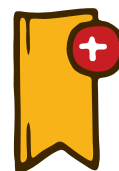
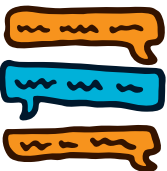




# Qualitative Content Analysis in Practice



Margrit Schreier



# Qualitative Content Analysis in Practice

# Qualitative Content Analysis in Practice



Margrit Schreier



Los Angeles | London | New Delhi  
Singapore | Washington DC

# Qualitative Content Analysis in Practice

**SAGE** has been part of the global academic community since 1965, supporting high quality research and learning that transforms society and our understanding of individuals, groups and cultures. SAGE is the independent, innovative, natural home for authors, editors and societies who share our commitment and passion for the social sciences.

Find out more at: [www.sagepublications.com](http://www.sagepublications.com)



# Qualitative Content Analysis in Practice



Margrit Schreier



Los Angeles | London | New Delhi  
Singapore | Washington DC

© Margrit Schreier 2012

First published 2012

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

SAGE Publications Ltd  
1 Oliver's Yard  
55 City Road  
London EC1Y 1SP

SAGE Publications Inc.  
2455 Teller Road  
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd  
B 1/I 1 Mohan Cooperative Industrial Area  
Mathura Road  
New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd  
3 Church Street  
#10-04 Samsung Hub  
Singapore 049483

**Library of Congress Control Number: 2011920596**

**British Library Cataloguing in Publication data**

A catalogue record for this book is available from the British Library

ISBN 978-1-84920-592-4

ISBN 978-1-84920-593-1

Typeset by C&M Digital (P) Ltd, Chennai, India  
Printed and bound by CPI Group (UK) Ltd, Croydon, CR0 4YY  
Printed on paper from sustainable resources



## CONTENTS

About the author	vi
Preface	vii
1 Introduction: What is qualitative content analysis?	1
2 What is qualitative about qualitative content analysis?	20
3 Will qualitative content analysis work for me? Decision aids	37
4 The coding frame	58
5 Building a coding frame	80
6 Strategies for building a data-driven coding frame	107
7 Segmentation and units of coding	126
8 Trying it out: The pilot phase	146
9 What to keep in mind: Evaluating your coding frame	166
10 The main analysis phase	194
11 How to present your results	219
12 Using software in doing qualitative content analysis	241
References	260
Index	269

## ABOUT THE AUTHOR

Margrit Schreier is Professor of Empirical Methods at Jacobs University Bremen. She started out with a BA in English Language and Literature (at New College, Oxford), before she went on to study Psychology at Heidelberg University (Dipl. Psych., Dr.) where she first became interested in qualitative research methods. She was also co-director of the Methods Center of the Bremen International Graduate School of Social Sciences (BIGSSS) and Editor-in-Chief of the *Journal of Media Psychology*. She has taught research methods courses for 15 years, including qualitative methods, experimental design, and mixed methods, to students from a variety of social science disciplines, and has given workshops on qualitative content analysis and other methods topics at GESIS, the Leibniz Institute for the Social Sciences, for more than five years. Her other research interests include the didactics of teaching qualitative research methods, media reception, the empirical study of literature, and health-related research. She has been a principal investigator in several DFG-funded research projects on these topics, and she has authored and co-authored an introductory methods textbook and more than 90 book chapters and articles. When she is not working, she enjoys cooking and photography. She is also a classical homeopath and has been surprised to find that homeopathy allows her to apply many of her methods skills – including qualitative content analysis!

## PREFACE

It has been a long time since my first encounter with qualitative content analysis (QCA) in the early 1990s as a student at Heidelberg University. Little did I anticipate at the time that I would one day teach the method, even less that I would write a textbook on the topic... Back then I was simply intrigued by what the method had to offer: a very systematic way of making sense of the large amount of material that would invariably emerge in the process of doing qualitative research.

As I began to teach qualitative research methods, including QCA, it quickly became clear that there was a need for a textbook on the topic. When I was talking to my German students, they obviously didn't know about the rich history of the method. When I was talking to my international students, chances were that they had never even heard of it. When I mentioned it to my students from the humanities, they would take it to mean a rich description of a work of art. When I mentioned it to my students from communication studies, they would say: Qualitative content analysis? But content analysis is a quantitative method, right?

At this point, I might simply have let things rest. Why try to convince students of the merits of a method they were barely familiar with? But alongside these confused looks, mostly from undergraduate students, I was increasingly approached by graduate students who wanted to use QCA in their work – but didn't know how exactly to go about it. I began to teach university courses and workshops on QCA on a regular basis, and it is above all from the studies and questions of the participants in these classes that this book has grown. There exist quite a few textbooks on quantitative content analysis, down to the details of how to actually go about it – but these focus on the quantitative version of the method. There also exist textbooks on QCA, but they tend to focus on the different traditions and concepts in QCA and less on the research process (and, sadly, one major exception was quickly sold out and never saw a second edition). In each class, students were asking me similar 'how to' questions. At first, I kept on thinking: If there only was a textbook I could use! And gradually this turned into: I must get down to it and write that textbook!

