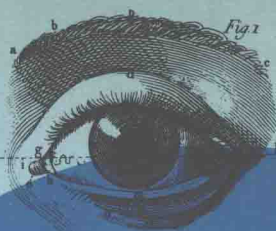




(Collect data)



Qualitative Data Analysis

Qualitative Data Analysis

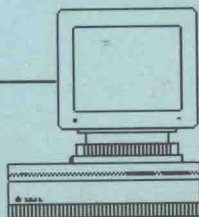
*A User-Friendly Guide
for Social Scientists*

IAN DEY



(Sort)

(Compute)



Qualitative data analysis

A user-friendly guide for social scientists

Ian Dey



London and New York

First published 1993

by Routledge

11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada

by Routledge

29 West 35th Street, New York, NY 10001

Reprinted 1993, 1995, 1996

© 1993 Ian Dey

Typeset in Times by Solidus (Bristol) Ltd, Bristol

Printed and bound in Great Britain by

Biddles Ltd, Guildford and King's Lynn

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data

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

Library of Congress Cataloguing in Publication Data

A catalogue record for this book is available from the Library of Congress

ISBN 0-415-05851-1 (hbk)

ISBN 0-415-05852-X (pbk)

Preface

A new book on qualitative data analysis needs no apology. By comparison with the numerous texts on statistical analysis, qualitative data analysis has been ill-served. There is some irony in this situation: even a single text might suffice for the standardized procedures of statistical analysis; but for qualitative analysis, oft-noted for the diffuse and varied character of its procedures, we might reasonably expect a multiplicity of texts, not just a few. Teaching a course on methods makes one especially aware of this gap. This book is my contribution to filling it, and I hope it will encourage – or provoke – others to do the same.

A contemporary text on qualitative data analysis has to take account of the computer. The days of scissors and paste are over. While those steeped in traditional techniques may still harbour suspicions of the computer, a new generation of undergraduates and postgraduates expects to handle qualitative data using the new technology. For better or worse, these students will not give qualitative analysis the same attention and commitment as quantitative analysis, if only the latter is computer-based. This book is written primarily for them. I hope it may also be of some interest to other researchers new to qualitative analysis and to those using computers for this purpose for the first time.

Although the methods presented here assume the use of specialist software to support qualitative analysis, those seeking an introduction to individual software packages must look elsewhere (for example, Tesch 1990). My intention is to indicate the variety of ways in which computers can be utilized in qualitative analysis, without describing individual software applications in detail. No one application – including my own package, Hypersoft – will support the whole range of procedures which can be employed in analysing qualitative data. The researcher will have to choose an application to support a particular configuration of procedures, and one of my aims is to permit a more informed choice by identifying the range of analytic tasks which can be accomplished using one software package or another.

The challenge of developing a software package to analyse qualitative

data has been a useful stimulus to clarifying and systematizing the procedures involved in qualitative analysis. It has also allowed me to write a text informed by what we can do with the computer. In my view, the advent of the computer not only enhances, but in some respects transforms traditional modes of analysis.

The book is based on my experiences as a researcher and teacher as well as a software developer. My research has involved a variety of qualitative methods, including observation, in-depth interviewing and documentary analysis; and through it I have learnt some of the procedures and paradoxes of qualitative analysis. As a teacher, I have become convinced of the merits of 'learning by doing', a perspective which has informed the skills-based methods course I have taught over the last few years with my colleague, Fran Wasoff. For those interested in skills acquisition, a text which provides a variety of task-related exercises and small-scale projects for students would be an invaluable asset. But this is not my aim in this book. Experience of teaching qualitative methods has also persuaded me of the value of a clear and uncomplicated introduction providing essential background knowledge and helping to structure the learning experience. This is what I hope this book will do.

