

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/285471178>

Qualitative Research Methods: A Practice-Oriented Introduction

Research · February 2023

CITATIONS

123

READS

310,847

1 author:



Arthur Cropley

University of Hamburg

284 PUBLICATIONS 8,301 CITATIONS

SEE PROFILE

ARTHUR CROPLEY

Introduction to Qualitative Research Methods:

A practice-oriented introduction

University of Hamburg



ISBN: 978-606-95073-6-0

© by Arthur J. Cropley

Editura Intaglio, 2022

www.edituraintaglio.com

editura.intaglio@gmail.com

All rights belong to Editura Intaglio Publishing House.

No part of this publication can be reproduced, stocked or distributed in any way, in any form, physically or digitally, without the written permission of the publisher.

Table of Contents

The purpose of this book	6
Chapter	Page
1 The definitive nature of qualitative research	9
Contrasting Approaches to Research	12
Emergence of Modern Interest in Qualitative Methods	20
The Focus of this Book	27
2 The essential nature of all research	29
Fundamental Issues in Research Design	31
Classifying Research Designs	36
Doing Good Research	42
3 The special nature of qualitative research	54
The Essence of the Qualitative Approach	56
The Special Properties of Qualitative Research	60

Examples of Qualitative Research Questions	66
The Relationship of Quantitative and Qualitative Methods	74
Why do People Conduct Qualitative Research?	81
4 Designing a qualitative study	87
The Phases of Qualitative Studies	89
Factors Affecting Reliability and Validity in the Phases	94
Key Aspects of Qualitative Research	101
Defining the “Object of Study”	104
The Sample in Qualitative Designs	111
Ethical Issues in Research Design	117
5 Collecting data in qualitative research	122
The Nature of Data in Qualitative Research	124
Methods for Collecting Data	129
Case Studies	136
Interviews	142
6 Mixed approaches to collecting data	156
The Core of the Mixed Approach	158
Mixed Models	165
Multi-Methods	169
Triangulation	173
7 Basic issues in analysing qualitative data	177
Basic Principles of Qualitative Analysis	179
Computer-Assisted Qualitative Analysis	188
The Role of Theory in Qualitative Analysis	195
Coding Content Units	199
8 The process of analysing qualitative data	207
Working out the Meaning of a Text	210
Analytic Induction and Grounded Theory	217
Other Ways of Analysing Qualitative Data	225

A Concrete Example	234
Completing a Qualitative Analysis	239
9 Requirements for a rigorous qualitative report	241
Criteria of a Rigorous Report	244
Achieving Originality	252
Advancing Theory	257
10 Rigorous, accessible, and transparent reporting	260
Writing a report	262
Transparent presentation: A case study	266
Analyzing and Interpreting the Data	271
Two Contrasting Examples	279
References	284
Appendix A: Checking the quality of a thesis	297
Checklist	300
Appendix B: Helpful introductory texts with a practical orientation	303

The purpose of this book

There are numerous critical discussions of qualitative methods. For example, a readable recent analysis that focuses on weaknesses in qualitative data analysis is to be found in Niederberger and Dreieck (2020). However, the purpose of the present book is not to add to these critical voices.¹ My goal is to make the *practical* aspects of qualitative research understandable and usable. Accordingly, the book is not centered on a *theoretical* analysis or evaluation of qualitative methods from the point of view of the philosophy of knowledge: for example, their epistemology (their understanding of what constitutes knowledge and the quality and general applicability of the knowledge they generate) or the rigorousness and standardisation of their heuristics (the techniques they use to acquire what they regard as knowledge).² Instead, I focus on

¹ However, I do not wish to underestimate their importance, either.

² Nonetheless, I found it impossible to ignore these issues entirely: See in particular Chapters 1 and 2.

everyday *practice*: I explain what concrete activities qualitative research encompasses and what advantages its advocates argue it brings. Why do people do qualitative research, what do they do, how do they do it, and is it OK?

- Why do people do such research?
- How do they go about doing it?
- What is special about the results it leads to?
- How can the results be presented in a convincing way?

The book aims at achieving effects in three domains: (a) the personal, (b) the scholarly, and (c) the practical. The *personal* goal is to demystify qualitative methods, give readers a feel for what can be done with qualitative research, and empower them to carry out qualitative studies. For this reason, it includes many down-to-earth examples couched in relatively everyday language. Many (although not all) of these come from my experience as a researcher. In the *scholarly* domain the book seeks to provide readers with a set of concepts for understanding the essence of qualitative methods and to familiarise them with specialized terminology for thinking, talking and writing about such methods, so that, for instance, they can explain and justify what they are doing in a professional way. Finally, in the *practical* domain, the goal is to show readers how to design a qualitative study, collect qualitative data, analyse the data and report the results in a scientifically rigorous manner. My intention is not to argue that qualitative methods are exempt from examination of higher issues such as their epistemology, but to start by saying clearly what they involve in practice. The ultimate goal is that at the end readers will be able to understand what qualitative researchers do and why and see where discussions of whether and how qualitative research meets traditional standards of scientific rigour could start. In addition, readers will know how to conduct and write up such research.

The book's contents involve:

- contrasting qualitative methods with quantitative approaches in order to highlight their central ideas,

- illustrating with examples the kinds of question that can be asked using qualitative methods,
- clarifying the nature of qualitative data,
- explaining how to collect such data,
- showing how to analyse qualitative data,
- providing guidelines on reporting the results of a quantitative study in a scholarly way.

At the beginning of each chapter there is a summary of the main points made in the chapter. In addition, there is a glossary of terms I would like readers to remember and use appropriately in the future. In addition, Appendix A offers a checklist for self-assessment of the quality of a qualitative report and Appendix B suggests further reading in order to obtain detailed practical advice.

Chapter 1: The definitive nature of qualitative research

The core property of qualitative research is that it examines the way people make sense out of their own concrete, real-life experiences in their own minds and in their own words and subsequently analyses these understandings in the sense of a behavioural science such as psychology, sociology, politics, education, health sciences or, nowadays, business and management, decision making, or innovation, to give some examples. When asked about how they understand some particular aspect of their lives people usually answer in everyday language using everyday concepts. The data that result are referred to as “narratives” (e.g., remarks in an interview, statements in a written text, but also videos, even works of art). Quantitative research, by contrast, focuses on the way the world is understood in researchers’ minds, usually using abstract concepts and terminology,

and the data usually consist of scores on measuring instruments of some kind (in psychology and education these are frequently tests or questionnaires). A second major difference is that qualitative research is concerned with the patterns and forms of variables, whereas quantitative research examines differences in amount or level of the variables being studied and cause and effect relationships among these variables. Although qualitative research was out of fashion for a long time, modern interest in it represents the re-emergence of an approach that has as long a history in social science as quantitative methods. Despite the existence of “paradigm wars,” the two approaches are not rivals, but are complementary.