So this is what this book is meant to be: a hands-on textbook, guiding students and other researchers through the process of conducting their own

QCA. It contains only a little about the history and the methodological background of the method because a lot has already been written about this. Instead, I focus on describing how QCA fits into the broader landscape of qualitative research, and I then move on to what makes up the major part of the book: a detailed description of the steps involved in QCA, illustrated by examples that are taken from published papers by other authors as well as from my own work and the work of several students. Many of these examples are originally in German; all translations from German into English are my own. An additional chapter containing more detailed descriptions of these examples can be accessed on the Sage website: [www.sagepub.co.uk/schreier](http://www.sagepub.co.uk/schreier)

I would like to thank Norbert Groeben who first introduced me to QCA; Ruth Rustemeyer whose book on QCA has inspired me over the years; Guenter Mey and Sascha Schroeder who have kept on asking me when I will finally get down to it and write my own; Özen Odağ for her steady reassurance over the years of working together that this book is worthwhile; Jürgen Hoffmeyer-Zlotnik and Patricia Lueder from GESIS as well as Guenter Mey (again) and Katja Mruck from [qualitative-research.net](http://qualitative-research.net) for inviting me to teach QCA; Claudia Gronewold, Simone Heil, Marie-Luise Herrmann, Özen Odağ (again), Marina Otten, and Katrin Viertel for letting me draw upon and quote extensively from their research; Adele Diederich and Jeannette Winkelhage for letting me draw on our joint research and working papers on prioritising in medicine; Immacolata Amodeo and Petra Lietz for their moral support; Patrick Brindle from Sage for his enthusiasm in getting the book project off the ground; Katie Metzler from Sage for her editorial suggestions, which have transformed a rather stuffy academic text into what is now hopefully a readable textbook; Irina Chiaburu for her tireless help with databases, files, and photocopying; Meropi Tzanetakis and several anonymous reviewers for their feedback on the manuscript in various draft stages; Jacobs University Bremen for granting me a sabbatical to write the book; and finally and above all, the many students and workshop participants whose questions over the years have inspired this book. I hope you will find it useful!

*Margrit Schreier  
Bremen  
June 2011*

# 1

## INTRODUCTION: WHAT IS QUALITATIVE CONTENT ANALYSIS?

### Chapter guide

You are a qualitative researcher, and probably you have already collected part of your data. Now you want to know: What does it all mean? There are many qualitative methods out there for analysing your data and interpreting its meaning, and qualitative content analysis (QCA) is one of them. This first chapter will give you a basic idea of the method and what it involves, before going into more detail in the following chapters. More specifically, we will look at:

- some basic features of QCA;
- the origin of quantitative content analysis in the social sciences;
- how QCA emerged from quantitative content analysis.

The chapter will conclude with an overview of the book.

### Some basic features of QCA

QCA is a method for describing the meaning of qualitative material in a systematic way. You do this by assigning successive parts of your material to the categories of your coding frame. This frame is at the heart of QCA, and it covers all those meanings that feature in the description and interpretation of your material.

---

#### KEY POINT

QCA is a method for systematically describing the meaning of qualitative material. It is done by classifying material as instances of the categories of a coding frame.

---

In the following, we will look in more detail at:

- the kind of material to which you can apply QCA;
- the goals of QCA;
- how QCA is done.

## What material is suitable for QCA?

In most general terms, QCA will be an option if you have to engage in some degree of interpretation to arrive at the meaning of your data. In a way, this is almost a tautology. Data never ‘speaks for itself’, it does not ‘have’ a specific meaning. Meaning is something that we, the recipients, attribute to the words that we hear or read, to the images that we see. This is a complex process in which we bring together our perception of the material with our own individual background: what we know about a topic, the situation in which we encounter it, how we feel at the time, and much more. Meaning is not a given, but we *construct* meaning. The assumption that meaning is not something that is inherent in a text, that the recipients take an active part in constructing meaning, was first put forward by Fredric Bartlett, a psychologist, as early as 1932. Ever since, it has become a staple of theory and research on reading and text comprehension (see Goldman, Graesser & van den Broek, 1999, on reading, processing and understanding different kinds of texts).

But meanings can be more or less standardised. Highly standardised meanings are also meanings by convention, and they also require some degree of interpretation. But because the meaning is such a standard one, the process of meaning construction no longer requires any effort; it has become automatic, and pretty much everyone with the same cultural background will agree that this is what the material means. When you are dealing with highly standardised meanings, there is really no need to apply a method like QCA (or any other qualitative method for data analysis). This would be the case, for instance, if you were interested in finding out about the number of men and women shown in magazine advertisements: little interpretation is needed to decide whether the persons in the picture are male or female. With such material, quantitative content analysis would be a good method to use (see Chapter 2 for the difference between qualitative and quantitative content analysis).

QCA comes into its own when you are dealing with meaning that is less obvious. If you were interested in finding out whether women in magazine advertisements are more often placed in trivial contexts than men, for instance, you would be dealing with a much less standardised meaning. What exactly is a trivial context? Not everyone would agree on this, and often you will only be able to tell whether a context is trivial or not by examining the image in some detail. It is in this sense that QCA is a suitable method for describing material that requires some degree of interpretation. When you are engaged in qualitative research, your data will usually be of the type that requires some interpretation.

As long as your material is of this kind, QCA will be an option. It does not matter, for instance, how you came by your data: whether you generated it in the process of doing your research (by doing interviews, or creating observation protocols) or whether you sampled material from other sources (such as newspapers or documentary archives). QCA can be applied to a wide range of

materials: interview transcripts, transcripts of focus groups, textbooks, company brochures, contracts, diaries, websites, entries on social network sites, television programs, newspaper articles, magazine advertisements, and many more (for an overview of qualitative methods for data collection see Marshall & Rossman, 2006, Chapter 4; on the interview in particular see Wengraf, 2001; Witzel & Reiter, in press).