A text introducing computer-based qualitative data analysis may need no apology, but my decision to illustrate analytic procedures using everyday material – mostly humorous – probably does deserve some explanation. The shortest explanation is that it works. Methods courses are notoriously dull. Pedagogical devices which work well enough for substantive issues can fail to engage students sufficiently in a course on methods. Students quickly tire of reading about methods, when what they want is to acquire and practise skills. In recent years I have been involved in teaching a methods course which aims to stimulate student interest and maintain motivation. One lesson I have learnt in teaching this course is that the problems students work on should be interesting and entertaining as well as instructive: that methods can be fun. We have used everyday material and humorous examples in our methods course, and it never fails to stimulate students' interest and engage their attention. I think this is a question of Mohammed coming to the mountain, rather than the mountain coming to Mohammed. It is better to introduce qualitative analysis on students' terms, rather than one's own. Students unfamiliar with research find familiar examples reassuring. They can relate to the material without effort. Because they can relax and even enjoy the substantive material, they can concentrate better on procedures and process. If students can easily grasp research objectives, and quickly become familiar with the data being analysed, they are more likely to find qualitative analysis a manageable and rewarding challenge.

In this book, I have mainly used humour as the medium through which to discuss the methodological problems of qualitative data analysis. Apart

from offering light relief, humour is a subject we can all relate to. Whereas substantive issues are likely to be of minority interest, humorous exemplars are accessible to all. We can analyse humour from any number of perspectives – anthropological, linguistic, psychological, sociological and so on. This is a significant advantage in a text which is addressing methodological issues germane to a number of subjects and disciplines. Humour might be thought distracting, but in fact I want to reduce the distractions which can derive from using substantive topics and issues as exemplars. By using humour as the subject of analysis, I want to ensure that attention remains focused on how to analyse data, and not on what is being analysed. Needless to say, the examples used are not intended to be taken too seriously. My main examples, from Victoria Wood and Woody Allen, are chosen for their entertainment value rather than any academic import.

Two other advantages accrue from using humour as a subject for analysis. Humour often turns on ambiguities in meaning, and therefore raises some of the central problems in analysing qualitative data. In particular, it precludes a merely mechanical approach to analysing data. Humour is also an experience which suffers from dissection: analysis kills humour, just as surely as vivisection kills the frog. This underlines the limits (and limitations) of analysis, which can describe, interpret and explain, but cannot hope to reproduce the full richness of the original data.

Familiarity with the data is also important because it is a prerequisite of qualitative analysis. This presents a problem in teaching qualitative analysis, which typically deals with large volumes of data. My 'solution' is to teach analytic procedures through very limited sets of data, with which students can become thoroughly familiar. Although this has drawbacks, I think it gives more feel for what qualitative analysis is about. It avoids students being overwhelmed by a mass of material, and gives them more confidence that they can analyse data effectively. It also helps to focus on method, and counter the almost fetishistic concern with the sheer volume of material produced by qualitative methods. Using limited data in this way may seem like dancing on the head of a pin; but, after all, it is learning the dance that matters, and not the pin.

Acknowledgements

My thanks are due to Elisabeth Tribe of Routledge for her support, to my colleagues for their encouragement and assistance, and to the members of my family for their forbearance while I was writing this book.

The author gratefully acknowledges permission to reproduce the following copyright extracts:

Allen, Woody (1978) 'If the Impressionists had been Dentists' *Without Feathers*, London: Sphere. © Woody Allen 1972. Reprinted by permission of Random House, Inc. and Hamish Hamilton.

Extracts from: Wood, Victoria (1985) *Up to You, Porky: The Victoria Wood Sketch Book*, London: Methuen; and Wood, Victoria (1990) *Mens Sana in Thingummy Doodah*, London: Methuen. © Victoria Wood. Reprinted by permission of the author.

Illustration 1.1 on p. 2, from Tesch (1990: 58), is reprinted by permission of the author.

Contents

List of figures, illustrations and tables	vi
Preface	xi
Acknowledgements	xiv
1 Introduction	1
2 What is qualitative data?	9
3 What is qualitative analysis?	30
4 Introducing computers	55
5 Finding a focus	63
6 Managing data	74
7 Reading and annotating	83
8 Creating categories	94
9 Assigning categories	113
10 Splitting and splicing	129
11 Linking data	152
12 Making connections	168
13 Of maps and matrices	192
14 Corroborating evidence	219
15 Producing an account	237
16 Conclusion	264
Appendix 1: 'If the Impressionists had been Dentists'	269
Appendix 2: Software	273
Glossary	275
References	277
Index	280