Chapter Glossary³

<i>emergent issue:</i>	an issue (often in the form of an unexpected preliminary finding) that is identified in the course of a piece of research and may surprise the researcher
<i>incompatibility thesis:</i>	the view that qualitative and quantitative methods are irreconcilable enemies (not shared in this book)
<i>introspection:</i>	a method of obtaining data by asking yourself about your own thoughts and behaviour in relation to some topic
<u><i>mixed methods:</i></u>	research designs in which qualitative

³ Each chapter in the book starts with a glossary of terms used in the chapter that are useful for enlarging your technical vocabulary. The definitions given here relate to the way I use terms in this book, and are not necessarily universal or even widely-used.

	and quantitative methods are combined (see Chapter 6)
<i>narrative:</i>	an account of real-life events given by a person who was (or is) involved in the events; narratives are most commonly constructed in the context of an interview, although other forms are possible. These include diaries or letters or even non-verbal forms of communication such as works of art
<i>narrative psychology:</i>	an approach to psychology based on people's accounts (narratives) of how they understand the world
<i>open item:</i>	an item in a questionnaire which allows respondents to formulate their own answer rather than requiring them to choose from among alternatives offered by the researcher
<i>positivism:</i>	the view that psychological states directly cause each other and that the strength of a causal relationships can be expressed in numbers
<i>social construction:</i>	the process through which people take over a view of what is "real" from other people rather than forming it from their own experience (see next glossary item)
<i>subjective reality:</i>	a person's individual view of what is "real," constructed on the basis of the person's own interpretation of his or her own experience with the world

Thomas theorem:

the view that what people believe to be real affects their thinking just as much as if actually it were real

Contrasting Approaches to Research

Simple everyday observation of people going about their lives shows that there are differences from person to person in behaviour: We do not all react in an identical, pre-programmed way like robots to any given circumstances, but in ways that vary from person to person. *Enlarging knowledge and understanding of these differences* is the task of *all* research in the social and behavioural sciences, regardless of whether it is qualitative or quantitative. The process of “enlarging knowledge and understanding” of people’s behaviour involves (a) identifying and describing differences from person to person or within a given person at different times or in different situations (e.g., defining the nature of differences, describing their effects, etc.), (b) giving the differences names, which usually come from the discipline in which the researcher is working (in the case of psychologists, for example, “feelings,” “motives,” or “self-image,” in the case of business studies, by contrast, “organizational structure,” “power relationships,” or “role conflict”), and (c) saying what causes the differences, how they relate to each other, what factors influence their development, and so on. For example, it is commonly observed that children differ from each other in how effectively and how quickly they learn in class: What is the precise nature of such differences? Where do they come from? What can we do about them?

Quantitative and qualitative research represent two differing ways of seeking answers to such questions. Of course, different disciplines explain differences in achievement in different

ways. For example, psychologists attribute them to factors such as cognitive processes, motivation, attitudes and values, or self-image, whereas sociologists emphasize group membership, socio-economic status, power relationships, gender, and the like. These discipline-related differences are discussed in greater detail on pp. 32-33. However, this book focuses on methods of investigation, not on differences in the conceptual frameworks of specific disciplines, so that differences in the concepts underlying different disciplines will not be pursued in detail here.

Quantitative research: Quantitative research starts by defining differences among people in terms of “standard” variables that apply to everybody, although different people display differing *levels or amounts* of the variables (such as intelligence or extraversion or tolerance for ambiguity). The evidence for the relevance of these “causative” variables is usually derived from *existing* knowledge, for example from earlier research: Reports on new research – such as dissertations – almost always start with a lengthy review of what is already known. Quantitative research usually then measures *how much* of each variable that has been identified as important different people have, often using highly standardized tests or questionnaires, and finally seeks to establish the existence of cause-and-effect relationships among the variables. In the example of classroom learning already mentioned, observed differences among people could be expressed in terms of pre-defined outcomes of classroom instruction such as mastery of the mother tongue, skill in using numbers, knowledge of local and world history, and the like (usually labelled “achievement”). These differences could then be *measured* using achievement tests with standard correct solutions from which points scores could be derived (i.e., achievement tests): *How much* language skill, facility with numbers, or knowledge about world history does each child in the class possess?

Now the question arises: What *causes* any differences that are observed: Why does one child know more than another? A common

psychological explanation is to invoke the concept of “intelligence.” This has a universal, so to speak “official” meaning defined in advance by psychologists and educators (not by schoolchildren) and *applies to everybody* – everybody has some of it. It involves things like ease of acquisition of knowledge of facts, speed of recall of information, skill at logical thinking, ability to concentrate, and similar abilities and skills. Everybody knows at least some facts, everybody takes some time to recall facts – some a lot of time, some only a little – everybody can concentrate, to some extent at least, and so on. Standard tests provide numerical information on how much of such skills and knowledge each individual person possesses (frequently, IQ scores).

Quantitative data: On the basis of formal measuring instruments specifically constructed to *measure* achievement on the one hand and intelligence on the other, each individual participant in a study can be given a numerical score on each of the variables involved. In the present example these might be (a) the *cause* of differences (intelligence expressed numerically as an IQ) and (b) the *magnitude* of those differences (expressed as an achievement score). These scores are based on a formal procedure specified by the experts, not for instance by participants themselves. For example, observable differences among the children in classroom learning could be measured using standardized tests, and expressed in a numerical form such as percentages or grade point averages of one kind or another, or at least in the form of an ordinal scale such as A, B, C, D, E. It is also possible to do the same for the children’s intelligence, using existing standardized intelligence tests to obtain an IQ score for each child. Subsequently, calculations can be carried out to determine whether the two sets of scores (in the present example, an IQ score and a grade-point average) go together: Are IQs associated with school grades and how strongly? The next step is to infer a cause-and-effect relationship of some kind (for example, IQ *causes* school achievement). This paradigm characterizes the quantitative approach to research in the social sciences and seems

to be highly rational.

However, the case study of a specific large-scale research project presented in the box that follows demonstrates the blind spots of the approach.⁴ Almost bizarrely, having invested considerable time and energy in a no doubt well-planned and professionally conducted and analysed quantitative study, the chief researcher stated in a newspaper interview (*The Australian*, Wednesday April 10, 2013, p. 29) that it would probably never be possible to measure the factors identifying and quantifying the differences among children objectively, and therefore impossible to learn more about them. In other words, according to this person *strict quantitative approaches are not and never will be capable of getting at the real heart of the matter being investigated*. Why is this? What alternative approach is there? What special insights can the alternative approach offer where quantitative approaches fail to get at the real issues? What is special about the alternative approach? How is research conducted using the alternative approach?

In the hope of identifying the school-specific variables which cause differences in students' grades in their final year of high school researchers studied no fewer than 14,000 students who had just completed secondary school. Variables studied included school *sector* (public versus private; higher fees versus lower fees), *location* (e.g., affluent versus low-income suburbs), *resources* (e.g., teacher salaries, student-teacher ratio), *demographics* (e.g., proportion of students from migrant families), and *gender mix* (single sex versus mixed sex). All of these can be objectively specified and quantified, and according to earlier research they all have an effect on academic achievement. The results showed that, in fact, although not irrelevant these variables accounted for a total of only about one-third of the variance of achievement – the researchers'

⁴ The study was presented at a national forum on research findings derived from the Longitudinal Surveys of Australian Youth (LSAY) in 2013 (NCVER, 2013).

modelling of the causal effects of such factors on achievement could not account for any higher a proportion. In other words, the factors specified by the researchers only told part of the story. The researchers speculated that the truly decisive properties of schools which affect students' achievement are variables like *pattern of educational leadership* or *quality of teaching* but concluded that it is impossible to investigate such factors because there is no way of *measuring* them *objectively*.

The qualitative idea of “reality”: Qualitative research attacks issues differently from the way outlined above. It is based on the fundamental idea that “reality” is subjective: Every human being constructs an individual, personal view of the way the world works on the basis of his or her specific interactions with it (including the people who are part of it). As a result, much of what people, including researchers, regard as reality actually consists of a set of impressions, inferences and opinions in each individual person's mind. As Woodman (2014, p. 465) put it, social sciences:

... posit that people carry *implicit theories* in their minds concerning almost all aspects of their reality and their behavior is based (at least in part) on these theories or “understandings” about how things are related to each other, how the world works, why others behave as they do, and so on. [emphasis added]

Hawking and Mlodinow (2010, p. 39) went further by arguing that, from the point of view of people interacting with their surroundings, there is no such thing as an objective reality: what an individual person regards as “reality” is actually a model constructed by that individual on the basis of observation and interpretation of events. However, it all seems real to the person in question. This point of view was summarised in a down-to-earth way many years ago by the pioneering sociologist, William Thomas, in what became known as

the “Thomas Theorem” (Thomas & Thomas, 1928, p. 572): “If men [sic] define situations as real, they are real in their consequences.” Put somewhat differently: What people regard as real might as well be real, since it affects the people as though it really were real.

