Online* 1.

Cortazzi, Martin, 2001. Narrative Analysis in Ethnography. Paul Atkinson, Amanda Coffey, Sara Delamont, John Lofland and Lyn Lofland (eds.), *Handbook of Ethnography*. London: Sage, 384-394.

Ezzy, Douglas, 2002. *Qualitative Analysis: Practice and Innovation*. Sydney: Allen & Unwin.

Flick, Uwe, 2002. *An Introduction to Qualitative Research*. London: Sage.

Franzosi, Roberto, 2004. Content Analysis. Melissa A. Hardy and Alan Bryman (eds.), *Handbook of Data Analysis*. London: Sage,

Friese, S. (2011). Using ATLAS.ti for Analyzing the Financial Crisis Data. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 12*(1). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1632/3150

Glaser, Barney G., 1978. *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory.* Mill Valley: The Sociology Press.

Glaser, Barney G., 1992. *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.

Glaser, Barney G., 2002. Conceptualization: On theory and theorizing using grounded theory. *International Journal of Qualitative Methods* 1.

George, Alexander L. and Andrew Bennett, 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Glaser, Barney G. and Anselm L. Strauss, 1967. *The Discovery of Grounded Theory : Strategies for Qualitative Research*. Chicago: Aldine.

Glaser, Barney G., 1992. *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.

Gläser, Jochen, and Grit Laudel, 2010. *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*. 4. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.

Hopf, Christel, 1982. Norm und Interpretation. Einige methodische und theoretische Probleme der Erhebung und Analyse subjektiver Interpretationen in qualitativen Untersuchungen. *Zeitschrift für Soziologie* 11: 307-329.

Hsieh, Hsiu-Fang, and Sarah E. Shannon, 2005. Three approaches to qualitative content analysis. *Qualitative Health Research* 15: 1277-1288.

Jary, David, and Julia Jary, 2000. *Collins Sociology Dictionary*. London: HarperCollins.

Kelle, Udo, 1997. Theory Building in Qualitative Research and Computer Programs for the Management of Textual Data. *Sociological Research Online* 2.

Kelle, Udo, 2003. Die Entwicklung kausaler Hypothesen in der qualitativen Sozialforschung: Methodologische Überlegungen zu einem häufig vernachlässigten Aspekt qualitativer Theorie- und Typenbildung. *Zentralblatt für Didaktik der Mathematik* 35: 232-246.

Kelle, Udo, 2005. "Emergence" vs. "Forcing" of Empirical Data? A Crucial Problem of "Grounded Theory" Reconsidered. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-line Journal]* Volume 6: Art. 27.

Kracauer, Siegfried, 1952. The challenge of qualitative content analysis. *Public Opinion Quarterly* 16: 631-642.

Lieberson, Stanley, 2004. Comments on the Use and Utility of QCA. *Qualitative Methods* 2, (2), 13-14.

MacMillan, Katie, and Thomas Koenig, 2004. The wow factor - Preconceptions and expectations for data analysis software in qualitative research. *Social Science Computer Review* 22: 179-186.

Mahoney, James, 2003. Tentative Answers to Questions about Causal Mechanims. Paper presented at the annual meeting of the American Political Science Association, Philadelphia, PA, August 28, 2003. http://www.allacademic.com/meta/p_mla_apa_research_citation/0/6/2/7/6/p62766_index.html (accessed 13 February 2011).

Mahoney, James, 2008. Toward a Unified Theory of Causality. *Comparative Political Studies* 41, 412-436.

Maxwell, Joseph A., 2004a. Causal Explanation, Qualitative Research, and Scientific Inquiry in Education. *Educational Researcher* 33 (2), 3–11.

Maxwell, Joseph A., 2004b. Using Qualitative Methods for Causal Explanation. *Field Methods* 16 (3), 243–264.

Mayntz, Renate, 2004. Mechanisms in the Analysis of Social Macro-Phenomena, *Philosophy of the Social Sciences*, 34, 237-259.

Mayring, Philipp, 1993. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim: Deutscher Studien Verlag.