Figures, illustrations and tables

FIGURES

1.1	The steps involved in data analysis – chapter by chapter	8
2.1	Describing a bit of data as a ripple in the flow of experience	18
2.2	Category relating two similar observations	19
2.3	Categorizing using inclusive categories	20
2.4	Nominal variable with mutually exclusive and exhaustive values	22
2.5	Ordinal variable indicating order between observations	23
2.6	Interval variable with fixed distance between values	23
2.7	Quantitative and qualitative data in dynamic balance	28
3.1	Qualitative analysis as a circular process	31
3.2	Three aspects of description in qualitative analysis	32
3.3	Categorizing as a method of funnelling data	43
3.4	Derivation of nominal variables with exclusive and exhaustive values	45
3.5	Formal connections between concepts	46
3.6	Formal and substantive connections between building blocks	47
3.7	Connections between chronological or narrative sequences	50
3.8	Causal connections between concepts	51
3.9	Qualitative analysis as a single sequential process	53
3.10	Qualitative analysis as an iterative spiral	53
4.1	A link between text held in separate locations	59
5.1	Deriving hypotheses about humour from the literature	69
5.2	Main themes for analysing humour	72
5.3	Integrating themes around issues of style and substance	73
6.1	Case documents kept in a hierarchical file system	79
6.2	Data stored in fields on a card-based filing system	80
7.1	Relating data to key themes	91
7.2	Mapping ideas to data within and across cases	92
7.3	Relating two ideas	93
8.1	Alternative category lists for analysing female stereotypes	101
8.2	Weighing up the degree of refinement in initial category set	106

8.3	Developing a more refined category list	108
9.1	Categorizing data – 1	113
9.2	Categorizing data – 2	114
9.3	Categorizing data – 3	114
10.1	Levels of subclassification of the subcategory ‘suffering’	137
10.2	Initial relationships between categories	140
10.3	Incorporating categories, and distinguishing more and less important lines of analysis	141
10.4	Reassessing relationships between categories – 1	142
10.5	Reassessing relationships between categories – 2	144
10.6	Reassessing position of categories in analysis	145
10.7	Revising analysis with minimum disturbance	148
10.8	Comparing subcategories of ‘substance’	148
10.9	Shifting the analytic emphasis	150
11.1	Single hyperlink between two bits of data stored separately	153
11.2	Multiple hyperlinks between bits of data stored separately	154
11.3	Linking dentists and patients	154
11.4	Observing the link ‘debunked by’ between databits	156
11.5	Linking and categorizing complement each other	157
11.6	Linking two databits	158
11.7	An explanatory link between two databits	160
11.8	Linking and categorizing two databits	160
11.9	Inferring an explanatory link between two databits	161
11.10	Explaining Mrs Sol Schwimmer’s litigation	163
11.11	Conditional and causal links in the tale of Kaufman and Tonnato	166
11.12	Connecting incongruous and cathartic humour	166
11.13	Linking data and connecting categories	167
12.1	The difference between associating and linking events	170
12.2	Association and linking as mutually related means of establishing connections	171
12.3	Following a trail of links through the data	181
12.4	Two trails of links through the data	181
12.5	Following a trail of different links through the data	182
12.6	A ‘chain’ of causal links in the data	183
12.7	Retrieving chronological links in the Claire Memling story	184
12.8	Vincent’s explanations linked to chronology of events in the Claire Memling story	185
13.1	Textual and diagrammatic displays of information	193
13.2	Map of relationship between two concepts	204
13.3	Map of complex relationships between four variables	204
13.4	The history of the universe through time	204
13.5	A small selection of symbols based on computer graphics	205
13.6	Differentiating concepts through different shapes and patterns	205