The point of view based on writers such as Woodman, Hawking and Mlodinow, or Thomas does not imply that we can make things be any way we like merely by imagining or believing they are that way: For example, a person who jumped out of a twentieth floor window would not, in the usual course of events, float gently down to earth, no matter how strongly the person in question believed that he or she could fly. What is subjectively constructed is the way people *make sense* out of events they observe – for example how they explain why the person who jumped out of the window fell. Most readers would know from observation of the world that without physical support people or objects drop, and would (correctly) predict that a person who jumped out of a window would plunge to the earth. But why is this? Like me, most people would probably attribute plunging to the ground to a force we call “gravity.” In fact, gravity is a construction for making sense out of our observation of what happens when objects are unsupported and, even if future science were ultimately to prove that there is no such thing as gravity, imagining that it exists has so far proved to be very helpful for dealing with the world, so it might as well be real (Thomas Theorem).

Social construction of reality: There is a second important aspect of the idea just outlined that each person has a personal, constructed reality based on interactions with the external world: Many constructions for making sense out of the world are not worked out individually on the basis of direct personal experience, but are taken over from other people. They are *socially constructed*.⁵ Gravity is an example; most people who believe that gravity exists

⁵ This issue is also discussed in other places (for instance p. 95ff.).

will not have thought it up for themselves but will have been told about it by other people, such as teachers. Of great interest here is that social constructions may not be universal: Different groups and sub-groups may construct the same aspect of the observable world differently. However, different constructions may be equally helpful for getting along in life; that is, there can be more than one “reality,” and alternative “realities” may all work quite well, no matter how absurd they look to groups not sharing them, as the following case study shows.⁶ The task of **qualitative** research is *to gain insights into these constructions of reality*, i.e., to tease out the nature of the world *as it is experienced, structured and interpreted* by people *in the course of their everyday lives* (including their interactions with other people).

In the traditional Trobriand Islands culture before the arrival of Europeans it was believed that a female became pregnant when a spirit entered her body and took on physical form in her womb. Men had nothing to do with the process. However, practical observation of the world showed that women often become pregnant after lying naked with men. The construction placed on this observation of the external world, and – prior to the arrival of Europeans – shared by everybody in the islands, was that it is easier for the spirit to enter a woman’s body if the entrance to her womb is open, and men are helpful for achieving this. On the basis of this construction, Trobriander women who did not want to become pregnant made use of an effective method of contraception: they avoided having sex with men. Despite its complete ignorance of the biology of sexual reproduction, the system worked quite well, and when missionaries explained the North American-European understanding of how women get pregnant, the Trobrianders found it absurd and laughable.

⁶ This example is based on reports by the anthropologist Bronislaw Malinowski (1929), although for the sake of simplicity I have taken some liberties with the details of his account.

Qualitative data: The information qualitative researchers collect in order to gain insights into people's constructions of the world is obtained by means of *narratives* in which the people being studied *communicate* the way they understand the world. Narratives most obviously involve speaking or writing, but they can be based on other forms of communication such as, for example, diaries, literary or journalistic works, audio or video tapes, or artworks, or can even result from close and systematic observation of people going about their lives in their usual manner (i.e., observation of behaviour). It is common in qualitative research for narratives to involve interviews, although this is not always the case. In fact, sometimes there is not even direct contact between the researcher and the people being studied, as in the example in the following box.⁷

Thomas and Znaniecki (1927) studied written autobiographies and letters home, in which Polish immigrants to the USA presented their own life histories and explained to relatives back in Poland what was happening to them in the new country. The researchers also examined newspaper reports and the archives of Polish organisations in Poland and Chicago. Thomas and Znaniecki analysed the contents of the documents by working out the more general *meaning* of what individual immigrants said about their lives and how they made sense out of what was happening to them. The researchers identified (a) typical patterns in the way people interpreted the events of their lives and (b) patterns of adjustment based on the way the people dealt with their situation once they had interpreted it. The researchers then divided individual people into "types" based on various combinations of these two dimensions (way of interpreting events and way of dealing with situations once they had been interpreted): the philistine type who rigidly avoids change, the bohemian type, who recklessly accepts

⁷ Several other examples of qualitative studies are given on pp. 66ff.

any change, and the creative type, who is discriminating in accepting some change but not all. (This form of analysis is referred to nowadays as *types and patterns analysis* and will be discussed in greater detail in Chapter 8, p. 230.)

Emergence of Modern Interest in Qualitative Methods

Recent discussions of research methodology in the social sciences have devoted a great deal of attention to qualitative approaches, and some writers have discussed qualitative methods as though they were new and revolutionary. However, despite the long tradition of quantitative research in the social sciences, qualitative methods have been used for at least as long as quantitative ones. In the 19th Century, when modern empirical behavioural science was beginning to take form, considerable use was made of qualitative methods both prior to the emergence of quantitative approaches and also more or less parallel to them. An example is *introspection*, in which researchers examined their own subjective experiences in order to obtain insights into how people make sense of themselves and the world around them. By the beginning of the 20th century there was already a considerable literature on this method (for a discussion see Titchener, 1912).

The (temporary) fall of qualitative methods: Around 160 years ago Fechner (1860) and colleagues in Germany and other countries hoped to discover objective, immutable, mathematical laws about human behaviour that would transcend the foibles of the individual person and make behaviour understandable in terms of *universal mathematical laws*, and thus predictable and controllable. An example is the Weber-Fechner law defining

the relationship between stimulus intensity and people's ability to detect changes in a stimulus ($S = k \cdot \log R$; the more intense the original stimulus, the greater a change in it must be in order for a human being to detect it – we do not detect tiny changes in the loudness of very loud sounds, for example, but do notice even small variations in soft sounds). Early researchers not only expanded the array of research methods being used in their own time, but also introduced quantitative methods that have endured to the present. Their approach focused on (a) properties capable of being measured and expressed numerically, (b) instruments for making measurements of these properties, and (c) methods for analysing the resulting numerical data.

For example, the Englishman, Francis Galton, designed and built instruments for objectively measuring pitch discrimination and similar sensory and motor abilities. Galton not only designed and built the special measuring instruments, but also developed techniques for *quantifying* abstract or intangible properties of human beings: among other things, the degree of exceptionality of performance (by *how much* do people differ from other people in achievement?) and the degree of closeness of family relationships. Quantification of these two variables made it possible for him to study the inheritance of intelligence (which he called “genius,” since the modern concept of “intelligence” did not yet exist). To enable him to express the *amount* by which variables were related to each other, he also worked out the basic ideas of what is now called “correlation,” and his protégé, Karl Pearson, developed the Pearson Product-Moment correlation coefficient, which is still one of the fundamental statistical procedures in use today.⁸ Throughout the considerably more than 100 years since the work of these pioneers, researchers in human sciences have used quantitative methods, and for much of that time this approach has been dominant.

⁸ Among other things, Galton used numerical methods to assess the efficacy of prayer, and came to the conclusion that, if anything, it had the opposite effect from the one intended.

The survival of qualitative methods: Nevertheless, qualitative approaches did not entirely disappear, so that recent renewed emphasis on qualitative methods does not constitute the emergence of something that has never been seen before. Rather, it involves a resurgence of interest in approaches that have long been known and employed. This fact does not detract from the usefulness of qualitative methods or make them any less valuable for looking at certain research questions (for examples of such questions see pp. 66ff.). However, it does weaken the apparent attitude of some writers who seem to suggest that qualitative methods represent a recent and newly discovered approach to research.