It also does not matter whether your material is verbal or visual. When QCA was first developed, it was used for analysing and interpreting texts, such as newspaper articles (see below). Today, too, QCA is most often applied to verbal data, such as interview transcripts, emails, archival material, and the like. This is why the data for QCA is often referred to as 'text'. But this is merely a matter of convenience and of habit, and many authors have stressed that content analysis can just as well be used for analysing visual material (for example, Krippendorff, 2004; Shapiro & Markoff, 1997). It is helpful to keep this in mind when reading about the 'text' or the 'textual material' to which QCA is applied: 'Text' is used as a generic term here, covering all kinds of qualitative material, visual as well as verbal.

#### **Checklist: When to use QCA**

- When you are dealing with rich data that requires interpretation
- On verbal data
- On visual data
- On data that you have sampled from other sources (documents, internet, etc.)
- On data that you have collected yourself (interviews, focus groups, etc.)

### **The goals of QCA**

In most general terms, the aim of QCA is to systematically describe the meaning of your material. The systematic nature of content analysis, including QCA, is a point on which pretty much all authors who have written about the method over the years agree (to name only a few: Früh, 2007; Groeben & Rustemeyer, 1994; Holsti, 1969; Krippendorff, 2004; Mayring, 2000; 2010; Shapiro & Markoff, 1997). But this very broad goal needs to be qualified in two respects.

In the first place, QCA will help you describe your material only in certain respects which you have to specify. QCA does not allow you to describe the full meaning of your material in each and every respect. This characteristic points to an important difference between QCA and other qualitative methods for data analysis, especially methods that are rooted in a hermeneutic tradition. These methods take you along a spiral path, taking more aspects into account and arriving at a yet more comprehensive sense of your material at

every step. QCA is not like this. With QCA, your research question specifies the angle from which you examine your data. If other important aspects strike you during the analysis, you can change your coding frame and include these as well. But these will again be specific, selected aspects. QCA does not give you a holistic overview of your material. Früh (2007) and Groeben (1987) write about this in more detail.

Focusing on selected aspects of your material is what distinguishes QCA from many other qualitative methods for data analysis. On the one hand, selected aspects are less compared to the full, comprehensive meaning of a text. On the other hand, qualitative data are very rich anyway – so rich that it is impossible for all practical purposes to really capture their full meaning. Also, qualitative research tends to produce a *lot* of data. And it is easy to get lost in what can be hundreds of pages of interview transcripts. When you are dealing with a lot of very rich material, it can actually be very useful to focus on selected aspects only, and QCA helps you do so.

There is a second sense in which the very general goal of describing the meaning of your material needs some qualification. This concerns the question of what your description is for: does it stand on its own, or do you use your description as a basis for conclusions about an external social reality? Do you, for instance, simply want to describe advertisements depicting men and women and the contexts in which they are shown? Or do you want to use the information that women are (or are not) shown in trivialising contexts more often than men to infer something about the values held by the culture and society in which these magazines are published?

This has been a highly controversial topic, with authors such as Groeben and Rustemeyer (1994), Lisch and Kriz (1978) and Rössler (2005) arguing that QCA can never do more than describe the material to which it is applied, and others such as Früh (2007), Krippendorff (2004) and Merten (1995) claiming that QCA proper invariably involves conclusions that go beyond the material under study. This controversy has its roots in the different disciplinary origins of the authors, and we will return to it in the context of validity issues (see Chapter 9). For the time being you should simply be aware that your answer will probably depend on your material. If you have generated your material in the research process, a simple description will often be enough. If you have conducted interviews, for example, you will probably use QCA to describe what your participants have said. But if you have sampled your material from other sources, chances are that you will want to go beyond description. If you are analysing company brochures, for example, you will usually want to say something about the company, not just about the brochures. Likewise, if you are analysing gender depiction in magazine advertisements, you will be interested in gender stereotypes and gender roles in that society, not just in describing the advertisements. In this case, QCA may not be enough to substantiate your conclusions. You may need additional evidence to show that the brochures do indeed provide valid information about the company,

and that magazine advertisements adequately reflect upon the gender roles and stereotypes that are prevalent in society (they may well lag behind actual developments!).

### What QCA involves

In our everyday interactions with others, we are constantly engaged in deciphering meaning and in interpreting communication. But what we do on an everyday basis is largely intuitive: we do not always listen carefully, we may jump to conclusions, and sometimes we hear what we want to hear, and not what the other person is actually saying. Qualitative data analysis is like everyday understanding in its concern with meaning. But at the same time it goes beyond our everyday activities. Each method of qualitative data analysis specifies a distinctive way – a *method* – of overcoming the shortcomings of our everyday understanding. The way in which QCA does this is as follows: first, it requires you to ‘translate’ all those meanings in your material that are of interest to you into the categories of a coding frame; second, it has you classify successive parts of your material according to these categories. The way in which this is done highlights three important characteristics of QCA: it is a systematic method, it is flexible, and it reduces data.

---

#### KEY POINT

QCA is systematic, flexible, and it reduces data.

---

#### QCA is systematic

The *systematic* nature of QCA is probably its most distinctive feature (see above). QCA is systematic in three important respects.

In QCA, you examine *all* your material and decide for each part where in the coding frame it fits (Rustemeyer, 1992; Shapiro & Markoff, 1997). This is very important and a point to which we will come back several times in the course of this book. Unless you examine each and every part of your material, chances are that you will overlook those parts that do not fit the concepts and ideas that you bring to your research. And you invariably have such notions, even if you do not have any hypotheses.