13.7	Incorporating detail by including subcategories	206
13.8	Adjusting for the empirical scope of categories	207
13.9	Mapping relationships for all cases	207
13.10	Comparing differences in scope through a bar chart	208
13.11	Using overlaps to indicate scale	208
13.12	Adjusting for scope in presenting classification scheme	209
13.13	Adjusting scope of most refined categories	210
13.14	Distinguishing exclusive and inclusive relationships	210
13.15	Making relationships between categories more explicit	211
13.16	Representing strength of different causal relationships	212
13.17	Comparing strength of relationships between categories	213
13.18	Integrating connections between categories'	213
13.19	Representing reciprocal connections between categories	214
13.20	Identifying positive and negative categories	214
13.21	Representing concurrence between categories	216
13.22	Using space to represent time	217
14.1	Concurrence between categories	226
14.2	Two routes through the data, arriving at different results	231
15.1	The whole is greater than the sum of the parts – 1	240
15.2	The whole is greater than the sum of the parts – 2	241
15.3	Tree diagrams representing different analytic emphases	243
15.4	Tree diagrams indicating different analytic emphases	244
15.5	Different writing strategies – sequential and dialectical	249
15.6	Decision-making laid out in algorithmic form	252
15.7	Procedures for assigning categories in algorithmic form	253
15.8	The two aspects of generalization	262
16.1	Linear representation of analysis	264
16.2	Loop representation of analysis	264
16.3	Analysis as an iterative process	265

ILLUSTRATIONS

1.1	Different approaches to qualitative research	2
2.1	Structured and unstructured responses to the question 'What are the main advantages and disadvantages of closed questions in an interview?'	16
2.2	Example of a grading and marking system	25
2.3	Grades with different mark bands	26
3.1	Personal ads	41
5.1	'The library'	64
5.2	Comments on feminist humour	68
6.1	'Two attendants at a Turkish Bath'	75
6.2	Recording data fully but inefficiently	77
6.3	Filing reference information – questions and sources	78

6.4	Data filed efficiently	78
7.1	'In the Office'	85
7.2	Using memos to open up lines of enquiry	89
7.3	Linking memos and data	90
8.1	Preliminary definitions of categories	103
8.2	Developing a more extensive category list	107
9.1	Two ways of identifying 'bits' of data	116
9.2	Overlapping bits of data	117
9.3	A preliminary category list	119
9.4	Checking memos prior to categorizing data	122
9.5	Contrasting definitions of the category 'temperament'	123
9.6	Inferring an emotional state from behaviour	123
9.7	Data stored following categorization of a databit	125
9.8	Categorizing Vincent's first letter	127
10.1	Comparing databits assigned to different categories	130
10.2	Databits assigned to the category 'suffering'	130
10.3	Subcategories of 'suffering'	134
10.4	Subcategorized databits for the category 'suffering'	138
10.5	Subdividing databits between subcategories	138
10.6	Comparing databits between categories	143
11.1	Possible links	156
11.2	Information held on linked databits	164

TABLES

3.1	Implicit classifications in everyday life	41
8.1	Alternative category lists	101
11.1	Result of linking and categorizing two databits	160
11.2	Multiple links between databits	162
11.3	Linking non-sequential databits	162
12.1	Concurrence between categories	172
12.2	Comparing databits between the different cells	173
12.3	List of indexed databits	173
12.4	Boolean operators for category retrievals	174
12.5	Retrieval based on categories assigned to proximate bits of data	175
12.6	Retrieval based on categories 'temperament' and 'suffering' assigned to proximate bits of data	176
12.7	Categories analysed as case variables	177
12.8	Cross-tabulating categories as case variables: 'temperament' and 'suffering' in Vincent's letters (N = 0)	177
12.9	Identifying connections between categories for databits assigned to category 'suffering' and databits linked to these by the link 'caused by'	186

12.10	Connecting 'X' categories 'transposing' and 'temperament' to 'Y' category 'suffering' through causal links between the databits	188
13.1	Comparing information across cases	194
13.2	Matrix with non-exclusive values	194
13.3	Using a matrix to explore variation in the data	195
13.4	Databits by case and category	197
13.5	Data indices by case and category	197
13.6	The number of assignments of each category by case	198
13.7	Recoding the data to express more meaningful values	200
13.8	Analysing subcategories as separate variables	200
13.9	Recategorizing variables as values of 'suffering'	201
13.10	Frequencies for the variable 'suffering'	202
13.11	Cross-tabulating 'occupation' and 'suffering'	202
15.1	Databits assigned to categories 'active' and 'passive'	257
15.2	'Passive' and 'active' responses by gender	258
15.3	Distribution of responses by case	259