In anthropology and sociology, researchers' interest in the qualitative approach persisted longer than in psychology, and it thus has a stronger influence on modern thinking there than in psychology and education. This is reflected in the fact that these disciplines dominated early modern discussions of qualitative methods. It is true that sociology experienced a "quantitative revolution" starting in 1935, when the "Columbia School" gained the ascendancy over the "Chicago School," and sociological research experienced about 25 years of domination by the quantitative methodology familiar to psychologists. However, qualitative methods have again established themselves in that discipline. To take an example, the modern idea of "grounded theory," probably the most widely used approach to analysing qualitative data, which will be discussed in more detail in Chapter 7 (p. 219ff.), dates back more than 50 years in sociology (e.g., Glaser & Strauss, 1967), although it would be regarded as relatively novel by many psychologists.

As Elliott (1999, p. 252) remarked, referring in the main to psychology, especially clinical psychology, until fairly recently "the phrase 'qualitative data' was generally used to refer to the analysis of nominal scale or categorical data." In other words, many psychologists regarded the qualitative approach as no more than a set of procedures for dealing with the unfortunate case of data that were unsuitable for parametric statistical analysis, for instance

because they were not interval data or did not derive from a normally distributed continuous variable. Renewed and greatly broadened interest in qualitative methods among psychologists is part of an interdisciplinary trend to be seen in modern social sciences. The acceptance of feminism as an approach to research strengthened this trend, as did the postmodernist movement. Heron (1992) called for “psychology in another key,” focused on “personhood” and the study of intentions and purposes. He expressly concluded that quantitative methods do not permit such an orientation. Some theorists even recognize the existence of a new sub-discipline of psychology, which they call “narrative psychology” (e.g., Sarbin, 1986).

Although qualitative approaches in the sense just described may still be relatively uncommon in psychology and educational science, it has become evident that qualitative ideas have become better known in recent years in these disciplines. Evidence of this is, among other things, the growing emphasis on qualitative methods in widely differing fields of research, such as psychoanalysis – where Schaefer (1992) emphasized the usefulness of autobiographies – and cognitive psychology. In the latter area, Bruner (1990) came to the conclusion that a self-designed, selectively organized and logically structured internal “biography” of a person’s own life is the core principle from which organized behaviour results. Humanistic psychology (Maslow, 1962, Rogers, 1980) emphasizes that people make their own lives, even if they need help to do so in a healthy way.

The paradigm wars: Qualitative and quantitative approaches to research in the human sciences are often contrasted as bitter opponents. In particular, proponents of qualitative research not infrequently seem to make an almost messianic claim that it represents the path to true enlightenment, whereas quantitative approaches define the road to damnation. In discussing this situation, Tashakkori and Teddlie (2010) amusingly referred to the “paradigm wars” (p. 3), and identified “paradigm warriors,” who

seem to be engaged in a fight to defeat quantitative methods and replace them with qualitative approaches. Tashakkori and Teddlie gave examples from both education and psychology. In the case of educational research Smith and Heshusius (1986) argued that qualitative and quantitative methods are *irreconcilable*. Lincoln and Guba (1985) appeared to seal the fate of quantitative methods by concluding that they are no longer acceptable. Thus, qualitative approaches seemed to have won the day.

However, it is not appropriate to talk of a “triumph” of qualitative methods over quantitative, although the former are considerably more prominent nowadays than in the past. In fact, the recent resurgence of “open science” (e.g. Open Science Collaboration, 2015) suggests the opposite of a triumph: many journals are increasingly regarding qualitative research reports as unpublishable, because data are not collected specifically to test a pre-formulated and pre-registered hypothesis, there is no control group and no random, double-blind assignment of participants to experimental and control groups, no objective scoring, and so on.⁹ Nonetheless, it is important to remember that *both qualitative and quantitative methods of doing research are useful*. As Glaser and Strauss (1967) wrote in one of the most famous papers on qualitative methods, “... each form of data is useful for both verification and generation of theory (p. 17).” A good 30 years later, Strauss and Corbin (1998) drew attention to the increasing use of research designs involving combinations of qualitative and quantitative methods and concluded that grounded theory can readily be applied in either quantitative or qualitative research. In fact, there is growing interest in research paradigms which employ qualitative and quantitative approaches in tandem – so-called “mixed” approaches – and these will be discussed in considerable detail later (see Chapter 6).

Thus, while it is true that this book contrasts the qualitative

⁹ For example, after a change of editor, the *Creativity Research Journal* no longer publishes qualitative research.

approach with the quantitative one, it does not adopt an adversarial position. Qualitative research and quantitative research are not seen as rivals, each trying to establish itself as the one true way. On the contrary, they are complementary or supplementary approaches. Both kinds of research are capable of adding to knowledge, although with differing strengths and weaknesses. Thus, the *incompatibility thesis* advanced by Lincoln and Guba (1985) is rejected. As will become more obvious later, the book also does not accept a second idea put forward by some radical qualitative theorists. This is the view that there is “... no difference between [qualitative research] and fiction” (for more detail, see Taylor, Bogdan & DeVault, 2015, p. 23). Thus, the book does not reject or ignore traditional, scientific norms in favour of flights of literary fancy. As I have pointed out elsewhere, qualitative research requires special care in reporting and makes particular demands of readers if it is to avoid becoming a form of artistry, whereas quantitative research is seen as more a matter of artisanship.

Weaknesses of quantitative research: To enthusiasts, qualitative research may even seem to be uniquely able to cast light on truly interesting questions, and to do this in a way that respects the humanity of the respondents, whereas quantitative approaches are presented as being less spiritually worthy and less able to deal with really worthwhile questions about real human beings. Indeed, it is worth noting that some elements of quantitative research design (e.g., “split plot” analysis of variance) were actually developed for research in agriculture, while traditional quantitative approaches can as easily be applied to research with animals or even insects as with human beings. Heron (1992) went so far as to argue that quantitative methods are incapable of studying “persons,” because a person is “self-determining”: In other words, quantitative methods are said to overlook *intention*, *purpose* and *understanding* in behaviour. Indeed, they can be used with cabbages, flatworms or bees, whereas this is not the case with qualitative approaches!

In fact, the claim that quantitative research somehow fails

to grasp the real essence of the people it studies is not completely without substance. The people being investigated sometimes do seem to become irrelevant very quickly: For instance, the first step in the data analysis in multivariate quantitative studies often involves calculation of a correlation matrix. In such matrices neither rows nor columns contain people any more – they have vanished almost at the very beginning! This feeling of anti-climax, almost of disappointment is not uncommon in master's and doctoral students who conduct quantitative studies involving close personal contact with participants, even when the data analysis leads to an unequivocal rejection of the null hypothesis, normally the desired result and a happy event. They often express the feeling that the real object of their interest has somehow slipped out of sight, leaving them with a “clean” result, to be sure, but one that is only a fragment of what they were really interested in. Both researcher and respondents have, as it were, been “written out” of the research, almost from the very beginning. In the example that follows (see the following box), the whole issue of training creativity had been reduced to the question, “Is there a difference in the mean scores of the trained and untrained groups?” This question was answered by testing the null hypothesis: “There is no significant effect of training on creativity,” and rejecting it. Missing entirely were matters such as what acting creatively meant to participants, how they understood what had happened, or whether they felt different.