Mayring, Philipp, 2000. Qualitative Content Analysis. Forum Qualitative Sozialforschung / *Forum: Qualitative Social Research* [On-line Journal] 1(2).

Merton, Robert K., 1968a. On Sociological Theories of the Middle Range. in Robert K. Merton (ed.) *Social Theory and Social Structure*, London: The Free Press, pp. 39-72.

Merton, R. K., 1968b. The Self-Fulfilling Strategy, in: R. K. Merton, *Social Theory and Social Structure*, London: The Free Press, 475-490.

Miles, Matthew B., and A. Michael Huberman, 1984. *Qualitative Data Analysis : A Sourcebook of New Methods*. Beverly Hills: Sage.

Miles, Matthew B., and A. Michael Huberman, 1994. *Qualitative Data Analysis : An Expanded Sourcebook, Second Edition*. Beverly Hills: Sage.

Morgan, David L., 1993. Qualitative Content Analysis: A Guide to Paths Not Taken. *Qualitative Health Research* 3: 112-121.

Mostyn, Barbara, 1985. The Content Analysis of Qualitative Research Data: A Dynamic Approach. Michael Brenner, Jennifer Brown and David Canter (eds.), *The Research Interview: Uses and Approaches*. New York: Academic Press, 115-145.

North, Douglass C., 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Patton, Michael Quinn, 1990. *Qualitative Evaluation and Research Methods*. Newbury Park: Sage.

Ragin, Charles C., 1987. *The Comparative Method. Moving Beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California Press.

Ragin, Charles C., 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. and Benoît Rihoux, 2004: Qualitative Comparative Analysis (QCA): State of the Art and Prospects. *Qualitative Methods* 2 (2), 3-13.

Roberts, Carl W., 2001. Content Analysis. Neil J. Smelser and Paul B. Baltes (eds.), *International Encyclopedia of the Social & Behavioral Sciences*. Oxford: Elsevier, 2698-2702.

Scharpf, Fritz W., 1997. *Games Real Actors Play. Actor-Centered Institutionalism in Policy Research*. Boulder: Westview Press.

Scott, John, and Gordon Marshall, 2005. *Oxford Dictionary of Sociology*. Oxford: Oxford University Press.

Silverman, David, 2001. *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*. London: Sage.

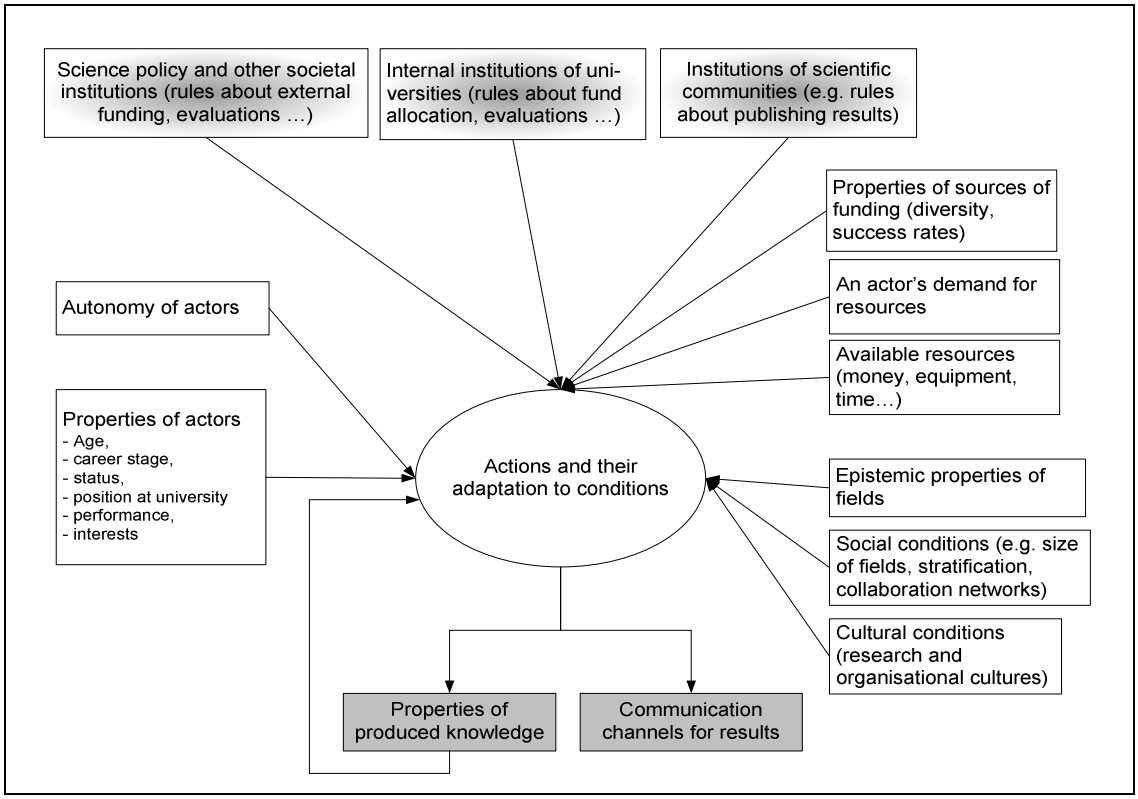
Strauss, Anselm, and Juliet Corbin, 1990. *Basics of Qualitative Research. Grounded Theory Procedures and Techniques*. Newbury Park: Sage.