Regardless of your material and your research question, QCA always involves the *same sequence of steps* (for content analysis in general see Neuendorf, 2002; for QCA in particular see Rustemeyer, 1992): deciding on a research question; selecting your material; building a coding frame that will usually comprise several main categories, each with their own set of subcategories; dividing your material into units of coding; trying out your coding frame through double-coding, followed by a discussion of units that were

coded differently; evaluating your coding frame in terms of the consistency of coding and in terms of validity and revising it accordingly; coding all your material, using the revised version of your coding frame, and transforming the information to the case level; interpreting and presenting your findings. Deciding on a research question, selecting your material, and interpreting and presenting your findings are an important part of all research, regardless of the method you are using. All other steps are specific to QCA.

### Steps in QCA

- 1 Deciding on your research question
- 2 Selecting your material
- 3 Building a coding frame
- 4 Dividing your material into units of coding
- 5 Trying out your coding frame
- 6 Evaluating and modifying your coding frame
- 7 Main analysis
- 8 Interpreting and presenting your findings

When we engage in understanding and interpretation of meaning in everyday contexts, we bring to this process our individual personalities, needs and moods; and all of this takes place in a specific situation. You may flare up at a chance remark that would not bother the next person in the least; and perhaps you only flared up today because you had a big fight with your partner the day before. In QCA, you go beyond your individual understanding at the given moment by checking whether your understanding stands the test of *consistency*. This can be consistency with how another person understands the same passage; it can also be consistency with what you take the passage to mean at another time (see Chapter 9 in more detail on how to go about this).

Consistency in this sense refers to what has been called *reliability*. The origins of reliability are in quantitative research where the criterion is used to assess the quality of instruments (see, for instance, Bryman, 2008, pp. 149ff.; Cresswell, 2009, pp. 190ff.). In general terms, an instrument is considered to be reliable to the extent that it yields data that is free of error. Checking for consistency between coders or between different points in time is one way of assessing the reliability of your coding frame. Note that this is not to say that your own, individual understanding is not worthy of consideration! It definitely is, and this plays an important role as you build your coding frame (see Chapter 5). But when it comes to classifying your material according to this coding frame, the goal of QCA is to go beyond individual understanding and interpretation (on the role of consistency and reliability in QCA see Boyatzis, 1998; Mayring, 2010; Rustemeyer, 1992).

### QCA is flexible

At the same time, QCA is also a highly *flexible* method – flexible in the sense that you will always have to tailor your coding frame to your material (Boyatzis, 1998; Rustemeyer, 1992). This is because you not only want your coding frame to be reliable – you also want it to be *valid*. In the methodological literature, an instrument is considered valid to the extent that it in fact captures what it sets out to capture (Krippendorff, 2004, Chapter 13; Neuendorf, 2002, Chapter 6; see Chapter 9 below for a more in-depth discussion of the role of validity in QCA). Your coding frame can be regarded as valid to the extent that your categories adequately represent the concepts in your research question, and to achieve this you have to adapt your frame so as to fit your material.

This is an important difference from quantitative content analysis. Here, concept-driven coding frames are sometimes used, and standardised coding frames have been developed, such as the coding frames by Gottschalk and Gleser (1967) for assessing the expression of emotions in textual material. Their assumption is that expressions of basic emotions and their indicators remain the same, regardless of the person expressing them and the context in which they are expressed. But in QCA, you are always to some extent concerned with describing the specifics of your material. And to do so, your coding frame has to match your material. Because of this, coding frames in QCA are always partly data-driven. You can make use of theory or of coding frames developed by other researchers, but you have to adapt these to the material that you are studying.

### QCA reduces data

Finally, QCA *reduces* data – and in this respect it is different from other methods for qualitative data analysis (Früh, 1992; Groeben & Rustemeyer, 1994). Most methods for qualitative data analysis are concerned with opening up your data, discovering new things about it, bringing it together in novel ways. This usually involves producing even more data – data about your data, as it were (see Chapter 2). QCA is different. It focuses your analysis on selected aspects, and in this process it reduces your material in two ways:

- In the first place, you do not take into account all the information provided by a case (be it a document, an interview transcript, etc.). Instead, you limit your analysis to those aspects that are relevant with a view to your research question.
- Second, the categories of your coding frame will usually be at a higher level of abstraction than the more concrete information in your material. By classifying the specific, concrete information in your material according to your coding frame, you lose these specifics. This is the price you pay for being able to compare one specific piece of information to another (within the same case or as part of another case). In the process of coding, you classify all specific information as instances of higher-order categories. Building a coding frame in QCA is all about finding the right

balance here. You will make your categories sufficiently abstract to allow for comparison and sufficiently concrete so as to preserve as many specifics as possible (see Chapters 4 and 5).

But note that as you engage in reducing specifics through the process of classification, you are at the same time producing new information (Früh, 1992). This is information *across cases*, telling you how your cases compare to each other with respect to the categories in your coding frame. You may lose specific information on the individual level, but you gain information on the aggregate level!

---

### Example of how you reduce your material using QCA

---

McDonald, Wearing and Ponting (2009) wanted to find out which elements of wilderness settings contribute to what has been called a peak experience: a brief experience of happiness, fulfilment, and spiritual insight. They asked 39 persons who had visited a wilderness setting to describe in their own words the most wonderful experience they had had there. They then used QCA to analyse these descriptions in a data-driven way, resulting in what they call seven core themes, i.e. seven important aspects of peak experiences in wilderness settings. This analysis reduces the interview material by focusing only on characteristics of peak experiences. Other aspects of the participants' responses were not included in the analysis, such as how they came to choose this wilderness setting in particular, or potential negative effects of their experience, such as forgetting to make an important phone call. It also reduces the material by subsuming all the individual details of the participants' experiences under these seven core themes. Theme 2, for instance, refers to escape from the man-made world. Different aspects of the man-made world are summarised here: the presence of other people as such, the sheer number of other people in a large city, as well as man-made media. At the same time, creating this category also produces new information, namely information about what there is in the man-made world that people want to escape from: number of people, the omnipresence of the media, and an inability to focus on one's surroundings because of a constant input from the outside world. Creating the category relates these to each other.