In a study involving the playing of games thought likely to facilitate creativity, a researcher spent 10 hours with each of about 30 people in either a control group or an experimental group playing various games with them, a total of about 300 person hours. Pre- and post-test data on various psychological aspects of creativity were collected and compared (experimental group vs. control group) by means of analysis of variance. Entering the data into SPSS took less than an hour, and the calculation of the F-values a few seconds. The value of the

F-statistic for the interaction between treatment and group was significant - the null hypothesis could be rejected. This meant that the research had been “successful”: After the treatment there were differences between trained and untrained subjects that had not been present prior to the training. However, instead of being pleased the researcher felt somehow flat and disappointed. So many meetings with so many people had somehow been reduced to a single F-value. What had become of the 300 person-hours of conversations, the human contact, the getting to know the subjects? Most of the information she had obtained in the course of her research had simply vanished. She voiced her feeling in one brief sentence: “Is that all?”

The Focus of this Book

It must be recognized that there is no single way of collecting data that all writers would agree is exclusively *the* way qualitative research is done. The present book:

1. gives considerable emphasis to conducting and analysing interviews, because these are most frequently used in qualitative social science research, although other procedures are not ignored;
2. emphasizes methods of analysing and reporting data that generate theory, or at least cast light on existing theory, i.e., that go beyond mere description;
3. focuses on making qualitative studies satisfy standards of scientific rigour.

Qualitative research has three main weak spots. These involve the problems of

- a) *arbitrariness* (qualitative research seems to some to

depend as much on the inventiveness and fantasy of the researcher as on hard facts);

b) *lack of objective proof* (qualitative methods do not permit the testing of hypotheses via objective - statistical - procedures);

c) *banality* (qualitative research is accused of doing no more than repeating what everybody already knows!).

The main issue, then, is how to deal with these weak spots in a way that meets the standards of scientific rigour. The following two chapters will provide a framework for comparing approaches to research and will show how quantitative approaches fit into the framework. Chapters 4–7 will look at these issues and outline methods for dealing with the weak spots in a rigorous way.

Chapter 2: The essential nature of all research

The aim of all research is to enlarge knowledge. In some disciplines, such as theology or philosophy, this process starts with fundamental tenets, which are taken to be self-evidently true. Knowledge is enlarged by showing that observed behaviour really does fit the fundamental tenets, even if it does not seem to do so at first: In other words, such research starts with the answer. In the case of empirical research, by contrast, knowledge is enlarged by observing the world in a systematic manner (i.e., collecting data), analysing what has been observed, drawing conclusions on the basis of this analysis, and publishing the findings: In other words, starting with the question. This is true of both quantitative and qualitative research. Both approaches to doing research involve a technology for carrying out the process just outlined in a way that yields rigorous, “robust”

knowledge. How, then, do they differ? The key differences between the two can be understood in terms of six basic dimensions: design, setting, method of data collection, kind of data, approach to data analysis, and strategies for making generalizations. The criteria of rigour in empirical research are reliability and validity although, when applied to research methodology these terms have a somewhat different meaning from what they mean when applied to tests, questionnaires, and the like.

Chapter Glossary

<i>action research:</i>	research aimed at understanding and dealing effectively with a specific issue or problem in a specific situation
<i>control:</i>	elimination of the effect of an irrelevant factor in a research study, often achieved by means of sampling (see glossary item below); not an issue in qualitative research
<i>ecological validity:</i>	the question of whether findings apply to the real world or only to settings contrived by a researcher.
<i>emic research:</i>	research aimed at analysing a situation in terms specified by the people experiencing it
<i>etic research:</i>	research aimed at analysing a situation in terms specified by the researcher
<i>macro-analytic research:</i>	research aimed at investigating a broad context

<i>micro-analytic research:</i>	research aimed at investigating a specific setting
<i>mixed methods:</i>	research designs involving a combination of qualitative and quantitative methods
<i>optimization:</i>	achieving the best combination of reliability and validity possible under specific circumstances
<i>reliable research:</i>	research findings which would be replicated by a different researcher working with different participants
<i>robust research:</i>	research findings which are not just fantasies of the researcher, but stand up to critical examination
<i>sampling:</i>	the selection of people for participation in a research project; done according to different principles in quantitative and qualitative research
<i>valid research:</i>	research which reveals genuine relationships among variables (internal validity) and/or which yields findings which apply in real life (external validity)

Fundamental Issues in Research Design

The task of all research is to *enlarge knowledge and understanding*. This task involves (a) identifying and describing phenomena (e.g., saying what their nature is, how they operate, giving them a name, etc.) and (b) improving understanding of their dynamics (e.g., saying where they come from, how they relate to each other,

what makes them change or develop, and so on). This is true for research in the social and behavioural sciences just as it is for any other area of science, although the social and behavioural sciences focus on people and the way they organize their lives, not on for example, changes in rock formations over millions of years or the behaviour of electrons when they are bombarded with energy. The explanations offered by researchers from different sciences differ from each other along a number of dimensions, as will be shown in the following sections.

External explanations: It is possible to explain differences among people by attributing the differences to the external environment. For instance, it seems rational to conclude that two children differ in reading ability because one child had in the past had greater access to books than the other. Even where some account is taken of factors “inside” the people in question, this could be restricted to looking at differences in biological structure and function, such as sensory perception – obviously, someone who is deaf would respond differently to auditory stimuli from someone with acute hearing. Indeed, some social sciences (e.g., sociology, anthropology) have developed explanations of differences from person to person that focus mainly or almost exclusively on the external world.

The explanatory concepts employed by different social sciences to explain the factors leading to differences from person to person differ from discipline to discipline. Examples of discipline-related differences are given in the following box:

Very broadly stated in an over-simplified way, the goal of sociological research is to identify the properties of communities that lead to the behaviour of people who are members of the community. Such properties are thought to guide behaviour more or less independently of the individual characteristics of the people involved. The sociological approach focuses on two central issues: (a) “proximate structural characteristics”

[emphasis added], on the one hand, and (b) “mediating processes of community social organisation” [emphasis added], on the other (Sampson & Wilson, 1995, p. 45). Proximate structural characteristics include factors such as socioeconomic status, level of education, housing, mobility, family structure, gender, age, or ethnic origin. Mediating processes include formation of groups, transmission of goals and values within groups, or enforcement of such goals and values by groups. Thus, sociology emphasizes social construction of reality. In a nutshell, anthropology studies the knowledge, beliefs, myths, symbols, moral tenets, customs, games, and similar factors that are shared by the people in a community (e.g., Harris, 2006). These can be summarised as involving culture, tradition and ritual. They promote social organisation and guide interactions among people. They are acquired through, among other things, child-rearing practices, and are transmitted by communication between people, both verbal and non-verbal, as well as by example (i.e., they are socially constructed). People exposed to different environments acquire different knowledge, beliefs, customs and the like, so that cultures in the sense just sketched out are local in nature.

To take a concrete example of the external approach, economic, sociological and political research on migration focuses on external factors such as poverty, war, famine, political or religious repression, family disruption, rapid social change, unemployment, and the like, thus seeing migration behaviour as a reaction to external pressures and privations. This way of looking at migration behaviour was criticized by Cropley and Lüthke (1994) for ignoring the role of the individual person’s personality. As these authors pointed out, migration occurs in all societies under all external conditions – provided that it is physically possible – and is often at its highest when external conditions are highly *favourable*. Furthermore, many emigrants come from comfortable

circumstances and not from the underprivileged or dispossessed. Thus, the idea that in deciding to emigrate (or not) people respond in a strictly logical and calculated way to favourable or unfavourable external circumstances does not hold up.

Internal approaches: By contrast, Cropley and Lüthke (1994) emphasized the importance of *internal* approaches that take account of structures and processes within people that regulate the way they react to events in the external world. A large body of appropriate concepts exists. In psychology and education researchers and theorists have developed ideas like “personality,” “perception,” “cognition,” “motivation,” “attitudes,” and many others to explain the internal factors that lead to behaviour. Psychological research is largely, although not exclusively, concerned with *describing* such *internal* factors and processes (their nature, magnitude, origins and development, method of action, and the like), and with *showing how they interact with each other*. The theoretical goal of these research activities is to enhance understanding of the nature and origins of “internal” differences among people. The practical goal is to help people cope with their lives better and, in the case of education, to learn better in schools and similar settings.