Turner, Barry A., 1981. Some Practical Aspects of Qualitative Data Analysis: One way of organising the cognitive processes associated with the generation of grounded theory. *Quality and Quantity* 15, 225-247.

Wilson, T.D., 2004. Talking about the problem: a content analysis of presearch interviews. *Information Research* 10 (1), http://informationr.net/ir/10-1/paper206.html (accessed 2 December 2011).

Appendix: Demonstration of Qualitative Content Analysis

In this project we addressed the following research problem: How does the performance-based block funding of Australian universities affect the content of research? We conducted semi-structured interviews with academics and university managers, which were fully transcribed. Before starting the actual extraction process, we constructed 12 multidimensional categories.



The following example of an extraction uses a section of an interview with a university manager, the head of the School of Biosciences. The identifier at the beginning of the paragraph contains the name of the interview file and a paragraph number.

> ...
>
> U2BioH-60#Yes. And to a large extent that's how everything runs. There are small sums of money that - I control the budget 100 per cent. I can do anything provided I pay for the salaries. So there are small sums of money that I will occasionally use strategically and my colleagues hate it when I do it. They shout and scream, but I will listen to people talking, I'll discuss it with one or two key individuals, I won't necessarily consult widely, and then say, okay, we're going to support this for six months conditional or three months conditional on something happening. But that's actually relatively rare.

We read the paragraph and decide that it contains information about an institutional rule, namely a funding rule of the academic unit. Therefore we store the relevant information in the variable "funding rules of the university", structured by the dimensions of the variable.[11]

time	**scope of the rule**	**subject of the rule**	**content of the rule**	**causes**	**effects**	**source**
time of the interview	School of Biosciences	budget	Is completely controlled by Head of School		Head uses small sums to strategically support research	U2BioH-60

Diagram x1: Extraction table for the variable "funding rules of the university"; extraction of one paragraph

This is done with all categories for each paragraph of each interview transcript. As a result, each category has a table with many rows, each of which contains one set of values (diagram x2). If a paragraph contains information for more than one category, the extraction procedure is undertaken for each category for which relevant information exists.