Introduction

Q. What colour is snow?

A. White.

To most of us, the answer 'white' may seem satisfactory, but to an Eskimo it would seem a joke: Eskimos distinguish between a wide variety of 'whites' because they need to differentiate between different conditions of ice and snow. So it is with qualitative data analysis: in a recent review of the field, Tesch (1990) distinguishes over forty types of qualitative research (Illustration 1.1). Just as the Eskimos distinguish varieties of white, so researchers distinguish varieties of qualitative analysis. There is no one kind of qualitative data analysis, but rather a variety of approaches, related to the different perspectives and purposes of researchers. To distinguish and assess these different perspectives fully would be a formidable and perhaps rather fruitless task, particularly as the boundaries between different approaches and their relation to what researchers actually do when analysing data is far from clear. But is there a basic core to qualitative research, as there is a basic colour 'white', from which these different varieties are derivative?

Different researchers do have different purposes, and to achieve these may pursue different types of analysis. Take a study of the classroom, for example. An ethnographer might want to describe the social and cultural aspects of classroom behaviour; a policy analyst might want to evaluate the impact of new teaching methods; a sociologist might be most interested in explaining differences in classroom discipline or pupil achievement – and so on. Different preoccupations may lead to emphasis on different aspects of analysis. Our ethnographer may be more interested in describing social processes, our policy analyst in evaluating results, our sociologist in explaining them. This plurality of perspectives is perfectly reasonable, remembering that social science is a social and collaborative process (even at its most competitive), in which (for example) descriptive work in one project may inspire interpretive or explanatory work in another (and vice versa).

Illustration 1.1 Different approaches to qualitative research

action research	ethnographic content	interpretive interactionism
case study	analysis	interpretive human studies
clinical research	ethnography	life history study
cognitive anthropology	ethnography of	naturalistic inquiry
collaborative enquiry	communication	oral history
content analysis	ethnomethodology	panel research
dialogical research	ethnoscience	participant observation
conversation analysis	experiential psychology	participative research
Delphi study	field study	phenomenography
descriptive research	focus group research	phenomenology
direct research	grounded theory	qualitative evaluation
discourse analysis	hermeneutics	structural ethnography
document study	heuristic research	symbolic interactionism
ecological psychology	holistic ethnography	transcendental realism
educational connoisseurship	imaginal psychology	transformative research
and criticism	intensive evaluation	
educational ethnography		

Source: Tesch 1990: 58

Given the multiplicity of qualitative research traditions, one might reasonably wonder whether there is sufficient common ground between the wide range of research traditions to permit the identification of anything like a common core to analysing qualitative data. On the other hand, the very notion of 'qualitative' data analysis implies, if not uniformity, then at least some kind of family kinship across a range of different methods. Is it possible to identify a range of procedures characteristic of qualitative analysis and capable of satisfying a variety of research purposes, whether ethnographic description, explanation or policy evaluation is the order of the day? The relevance and applicability of any particular procedure will, of course, depend entirely on the data to be analysed and the particular purposes and predilections of the individual researcher.

Having identified a multiplicity of perspectives, Tesch manages to reduce these to three basic orientations (1991: 17–25). First, she identifies 'language-oriented' approaches, interested in the use of language and the meaning of words – in how people communicate and make sense of their interactions. Second, she identifies 'descriptive/interpretive' approaches, which are oriented to providing thorough descriptions and interpretations of social phenomena, including its meaning to those who experience it. Lastly, there are 'theory-building' approaches which are orientated to identifying connections between social phenomena – for example, how events are structured or influenced by how actors define situations. These distinctions are not water-tight, as Tesch herself acknowledges, and her classification is certainly contestable. No one likes to be pigeon-holed (by some one else), and nothing is more likely to irritate a social scientist than

to be described as atheoretical! However, Tesch does suggest a strong family resemblance between these different research orientations, in their emphasis on the meaningful character of social phenomena, and the need to take this into account in describing, interpreting or explaining communication, cultures or social action.