---

### SUMMARY

---

QCA is a method for systematically describing the meaning of qualitative material. This is done by classifying parts of your material as instances of the categories of a coding frame. The method is suitable for all material that requires some degree of interpretation. This can be verbal or visual, and it can be material that you generated for your research, or material that you sampled from other sources. QCA requires that you focus your analysis on selected aspects of your material, as indicated by your research question; in this respect it differs from other qualitative methods of data analysis. QCA is systematic, flexible, and it reduces data. The

method is systematic in three ways: all relevant material is taken into account; a sequence of steps is followed during the analysis, regardless of your material; and you have to check your coding for consistency (reliability). It is flexible in that your coding frame must be adapted so as to fit your material, i.e. to be valid. And it reduces your material by limiting your analysis to relevant aspects of the material. Also, through classifying specific information as an instance of a category, you subsume the specific information under a more general concept. At the same time, categorising also produces new information about how your cases compare.

---

## The origins of quantitative content analysis in the social sciences

We will now look at the history of quantitative content analysis, originating in communication studies. The history of the method can be divided into three broad phases: a first phase of early applications, lasting from approximately the eighteenth century until the early twentieth century; a second phase when quantitative content analysis came into its own as a research method, lasting until the late 1940s; and a third phase of interdisciplinary and methodological elaboration that is continuing until the present day. Developments during these phases actually overlap to some extent. This is why only approximate beginnings and ends of phases are given.

### First phase: Early applications

People were interested in the systematic analysis of text a long time before 'content analysis' was formally developed as a method in the social sciences (for more detail see Krippendorff, 2004; Merten, 1995).

In the late seventeenth and eighteenth century, Church potentates were worried that non-religious or unorthodox material might be distributed in the name of the Church. To prevent this, they commissioned analyses of religious texts. In eighteenth-century Sweden, for instance, a collection of hymns was analysed for the frequency of certain key words (such as *God*, *Kingdom of Heaven*) to determine whether these songs were in line with Church teachings (it was concluded that they were; see Dovring, 1954).

---

### Example of an early use of newspaper analysis

---

In 1893 Speed published an analysis of the themes covered by different New York newspapers, comparing the years 1881 and 1893. He concluded that over time the coverage of themes such as gossip and scandal had increased at the expense of religious and scientific content (Speed, 1893).

As the newspaper gained in popularity and turned into the first 'mass medium' in the second half of the nineteenth century, there was also an increasing interest in the content distributed by this medium. Journalism schools were founded, and founders and teachers wanted to instil in their students ethical standards of journalism and 'objective' reporting. Scientists also wanted to know whether newspaper content was in fact objective, ethical, and 'edifying'. In this context, quantitative descriptions and differentiations of newspaper content became the focus of early content analysis.

Often, this involved comparisons, both within one newspaper (by following its development over time) and by comparing different newspapers in terms of the themes that were covered. In determining the relative importance of the different themes, researchers did not rely on the number of articles alone. They also took into account number of words, percentage area of a page taken up by an article on a given topic, letter size of headlines, placement on the page, and the like (Merten, 1995, Chapter 1.2 provides many detailed examples).

### Second phase: Content analysis coming into its own

During the second phase, content analysis was developed into a research method in the empirical social sciences. Two developments contributed to this: the rise of the social sciences and an increasing interest in the effects of communication content in the media (Krippendorff, 2004; Lissmann, 2008; Merten, 1995).

In the 1930s and 1940s, other social science disciplines such as sociology and psychology were gaining in importance, and researchers from these disciplines introduced new concepts, such as social stereotypes or attitudes. These social science concepts affected the analysis of communication content in two ways. First, these concepts were theory-based, requiring far more sophistication of conceptualisation and measurement than had been customary in early quantitative newspaper analysis. Second, the concepts suggested new directions in the analysis of mass media content. The concept of stereotypes, for instance, created an interest in how certain key issues were represented in media products. Simpson (1936), for example, analysed representations of black Americans in the press, and Martin (1936) examined the representation of nationalism in children's books from different countries.

#### — Example of the use of stereotypes in early content analysis —

For her analysis of nationalism in children's books, Martin (1936) selected 24 popular children's books from 12 different countries in their English translation (among them *Pinocchio*, *Heidi*, *Jungle Book*). For her analysis, she identified symbols of nationalism which she then analysed in three respects: according to subject, whether the symbol referred to the country of origin or a different

country, and whether the evaluation of that country was positive or negative. Altogether, she identified approximately 1,000 such symbols in each of the books! As it turned out, books from the different countries did not differ significantly in the amount and type of national symbolism used. But Martin did find more nationalist symbolism in children's books that were published after times of crisis (compared to books published before or during a national crisis).

The concept of attitude proved to be of special importance for the further development of content analysis. Whereas in the past, analysis of newspaper content had been limited to themes, Lasswell (1941) now examined how such themes were evaluated (see also the above example of Martin's analysis of children's books). For each theme, he also recorded the direction of the evaluation, differentiating between negative, neutral, and positive evaluations. This analysis of the evaluative dimension of representations in the media has become standard in quantitative content analysis in communication studies.

Whereas early analysis of newspaper content had been limited to the content itself, the 1930s brought an increasing interest in inferences from communication content to the recipients. This new interest was linked to the advent of the new media of radio and film. There was now increasing competition within the media landscape, raising the question of how to attract and hold the interest of the audience. More media also meant more advertising, and with more advertising there came a concern with designing maximally effective messages.