[11] The extraction is supported by a computer software that we developed ourselves. It is integrated into Microsoft Word. The reason for not using one of the numerous commercially available programs is that most of them are constructed to support the coding of text rather than extracting text (Coffey, Holbrook, and Atkinson 1996); (MacMillan and Koenig 2004): 182-183). Using them for methods based on extracting information turned out to be impossible or very difficult.

time	scope of the rule	subject of the rule	content of the rule	causes	effects	source
time of the interview	University	strategic funds	to match infrastructure grants			U2ResOff-73
time of the interview	University	strategic funds	to match external infrastructure funding, in certain cases of need			U2ResOff-80
time of the interview	School of Biosciences	budget	Is completely controlled by Head of School		Head uses small sums to strategically support research	U2BioH-60
until last year	Faculty of Science	distribution of funds from DEST formula	distributed money pretty much as earned to faculties		got 70% from teaching, 30% from research	U2-BioH-73
since 2001	university	strategic funds	for major centres of excellence, major equipment	easier to do from above, direct to areas of research strength		U2DVC-31

Diagram x2: Extraction table for the variable "funding rules of the university"; extraction for all paragraphs (extract)

In the next step, the extracted information is sorted and summarised. In our example, we merged rows containing identical information and sorted the variable "funding rules of the university" firstly by the scope of the rule, secondly by the subject of the rule, thirdly by the time at which the rule applies, and fourthly by the content.

time	**scope of the rule**	**subject of the rule**	**content of the rule**	**causes**	**effects**	**source**
time of the interview	School of Biosciences	budget	Is completely controlled by Head of School		Head uses small sums to strategically support research	U2BioH-60
until last year	Faculty of Science	distribution of funds from DEST formula	distributed money pretty much as earned to faculties		got 70% from teaching, 30% from research	U2-BioH-73
since 2001	University	strategic funds	to match infrastructure grants, major centres of excellence, major equipment, in certain cases of need	easier to do from above, direct to areas of research strength		U2ResOff-73 U2ResOff-80 U2DVC-31

Diagram x3: Extraction table for the variable "funding rules of the university" sorted and summarised (extract)

This table (together with the tables of the other variables) is subject to further analyses and a basis for the interpretation of the data.

Bisher veröffentlichte discussion paper des ZTG

Die discussion paper werden von Martina Schäfer, Hans-Liudger Dienel, Leon Hempel und Dorothee Keppler herausgegeben. Sie sind als pdf-Datei abrufbar unter www.ztg.tu-berlin.de.

Nr. 01/02 Susanne Schön, Dorothee Keppler, Brigitte Geißel: Gender und Nachhaltigkeit. Sondierung eines unübersichtlichen Forschungsfeldes.

Nr. 02/02 Alexander Peine, Rainer Haase, Hans-Liudger Dienel: Prozessevaluation – Evaluation der interdisziplinären Zusammenarbeit im Rahmen der Forschergruppe Sentha.

Nr. 03/02 Martina Schäfer: Kundenvielfalt erfordert Marktvielfalt – Eine Untersuchung der Potenziale von vier verschiedenen Bioeinkaufsformen.

Nr. 04/02 Heike Walk: Global Governance – Hoffnung und Ernüchterung in der internationalen Klimapolitik.

Nr. 05/03 Susanne Schön: Co-operation Management as a Part of Engineering Education.

Nr. 06/03 Leon Hempel, Eric Töpfer: On the Threshold to Urban Panopticon? Objectives and Results of the "Urbaneye" Project on the employment of CCTV in Europe.

Nr. 07/03 Dörte Ohlhorst: Der Weg ist das Ziel... Radfernwanderwege als nachhaltige Verknüpfung kontrastreicher Regionen.

Nr. 08/03 M. Schophaus, H. L. Dienel, C. F. von Braun: Von Brücken und Einbahnstraßen. Aufgaben für das Kooperationsmanagement interdisziplinärer Forschung.

Nr. 09/03 Leon Hempel, Hans-Liudger Dienel: Tele City Vision –Perceptions of ICT and its Impacts on City Competition.

Nr. 10/03 Martina Schäfer, Benjamin Nölting, Lydia Illge: Zukunftsfähiger Wohlstand. Analyserahmen zur Evaluation nachhaltiger Wohlstandseffekte einer regionalen Branche.

Nr. 11/04 Gabriele Wendorf, Doris Felbinger, Bettina Graf, Sabine Gruner, Helga Jonuschat, Olaf Saphöster: Von den Qualitäten des Wohnumfeldes zur Lebensqualität? Das Konzept des „Atmosphärischen" als Ausgangspunkt einer integrierten Theorie.