Etic versus emic approaches: A distinction that is closely related to “internal” and “external” involves the difference between collecting data (a) in the form of *self-selected and self-organized* statements made by people who are knowledgeable about a particular situation which is of interest to a researcher (narratives) and (b) in the form of respondents’ reactions to statements about the situation formulated and organized by the researcher (often by means of questionnaires or similar instruments). This distinction is well known in anthropology and ethnology, but less widely recognized in psychology or education. Pike (1967) referred to the difference as involving “etic” versus “emic” research. A more recent discussion, largely from the point of view of anthropology, is to be found in Friedman and Schustack (2012).

For the purposes of this book, an etic investigation can

be defined as being carried out *from the outside*: The researcher selects concepts for describing the situation being investigated which make it understandable to him or her and which are clearly “scientific” in that they are well anchored in existing knowledge (i.e., the basic nature of the concepts is already known in existing theory and research). An emic investigation, on the other hand, can be regarded as carried out *from the inside*. It is based on self-selected and self-organized statements made by people who have experience of the situation being investigated, regardless of what the researcher thinks. These people may or may not be consciously aware of the feelings, opinions, attitudes, values, hopes, ambitions, and the like which they express. What the respondents, for whom the situation being discussed is part of their normal life, say makes life understandable *to them*, regardless of whether the statements support the researcher’s understanding of the situation or not. In principle, anybody who is familiar with a given situation is in a position to deliver emic information about it. This ability to deliver “insider” information about a setting and events in the setting is referred to later as involving “narrative competence” (see p. 112).

Macro-analytic versus micro-analytic research:

In anthropology, ethnography, and sociology especially, the goal of research is often to reveal, describe, categorize or interpret characteristics of a broad field of action such as a culture, a society, or a group within a society such as for instance an occupational group. This approach (adding to understanding of a broad field) involves *macro-analytic* research. In psychology, by contrast, the focus is more often on individual people and their interactions with each other and with the circumstances in which they find themselves. This involves *micro-analytic* studies, which focus on small sub-groups or even on individual people. As Lewin (1997, p. 145) put it, whereas macro-analytic studies are concerned with working out “general laws of group life,” the micro-analytic approach focuses on the “diagnosis of a *specific situation*” [emphasis added]. The intention of this focus on the specific is sometimes to solve a

particular specific problem.

A “classic” example of a macro-analytic study is Hofstede’s (Hofstede, Hofstede & Minkov, 2010) analysis of “cultural dimensions.” On the basis of a study of more than 100,000 IBM employees in 53 countries, Hofstede identified six dimensions through which he argued *any society* could be classified (i.e., the dimensions are universal): “Individualism vs. Collectivism,” “Power Distance,” “Uncertainty Avoidance,” “Masculinity vs. Femininity,” “Long-Term Orientation,” and “Indulgence vs. Self-Restraint.” Hofstede also worked out a way of calculating index values for any country on any dimension, so that a profile or “cultural map” could be calculated for each country: For instance, one country might be highly individualistic with moderate power distance, low uncertainty avoidance, high masculinity, a short term orientation and high self-indulgence, whereas a different country might be collectivist, and have low power distance, high uncertainty avoidance, extreme masculinity, a short term orientation and high self-restraint.

This focus on a specific and concrete problem probably reaches its peak in *action research*, where a “spectator” (a researcher) and an “agent” or agents (people actually experiencing the problem being investigated) carry on a discourse aimed at jointly solving the specific practical problem, not at the discovery by the outsider – the researcher – of universal laws of human behaviour. Both quantitative and qualitative research can be macro-analytic or micro-analytic, although the general thrust of quantitative research is macro-analytic, whereas qualitative research is largely micro-analytic.

Classifying Research Designs

All research methods, whether qualitative or quantitative, can be examined systematically by looking at them within a general framework encompassing six broad dimensions. Such a framework is outlined below.

The dimensions of research designs: Regardless of orientation, the empirical part of all research projects can be examined in terms of six fundamental dimensions:

- *design* (e.g., experimental, quasi-experimental, *ex-post-facto* design, non-experimental design)
- *setting* (e.g., laboratory, real-life setting)
- *data collection procedures* (e.g., special apparatus, tests, questionnaires, observation schedules, interviews)
- *kind of data* (e.g., ratio/interval, ordinal, nominal, non-numerical data, narratives)
- *kind of analysis* (e.g., statistical analysis, meaning based analysis)
- *strategies for generalizing findings* (e.g., hypothesis testing, hypothesis generation).

The essence of an experiment: The meaning of terms such as “laboratory,” “real-life setting” or “questionnaire” should be obvious to readers, and expressions like “ratio/interval data,” “statistical analysis” belong to the everyday language of empirical social sciences. However, some terms are less obvious (e.g., “narrative”), and these will be discussed in this or later chapters. The first thing that is needed is a brief explanation of “experimental” and related terms. Although researchers in the social sciences frequently speak of “experiments,” they often use the term rather carelessly to refer to any situation where data are gathered in order to investigate a question or issue. What these people really mean is “empirical.” In this book, the word “experimental” is used more

strictly: An experiment involves a research design in which there is:

- a) at least two clearly defined variables, one of which (the independent variable) is thought to “cause” the other (dependent variable),
- b) a defined group (the *experimental group* – or more than one) that receives a *treatment* (the independent variable is manipulated in some way),
- c) one or more *control groups* (that do not receive the treatment),
- d) *systematic sampling* of the population under study, and
- e) *random assignment* of participants to experimental and control groups.

The purpose of these procedures is *control of confounding variables*, so that it is possible to make unequivocal statements about cause-and-effect relationships between independent and dependent variables. Confounding variables are unintended pre-existing differences among participants in a study that influence the effects of the treatment to an unknown degree and distort or disguise the effects of the actual intended treatment. Suppose that we were interested in differences in the quality of teaching in two schools, one of which was thought to be very good, the other less good. The children in the first school had an average IQ of 130, those in the second an average of 90. We might indeed discover significant differences in achievement between the schools but would not know whether these were caused by between-school differences in teaching quality or between-student differences in ability (or of course by other factors): IQ would be a confounding variable. The usual way of dealing with such a problem would be to eliminate IQ as a variable (i.e., to *control* it). This is often done by means of sampling, for instance by selecting samples from the respective schools that were matched on IQ. A second possibility is to control potential confounding variables’ effects by statistical means: If it is known how strongly a confounding variable correlates with the result of the treatment, the confounding variable’s influence can be

removed mathematically. For instance, in the example just given it would be possible to calculate how much IQ correlates with school achievement, and then to adjust the difference in achievement scores between the schools by removing the amount of difference attributable to IQ.

As designs depart further and further from these ideal conditions – for instance because random assignment is impossible (*quasi-experiment*) or because the treatment has already occurred prior to the conduct of the research (*ex-post-facto* design) – they depart further and further from the conditions necessary for a true experiment, eventually becoming clear-cut non-experimental designs. These may suffer from one or more of:

- a) lacking a specific treatment,
- b) using no control group(s),
- c) depending on opportunity samples consisting of whoever is available and willing to participate i.e., no random sampling,
- d) having no control over confounding variables.

Stereotypes of quantitative and qualitative research: For the purposes of the present discussion, the dimensions listed above will be treated as dichotomies, although it must be emphasized that this represents an oversimplification – for instance, it is obvious that research settings form a continuum ranging from “pure” laboratory settings where the circumstances are completely controlled by the researcher, through settings that have some properties of the laboratory, to “natural” settings where events occur as they do in real life, whether this helps the researcher or not.