The increasing interest in the effects of media messages was also closely related to the Second World War. In 1939, the US Government made Harold Lasswell head of the department for the analysis of wartime communication. Lasswell had already begun to develop propaganda analysis in the 1920s (Lasswell, 1927). Under his directorship, ongoing content analysis of propaganda issued in Nazi Germany was carried out (e.g. Lasswell & Jones, 1939; see also the overview in Schramm, 1997). Unlike the previous quantitative analysis of newspaper content, propaganda analysis required that communication content be placed in the context of both its production and its reception.

With propaganda analysis, Lasswell not only opened up a new substantive area for the application of content analysis (the first among many). He also began to refine the method, adding considerations concerning sampling, the building of categories, and assessing agreement between coders as a quality measure. This marked the beginning of a period of methodological reflection on content analysis as a research method, starting in the year 1941.

#### KEY POINT

1941 was the 'birth year' of content analysis.

This was the year when content analysis really came into its own. A conference on mass communication was held in Chicago which was attended by all leading scholars in the field, and the focus of the conference was on content analysis (Waples, 1942). Soon afterwards, Berelson and Lazarsfeld (1948) published a first introduction to the method. Based on this book, Berelson (1952) published what was to become the first leading textbook on content analysis.

### Third phase: Interdisciplinary and methodological elaboration

From the 1940s, content analysis began to attract attention as a research method outside communication studies. This trend continued during the subsequent years, and content analysis came to be used in many diverse disciplines such as political science, psychology, education, and literary studies (Krippendorff, 2004; Merten, 1995).

In political science, Lasswell's propaganda analysis had already paved the way for content analysis. The beginnings of the use of content analysis in psychology were marked by the work of Gordon Allport (1942; 1965). He used personal structure analysis (a variant of content analysis: Baldwin, 1942) to analyse 301 letters written by one woman ('Jenny'), identifying key themes and drawing conclusions from these themes on Jenny's attitudes and personality. Other landmark applications of content analysis in and to psychology include Bales's development of a multidimensional coding frame for analysing the verbal interactions between the members of small groups (interaction process analysis: Bales, 1950), and using the method for analysing free responses to the Thematic Apperception Test (Smith, 2008), a personality test for assessing the strength of people's motives (such as power, success, or affiliation).

In educational research, content analysis was, for instance, used to analyse texts in terms of their readability. Flesch (1948) developed a readability formula that was based on average sentence length, average word length, number of personal words, and number of personal sentences in a text. At the same time, this is an example of content analysis that does not in fact focus on textual content features, but on the formal characteristics of a text. Formal features were also the focus in applying content analysis to literary studies where it has, for example, been used to differentiate between potential authors of texts of unknown authorship. Yule (1944) used stylistic content analysis to establish that of two authors, Thomas à Kempis and Charlier de Gerson, the former was more likely to have written the text *De Imitatione Christi*.

This use of content analysis in different disciplines was accompanied by an increasing methodological differentiation. To adapt the method to the research questions that were asked in the different disciplines, ever new variants of content analysis were developed, such as contingency analysis (Osgood, 1959), value analysis (White, 1944), the semantic differential (Osgood, 1952), and

others already mentioned above (readability analysis; personality structure analysis; analysis of motives, etc.). Overall, these developments were characterised by increasing attention to the context in which communication content is produced or received, and by changing the focus of the analysis from the frequency of selected textual characteristics to their interrelation. These elaborations of the method were discussed at a second landmark conference on content analysis which took place in 1955 (Pool, 1959). Later conferences in 1967 and 1974 increasingly focused on the use of computers in content analysis (Gerbner et al., 1969; Stone, 1975).

---

### SUMMARY

The history of content analysis can be divided into three phases: early applications, content analysis coming into its own, and interdisciplinary and methodological elaborations. Early applications focused on the quantitative description and differentiation of newspaper content, often from a comparative perspective. The second phase was characterised by more sophisticated conceptualisation and measurement as well as an increasing interest in the effects of content on the recipients. During the third phase, content analysis came to be used in other social science disciplines. As the method was applied to novel kinds of research questions, ever more variants were developed. This was accompanied by increasing attention to the context of production and reception and to the interrelation of selected textual characteristics.

---

## The emergence of qualitative content analysis

### Critique of quantitative content analysis

In his textbook on content analysis, Berelson (1952, p. 18) defined the method as follows: 'Content analysis is a research technique for the objective, systematic, and quantitative description of the manifest content of communication.' But this definition, with its strong focus on content analysis as a quantitative method, was contested even in that same year by Kracauer (1952). Kracauer argued against a purely quantitative type of content analysis on three grounds:

- Meaning is often complex, holistic, and context-dependent.
- Meaning is not always manifest and clear at first sight. Sometimes it is necessary to read a text in more detail to determine what exactly it means.
- Some aspects of meaning may appear only once in a text. This does not necessarily imply that such aspects are less important than aspects that are mentioned more frequently.

Based on these considerations, Kracauer favoured a more qualitative type of content analysis that does not limit itself to manifest content and frequency

counts. This suggestion was taken up by George (1959) who had been one of the researchers engaged in the analysis of wartime propaganda. The analysis of propaganda, George argued, requires an analysis of strategy, and strategy often manifests in what he called non-frequency indicators, namely the single occurrence of a certain phrase or word throughout a text. In fact he preferred the term 'non-frequency' to the term 'qualitative' in characterising the variant of content analysis that he was proposing: 'We employ the term "non-frequency" to describe the type of nonquantitative, nonstatistical content analysis which uses the presence or absence of a certain content characteristic or syndrome as a content indicator in an inferential hypothesis' (George, 1959, p. 8). In this way, George took up one of Kracauer's criticisms of quantitative content analysis, namely its focus on frequency counts (for a similar conceptualisation of qualitative content analysis see Holsti, 1969).