Nr. 12/04 Susanne Schön, Benjamin Nölting, Martin Meister: Konstellationsanalyse. Ein interdisziplinäres Brückenkonzept für die Technik-, Nachhaltigkeits- und Innovationsforschung.

Nr. 13/04 Jörg Potthast, Hans-Liudger Dienel: „Die Zeiten des natürlichen Fortschritts sind vorbei." Verkehrssicherheit als Gegenstand von Forschung und Politik. Vertiefung der SMARTBENCH-Teilstudie Frankreich.

Nr. 14/04 Achim Brunnengräber, Kristina Dietz, Bernd Hirschl, Heike Walk: Interdisziplinarität in der Governance-Forschung.

Nr. 15/05 Elke Baranek, Corinna Fischer, Heike Walk: Partizipation und Nachhaltigkeit. Reflektionen über Zusammenhänge und Vereinbarkeiten.

Nr. 16/05 Dorothee Keppler: Nachhaltigkeitskompetenzen. Zur Bedeutung geschlechtsspezifischer Kompetenzunterschiede für eine nachhaltige Technikentwicklung.

Nr. 17/05 Tina Boeckmann, Pamela Dorsch, Frauke Hoffmann, Dörte Ohlhorst, Ulrike Schumacher, Julia Wulff: Zwischen Theorie und Praxis. Anregungen zur Gestaltung von Wissenschafts-Praxis-Kooperationen in der Nachhaltigkeitsforschung.

Nr. 18/05 Benjamin Nölting, Tina Boeckmann: Struktur der Land- und Ernährungswirtschaft in Brandenburg und Berlin – Anknüpfungspunkte für eine nachhaltige Regionalentwicklung.

Nr. 19/05 Hans-Liudger Dienel: Grupy nieprofesjonalnych planistów i opinie mieszkańców. Nowa metoda uczestnictwa mieszkańców na przykładzie opracowania projektu dla jednej dzielnicy Berlina (Übersetzung Bürgergutachen „Zukunft Sparrplatz" der Senatsverwaltung für Stadtentwicklung Berlin 2001).

Nr. 20/05 Adina Herde: Kriterien für eine nachhaltige Ernährung auf Konsumentenebene.

Nr. 21/05 Christin Wemheurer, Jens Eitmann: Coaching in der ökologischen Landwirtschaft.

Nr. 22/05 Dorothee Keppler: Zur Evaluierung integrativer Arbeitsmarktkonzepte für Menschen mit Benachteiligungen.

Nr. 23/06 Benjamin Nölting: Die Politik der Europäischen Union für den ländlichen Raum. Die ELER-Verordnung, nachhaltige ländliche Entwicklung und die ökologische Land- und Ernährungswirtschaft.

Nr. 24/06 Dorothee Keppler, Eric Töpfer: Die Akzeptanz und Nutzung erneuerbarer Energien in der "Energieregion" Lausitz.

Nr. 25/07 Benjamin Nölting, Dorothee Keppler, Birgit Böhm: Ostdeutschlandforschung trifft Nachhaltigkeitsforschung - fruchtbare Spannungsfelder für die Entwicklung neuer Perspektiven in Ostdeutschland.

Nr. 26/08 Dorothee Keppler: "Das persönliche Engagement derer, die hier sind, das ist doch das eigentlich Wertvolle". Die Bürgerausstellung als Forum für die Stimmen von BürgerInnen zur Zukunft der Energieregion Lausitz.

Nr. 27/08 Benjamin Nölting: Social-ecological research for sustainable agriculture and nutrition.

Nr. 28/08 Christine Dissmann, Nina Gribat, Benjamin Nölting: Bilder des Wandels – Wandel der Bilder. Analysen zu Ostdeutschland.

Nr. 29/09 Leon Hempel, Michael Carius & Carla Ilten: Exchange of information and data between law enforcement agencies within the European Union.

Nr. 30/09 Benjamin Nölting, Silke Reimann, Carola Strassner: Bio-Schulverpflegung in Deutschland Ein erster Überblick