Thus encouraged, we can look for a basic core of qualitative data analysis, though not in some consensus about research perspectives and purposes, but rather in the type of data we produce and the way that we analyse it. Is there something about qualitative data which distinguishes it from quantitative data? And if qualitative data does have distinctive characteristics, does this also imply distinctive methods of analysis? My answer to both these questions is a qualified 'yes'. In Chapter 2 I distinguish between qualitative and quantitative data in terms of the difference between meanings and numbers. Qualitative data deals with meanings, whereas quantitative data deals with numbers. This does have implications for analysis, for the way we analyse meanings is through conceptualization, whereas the way we analyse numbers is through statistics and mathematics. In Chapter 3, I look at how we conceptualize qualitative data, including both the articulation of concepts through description and classification, and the analysis of relationships through the connections we can establish between them.

I said my answers were qualified, for though we can distinguish qualitative from quantitative data, and qualitative from quantitative analysis, these distinctions are not the whole story. We can learn as much from how meanings and numbers relate as we can from distinguishing them. In social science, number depends on meaning, and meaning is informed by number. Enumeration depends upon adequate conceptualization, and adequate conceptualization cannot ignore enumeration. These are points I take up in Chapters 2 and 3. My aim is to introduce the objects and methods of qualitative analysis, as a basis for the subsequent discussion of procedures and practice.

It is easy to exaggerate the differences between qualitative and quantitative analysis, and indeed to counterpose one against the other. This stems in part from the evolution of social science, most notably in its efforts to emulate the success of the natural sciences through the adoption of quantitative techniques. The fascination with number has sometimes been at the expense of meaning, through uncritical conceptualizations of the objects of study. Nowhere is this more apparent than in the concepts-indicators approach, where specifying the meaning of concepts is reduced to identifying a set of indicators which allow observation and measurement to take place – as though observations and measurement were not themselves 'concept-laden' (Sayer 1992). The growing sophistication of social science in terms of statistical and mathematical manipulation has not been matched by comparable growth in the clarity and consistency of its conceptualizations.

Action breeds reaction. In response to the perceived predominance of quantitative methods, a strong undercurrent of qualitative research has emerged to challenge the establishment orthodoxy. In place of the strong stress on survey techniques characteristic of quantitative methods, qualitative researchers have employed a range of techniques including discourse analysis, documentary analysis, oral and life histories, ethnography, and participant observation. Nevertheless, qualitative research is often cast in the role of the junior partner in the research enterprise, and many of its exponents feel it should have more clout and more credit. This encourages a posture which tends to be at once defensive of qualitative methods and dismissive of the role of the supposedly senior partner, quantitative research.

Beneath these rivalries, there is growing recognition that research requires a partnership and there is much to be gained from collaboration rather than competition between the different partners (cf. Fielding and Fielding 1986). In practice, it is difficult to draw as sharp a division between qualitative and quantitative methods as that which sometimes seems to exist between qualitative and quantitative researchers. In my view, these methods complement each other, and there is no reason to exclude quantitative methods, such as enumeration and statistical analysis, from the qualitative toolkit.

Reconciliation between qualitative and quantitative methods will undoubtedly be encouraged by the growing role of computers in qualitative analysis. The technical emphasis in software innovation has also encouraged a more flexible and pragmatic approach to developing and applying qualitative methods, relatively free from some of the more ideological and epistemological preoccupations and predilections dominating earlier discussions. The development of software packages for analysing qualitative data has also stimulated reflection on the processes involved, and how these can be reproduced, enhanced or transformed using the computer. The development of computing therefore provides an opportune moment to consider some of the main principles and procedures involved in qualitative analysis. I outline the general contribution of the computer to qualitative analysis in Chapter 4. In doing so, I take account of how computers can enhance or transform qualitative methods. This is a topic I address explicitly in Chapter 4, but it also forms a recurrent theme throughout the discussion of analytic procedures in the rest of the book.

On the other hand, software development has also provoked concerns about the potentially damaging implications of new technological forms for traditional methods of analysis. Some developers have emphasized the potential danger of the software they themselves have produced in facilitating more mechanical approaches to analysing qualitative data, displacing traditional analytic skills. This concern has highlighted the need to teach computing techniques within a pedagogic framework informed by documented analytic principles and procedures. Paradoxically, however,