### Development of QCA

On the one hand, these early criticisms of quantitative content analysis established a sharp dichotomy between a quantitative and a qualitative variant of the method, reflecting the division between adherents of the quantitative and the qualitative research paradigm. On the other hand, this sharp contrast becomes blurred on closer inspection. Berelson himself wrote that some research questions require a more qualitative approach, and George, by choosing the term 'nonfrequency analysis', attempted to evade the distinction altogether. As quantitative content analysis evolved and became more sophisticated, it was increasingly applied to less manifest content. In consequence, many proponents of quantitative content analysis argued that the distinction between a qualitative and a quantitative type of content analysis was artificial, that 'qualitative' and 'quantitative' was merely a matter of degree (cf. Früh, 2007; Holsti, 1969; Krippendorff, 2004; Lisch & Kriz, 1978; Merten, 1995; see also the overview in Groeben & Rustemeyer, 1994). In this way, especially in England and the US, quantitative content analysis opened up towards these first attempts at establishing a qualitative version of the method and ultimately came to embrace it.

Because of this development and the continuing dominance of quantitative content analysis, QCA has not been well known as a research method, especially in English-speaking countries, until recently. Many qualitative researchers do not mention QCA at all (cf. Gibbs, 2007; Mason, 2002; Miles & Huberman, 1994; Silverman, 2000) or present a very quantitative version of the method (compare the descriptions in Berger, 2000, or in Bernard & Ryan, 2010). Other authors equate QCA with the whole range of qualitative methods for data analysis, subsuming other methods such as discourse analysis, conversation analysis, or objective hermeneutics under QCA (Krippendorff, 2004; Lamnek, 2010). Yet other authors describe what is essentially QCA, but call it by a different name, such as 'thematic coding' (cf. Boyatzis, 1998; Saldana, 2009) or 'qualitative media analysis' (Altheide, 1996). It is only

recently that QCA has been described as a distinct method in the Anglo-American literature (e.g. Elo & Kyngäs, 2008; Hsie & Shannon, 2005), although some authors (e.g. Klotz & Prakash, 2008) continue to use the term in the same way as did Kracauer (1952) or George (1959), i.e. to refer to an analysis of the presence versus absence of specified themes or features.

On the Continent, especially in Germany, the situation has been different, and there have been a number of conceptualisations of a genuinely qualitative QCA. These include Ritsert's (1972) concept of an anti-ideological variant of the method (see also Vorderer & Groeben, 1987), Rust's (1980) development of a 'strict and qualitative' type of QCA, and flexible QCA (Groeben & Rustemeyer, 1994; Hussy, Schreier & Echterhoff, 2009; Rustemeyer, 1992; other variants can be found in Bilandzic, Koschel & Scheufele, 2001; Gläser & Laudel, 2009; Kuckartz, 2009; Mathes, 1992). A major proponent of QCA in Germany has been Philipp Mayring (2010). He distinguishes between several distinct variants of the method, such as summative and structural QCA. We will look at these in more detail in the context of developing a coding frame in QCA (Chapters 5 and 6).

### What is different about QCA?

Who is right? Those who argue that there is no need for QCA because quantitative content analysis can do it all? Or those who have elaborated QCA as a distinct research method? There is certainly no sharp line dividing quantitative content analysis and QCA. Nevertheless all versions of QCA share some characteristics which distinguish the method from quantitative content analysis (see Table 1.1).

---

#### KEY POINT

There is no sharp line dividing quantitative and qualitative content analysis. Nevertheless all versions of QCA share some characteristics which make it a method in its own right.

---

The most important difference was suggested by Kracauer in the early 1950s: the focus of QCA is on *latent* meaning, meaning that is not immediately obvious, whereas quantitative content analysis focuses on manifest, literal meaning (Berelson, 1952; Kracauer, 1952; see the discussion in Groeben & Rustemeyer, 1994; Lisch & Kriz, 1978).

Because manifest meaning is fairly obvious at first sight, you can usually identify it by looking at a small segment of material, such as a single sentence or paragraph. To detect latent meaning, on the other hand, you often have to take context into account. This can be the entire text from which a passage is taken – or even the publication venue or additional background information. If you come across a passage praising the foresight of George W. Bush in the Iraq war, for instance, and

**Table 1.1** Differences between quantitative content analysis and QCA

<i>Quantitative content analysis</i>	QCA
Focus on manifest meaning	Focus on latent meaning
Little context needed	Much context needed
Strict handling of reliability	Variable handling of reliability
Reliability checks more important than validity checks	Validity checks just as important as reliability checks
At least partly concept-driven	At least partly data-driven
Fewer inferences to context, author, recipients	More inferences to context, author, recipients
Strict sequence of steps	More variability in carrying out the steps

you know that this comes from a news broadcast in 2008 on the Fox News channel (known for its Republican sympathies), you will take the passage literally. But if you know that it comes from an article published in *Mad* (a major US satirical magazine), you will take it to be ironic and mean the very opposite of what it says. QCA therefore requires you to take *context* into account.