An example of this continuum is seen in a study of mother-child interactions. At one extreme (the pure laboratory pole), observation was carried out at the university through a one-way mirror. The mothers and children were placed in a soundproof room, empty of all furniture except standardized tables and

chairs, and their behaviour was rated using rating scales. Every mother was required to speak exactly the same words to her child, regardless of the child's reaction. At the other extreme (the natural setting) was a study following each mother in her real life as she coped with her regular day-to-day schedule, regardless of the fact that the telephone might ring or the child graze its knee, or a tradesperson come to do repairs, or the games chosen by the mother be different from family to family, without any regard for the researcher's plans or needs. It is easy to imagine intermediate positions between these poles, thus defining a *continuum*, with "pure" natural conditions at one extreme and "totally lab-controlled" conditions at the other, but with intermediate positions such as "largely but not entirely controlled" or "almost natural" between the poles.

The continuum from "pure" lab conditions to "pure" natural conditions allows a highly differentiated view of the difference between strict lab studies and studies conducted in real-life settings. None the less, thinking in terms of dichotomies greatly simplifies discussion, and this simplified view will be adopted here for ease of understanding. Six dichotomous dimensions based on the dimensions listed in Table 2.1 yield 64 possible combinations of conditions, i.e., 64 "types" of research design, and in theory all of these are possible. For instance, it would be possible to select an experimental and a control group using random assignment (conditions for a true experiment) and then conduct interviews (thus employing a data collection procedure more characteristic of non-experimental designs). However, there is a correlation among the dimensions listed above (design, setting, data collection procedures, etc.), perhaps the most obvious example being that setting tends to go together with design: For instance, in real-life field settings it is almost impossible to apply a true experimental design. Similarly, it is difficult to collect data with laboratory instruments in field settings, while it is impossible to conduct numerical analyses

with purely narrative data. In practical terms, however, it is helpful to focus on only two of the theoretically possible combinations of design, setting, data collection procedure, kind of data, kind of analysis, and generalization strategies. These are:

- 1. An experimental design in a laboratory, yielding numerical data collected via objective procedures and leading to statistical testing of hypotheses. This represents an idealized quantitative approach.
- 2. A non-experimental design in a real-life setting, involving collection of verbal reports describing respondents’ construction of the situation in question, and content-based analysis of these reports. This is an idealized qualitative approach.

Such oversimplified stereotypes are depicted in Table 2.1. It should be noted that the table actually confounds two dimensions of research: what is called “approach” in the table (qualitative vs. quantitative) and orientation (i.e., applied vs. basic). In effect, the table compares *basic* quantitative research with *applied* qualitative research. As a result, it may create the impression that there is no such thing as basic qualitative research or applied quantitative research. In fact, both exist. Nonetheless, there is a tendency for qualitative research to be more frequently of the applied kind than of the basic.¹⁰

Table 2.1: Characteristics of idealized quantitative and qualitative approaches

Dimension	Quantitative Approach	Qualitative Approach
Design	experimental	non-experimental

¹⁰ In fact, most research projects display some departures from the “perfect” types depicted in the table, but focusing on the idealized patterns to be found there helps make the contrast between the two approaches clearer.

Setting	laboratory	field
Data Collection	instruments (e.g., electronic or mechanical devices, tests, questionnaires)	collection of “narratives,” either already existing (e.g., diaries), or new (e.g., interviews)
Data Type	numerical (ratio, interval, ordinal)	descriptive (e.g., interview protocols, written records, videos)
Analysis	descriptive and analytic statistics	analysis aimed at revealing meaning
Generalization	focus on testing hypotheses	focus on generating hypotheses

Doing Good Research

What makes qualitative research scientifically rigorous? The stereotypical quantitative approach depicted in Table 2.1 corresponds well with what many people regard as the essence of the scientific method. The stereotype of qualitative research, by contrast, may seem to involve nothing more than interacting with people and drawing conclusions about them, something that everybody does all the time. How can this be scientific? As Kromrey (2009) put it, qualitative research may even find itself competing with common sense, which may not claim to be scientific but is nonetheless very useful in helping people get along in life. When research findings agree with common sense they are derided as banal and simply stating the obvious, but when they disagree with it they may be treated with suspicion – a typical “damned if you do, damned if you don’t” situation.

Shaughnessy, Zechmeister and Zechmeister’s (2014) discussion of this issue is very helpful. They agreed that both

“scientists” and “non-scientists” learn about behaviour by observing people going about their lives. The crucial difference lies not in what people in the two groups do or how much of it they do (i.e., it is not quantitative), but in how they do it (i.e., the difference is itself qualitative). The differences lie in three main areas:

1. *Organization of observations*: Non-scientific observation is casual, informal and may be unconscious, whereas scientific observation is conscious, systematic and purposefully directed towards goals specified beforehand.
2. *Awareness of distorting factors*: Non-scientists have little or no awareness of distorting factors in their interactions with other people, whereas scientists should be acutely aware of them.¹¹
3. *Record keeping*: In the case of non-scientists, record keeping relies on casual recollection, with all its distortions, which have been well known at least since the work of Bartlett (1932). Scientific record keeping, by contrast, is formal and systematic, and involves written or electronic records.

Three further dimensions should be added to these:

4. *Analysis*: Non-scientists draw conclusions in an unsystematic way with little awareness of the effects of their own expectations and biases and little distinction between these and systematic conclusions. Scientists are aware of distorting factors and distinguish clearly between well-grounded conclusions and speculations, tentative hypotheses, even wishful thinking.
5. *Generalizing*: Non-scientists may generalize with reckless abandon, simply assuming that the conclusions referred to in 4 above apply automatically in a wide variety of other settings, whereas scientists are keenly aware of the limits to generalization, preferring to state many of them as

¹¹ These distorting factors are discussed in greater detail on p. 72 and p. 98 (e.g., selective perception, dissonance reduction, etc.).

suggestions for clarification through further study.

6. *Communicating*: Non-scientists may be vague about details and chiefly interested in persuading other people to accept their views (as in much journalism), whereas scientists carefully describe the procedures that led to their conclusions, and draw attention to aspects of their work where caution may be called for. This is at the core of differences between truly scientific writing and areas such as popular science, or journalism. Writers in the latter domain are chiefly interested in persuading a mass audience to accept their views, whereas scientific writers are chiefly concerned with providing sufficient information for specialist readers to draw their own informed conclusions.

Reliability and validity: The question that now arises is how to judge the strengths and weaknesses of the two research approaches. Put in a down-to-earth way, this involves the issue of convincing readers that the “findings of an inquiry are worth paying attention to” (Lincoln & Guba, 1985, p. 290). To do this, it is first necessary to specify the key characteristics of good research and the sections that follow are concerned with developing a framework for doing this. The matter can be addressed by looking at two concepts: reliability and validity. Validity has two forms: internal validity and external validity (see below). Although the basic ideas are similar, reliability and validity have a different meaning in the present discussion from their meaning in test theory, and it should not be assumed that the passages that follow simply repeat test-related discussions.

The key ideas for the present purposes are as follows:

1. The *reliability of research* refers to the likelihood that its findings would be replicated by a different researcher working with different participants. Highly reliable findings have a strong likelihood of being replicated, whereas unreliable findings do not. From the point of view of

qualitative research, the practical importance of reliability is that unreliable findings cannot be *generalized*, i.e., the information they provide cannot be used to understand settings over and above the one in which the information was obtained. Unreliable findings may be very informative or interesting, but they are the result of a particular interaction between a specific researcher and the specific participants in a particular study.