In quantitative content analysis, reliability by double-coding is the most important quality criterion. This is closely related to the focus of quantitative content analysis on manifest meaning. If two persons independently code the same passage, they are more likely to code it the same way if the meaning of the passage is manifest. The more hidden the meaning is, the more context you need in order to infer it, and the more likely it is that two people will read it differently (Neuendorf, 2002). In quantitative content analysis, high reliability is fairly easy to achieve because of the focus on manifest meaning; in QCA, when looking at latent meaning, reliability is more difficult to achieve. Therefore it is only to be expected that reliability is handled differently in QCA (see Chapter 9 for details). In the first place, in QCA consistency scores between coders are acceptable. Second, agreement between coders is not necessarily quantified. Third, in QCA validity is considered to be just as essential as reliability.

This takes us to the next characteristic of QCA, namely the importance of *validity* as a quality criterion (Holsti, 1969; Lisch & Kriz, 1978; Rustemeyer, 1992). This importance is closely related to the role of theory and description in quantitative content analysis and QCA, respectively. In quantitative content analysis, coding frames will usually be partly concept-driven, and you may want to use the method for hypothesis testing. In QCA, on the other hand, your coding frame will usually be partly *data-driven*, and you may want to use the method primarily for describing your material. Overall, theory and prior research play a greater role in quantitative content analysis, and working in a data-driven way is more important in QCA. In quantitative content analysis, theory validates the concept-driven parts of your coding frame. But in QCA you have to make sure that the data-driven parts of your coding frame really capture what is there in your material. Because of this, a validity check is just as important in QCA as a reliability check.

Some authors have argued that QCA is also more likely to be used in making inferences about the context of production, the authors, or the effects on the recipients (Groeben & Rustemeyer, 1994). If you want to draw such inferences that go beyond your material, checking for the validity of your conclusions becomes even more important.

Finally, quantitative content analysis always follows a certain series of steps. So does QCA (although some versions do not: see Altheide, 1996; Lamnek, 2010), but there is more variety in QCA. If you are building a data-driven coding frame, for instance, you may do so based on 5% of your material, 20%, or even all of it. In checking for the reliability of your coding, you may quantify the consistency of the coding, or you may simply sit down with the other coders and explain why you coded a passage in a certain way. Overall, QCA is therefore more *flexible* than quantitative content analysis.

---

#### SUMMARY

Already in 1952, quantitative content analysis was criticised on three accounts: meaning is often complex, it may be latent, and it may appear only once in a text. Non-frequency analysis, an early version of QCA, was suggested as an alternative. As quantitative QCA became more sophisticated, proponents of quantitative content analysis came to consider the distinction between quantitative and qualitative content analysis as a matter of degree. Especially in the Anglo-American context, quantitative content analysis encompassed QCA. On the Continent, especially in Germany, QCA was developed as a method in its own right. Recently, it has also come to be recognised as a distinct method of qualitative data analysis in an Anglo-American context. There is no sharp line dividing quantitative content analysis and QCA. Nevertheless all versions of QCA share certain characteristics: focus on latent meaning; attention to context; variable handling of reliability; validity checks just as important as reliability checks; at least partly data-driven; more inferences to context, author, and recipients; more flexibility in going through the steps.

---

### Outlook: What lies ahead

By now, you have gained a first impression of what QCA is all about and how it evolved. In the following chapters, we will look at QCA and the steps involved in carrying out a QCA in more detail.

Although QCA is a qualitative research method, it has its roots in both the qualitative and the quantitative research tradition. In Chapter 2 we will look at some important features of qualitative research and examine in what way QCA does (or does not) exemplify these features.

While QCA is a highly useful method for qualitative data analysis, especially when you are dealing with a large-scale study, there are some research objectives for which it is more suitable than others – and some for which it is not suitable

at all. In Chapter 3 we will compare QCA to other methods for qualitative data analysis. On this basis, I will help you decide whether QCA would be a good method for you to use, considering your material and your research question.

The following chapters will then guide you in actually carrying out your own QCA. In Chapter 4, I will explain in more detail what a coding frame is and how to structure a frame. On this basis, we will then look at the steps involved in building a coding frame in general (Chapter 5) and strategies for generating data-driven categories in particular (Chapter 6). To apply a coding frame, you first have to divide your material into smaller parts; this so-called process of segmentation will be covered in Chapter 7. Once you have generated a first version of your coding frame and segmented your material, you are ready to try out your frame in a trial coding. The trial coding is at the core of the pilot phase which we will describe in Chapter 8. Based on your trial coding, the next step is to evaluate your coding frame and modify it accordingly (Chapter 9). Chances are that your frame is now suitable for starting on your main analysis, i.e. assigning all your material to the categories of your frame (Chapter 10). Once you have completed your QCA, you will want to present your findings; Chapter 11 provides you with an overview of strategies for doing so.

Nowadays, qualitative researchers are increasingly making use of software to support their analysis. In Chapter 12 we will take a look at the kinds of software packages that are available today and how they can support you during the different steps of carrying out a QCA.

To illustrate the process of QCA, many examples will be used throughout the book. Nevertheless, you may find yourself wanting to look at more sample studies. You will find an additional chapter describing examples of studies using QCA from a variety of different social science disciplines on the website accompanying this book.

## Frequently asked questions

### Is content analysis really a qualitative method?

You will probably find as many opinions about this as there are researchers using the method. Content analysis as used in communication studies is typically quantitative content analysis. As I explain earlier in this chapter, QCA developed out of quantitative content analysis. And while there is no sharp line dividing QCA from quantitative content analysis, the various versions of QCA share a number of features which distinguish the method from quantitative content analysis. Because these are features that QCA shares with other qualitative research methods, I would argue that QCA is indeed a qualitative method. In the next chapter, we will look in more detail at what QCA has in common with the qualitative research tradition and where it differs from this tradition and is closer to the quantitative framework.