2. The *internal validity of research* refers to the likelihood that apparent relationships among variables are genuine and are not, for instance, the result of some other undetected factor(s) or even the result of some aspect of the data-collection procedure. This is easiest to understand by thinking of cause-and-effect relationships: Suppose, for instance that a group of boys did better on a maths test than a group of girls. It might be concluded that gender causes superiority in maths. However, if attitudes to maths and previous experience with maths were much more favourable among the particular group of boys in the study in question (for whatever reason), these advantages might well be the real reason for the differences in test scores. Because of lack of control of such *confounding variables* the research would have low internal validity.

3. The *external validity of research* involves the extent to which its findings continue to apply in new settings that are unlike the circumstances under which the study took place, especially real-life settings. For instance, if a researcher found in a study that smiling at people made them behave in a friendly way in a large unruly queue at an airport, external validity would be high if smiling produced the same beneficial effects in other unrelated settings such as after a car crash or during a riot at a football match.¹²

¹² A helpful concept in this context is that of “ecological” validity. As a strict methodological concept, *external* validity refers to the question of whether the

It is important to note a number of important issues in discussing reliability and validity as terms to describe research designs:

1. *Reliability and validity are continuous variables.* They are not discrete properties of a research project that are either entirely present or entirely absent – research is not either reliable/valid or not. Reliability and validity range, at least in theory, from non-existent (expressed as a coefficient of 0.00) to perfect (1.00), with all intermediate values being possible. Thus, one project may have low reliability/validity, another high reliability/validity, others fairly high or fairly low levels, still others intermediate levels, and so on.
2. *Reliability and validity do not necessarily co-vary.* It is possible for a project to have, let us say, high reliability but low validity, or the reverse. In practice, for reasons that will be mentioned below, the two properties tend to correlate, but this is not theoretically necessary.
3. *High reliability and validity of research are norms or ideals* that researchers would like their studies to approach, the closer the better. They are not states that are either present or absent – a study is not either “reliable” and therefore good, or else “unreliable” and therefore bad (or “valid” and therefore good versus “invalid” and bad), but as reliable (or valid) as possible.
4. *The level of reliability/validity of a study is only a theoretical property.* Except perhaps in the most unusual cases, reliability/validity coefficients of research are never

findings of a study can be generalized to research settings other than the one in which the findings were obtained. Do findings obtained with a sample drawn from one population also apply to other populations?. Ecological validity, on the other hand, involves the question of whether the findings of a particular study *apply to the real world*, or only to settings contrived by a researcher. Thus, ecological validity is intuitively a stricter version of external validity. However, for the discussion of qualitative research methods, the crucial point is that it emphasizes real-world practical usefulness.

actually calculated. Thus, when it is said that a particular study has, let us say, “high internal validity,” this usually mean that it *looks as though* it has this property because it possesses characteristics that are known to be related to high internal validity (such as random sampling, use of a control group, etc.).

5. *Statements about the reliability/validity of research are probabilistic* in nature – a given project may have a high *likelihood* of being reliable/valid because it keeps to the rules of good practice, but it is not actually known whether the study really achieved these.

6. *Reliability and validity are potentials rather than established facts.* General statements that particular approaches such as a quasi-experimental design or use of narratives as the method of data collection have “high” reliability or validity mean only that studies that follow the rules for that kind of approach have the *potential* to be highly reliable/valid. To take an example, qualitative studies are often said to have “high” external validity (i.e., their findings are very plausible as statements about the real world). This means only that specific, individual studies employing a qualitative approach offer the opportunity of achieving high external validity, or that they are more likely to yield findings with high external validity, not that this is guaranteed. Naturally, it is possible to conduct research badly, regardless of the particular approach, with resulting low reliability/validity. Thus, it is appropriate to say that an approach is “favourable” for achieving high reliability/validity. Similarly, when it is said that a particular approach has “low” reliability/validity this does not mean that it is impossible for research of this kind to be reliable/valid, but that it is more difficult. For instance, as a group, experimental designs have “low” external reliability. This means that it is frequently difficult

to transfer the findings of such studies in a plausible way to the world outside the laboratory. However, a particular experimental study may overcome this problem and be very plausible.

7. *It is possible to enhance or increase the probability of a study's reliability and validity.* This is done by building into it characteristics that are known to be favourable to these properties. In the case of quantitative research this means via *sampling and control*. How to enhance the reliability and validity of qualitative research is discussed specifically in Chapter 7. Of course, it is also possible to “threaten” or “endanger” reliability/validity by ignoring the rules of good practice. It is also possible for strengths in one area of a study (e.g., design, data collection procedure, method of data analysis) to compensate at least partially for weaknesses in another. This means that it is most appropriate to think of certain conditions as “favourable” for reliability and validity, others as placing them “at risk.”

8. *Specific, individual studies may overcome some of the weaknesses of a particular approach.* Thus, one kind of research could have, let us say, a low probability of high external validity, but in an actual study the validity might be high. Although qualitative studies as a group run the risk of low internal validity, it is possible to imagine specific qualitative studies with high internal validity. In reporting a study (see Chapter 8 for appropriate discussions) emphasis should be given to the way in which an attempt was made to minimize the effects of factors likely to place reliability and validity at risk.

9. *It is usually necessary to compromise on reliability and validity.* Particular research approaches (e.g., qualitative vs. quantitative) possess inherent strengths that enhance reliability/validity and, at the same time, weaknesses that endanger them. Thus, a particular project is usually

a compromise, with some aspects favouring reliability/validity, some placing it at risk.

Rigorous research: Every research approach should seek to:

- a) *optimize reliability* (i.e., go as far as possible towards ensuring that the findings are not simply the result of the preconceptions and expectations of a single researcher or participant or of the special and unique interaction between them, but would also emerge if the cases in question were analysed by a different researcher following the procedures described in the research report or if the same topic were investigated with different participants;
- b) *optimize internal validity* (i.e., do whatever can be done to ensure that findings draw attention to genuine relationships among variables);
- c) *optimize external validity* (i.e., show as convincingly as possible how findings can be applied to other settings and to real life – ideally, findings should offer insights going beyond the particular cases studied in the research project, and encompass people in general).

The term “optimize” is used here because reliability and validity can be regarded as ideals that can never be reached entirely, but towards which researchers can constantly strive.

Although it has been pointed out that reliability and validity can vary independently of each other, there is some degree of correlation among the various dimensions of a research project: Conditions that favour high reliability and high internal validity (see the discussions above of what these are) often run the risk of endangering external validity, although this is not necessarily the case. The task is then not that of doing research that is perfectly valid (or reliable), but of making the combination of validity and reliability as good as possible (i.e., optimized) A research design that strives for optimized reliability, internal validity, and external validity and is aware of and takes account of weaknesses in these,

can be said to be “rigorous.”

Research dimensions, reliability and validity:

There is a systematic relationship between the dimensions of a research study listed in Table 2.1 and reliability and validity. These relationships are summarised in Table 2.2.

Table 2.2: Dimensions of research and reliability and validity

Dimension			
Design	Reliability	Internal Validity	External Validity
<i>True experiment</i>	Highly favourable	Highly favourable	At risk
<i>Quasi-experiment</i>	Favourable	Favourable	Lower risk than experiment
<i>Ex-post-facto</i>	Favourable	Favourable	Lower risk than experiment
<i>Non-experimental</i>	At risk	At risk	Highly favourable
Setting			
<i>Laboratory</i>	Highly favourable	Highly favourable	At risk
<i>Real life</i>	At risk	At risk	Highly favourable
Data Collection			

<i>Instruments, tests, questionnaires, etc.</i>	Highly favourable	Highly favourable	At risk
<i>Narratives</i>	At risk	At risk	Highly favourable
Data Type			
<i>Numerical</i>	Highly favourable	Highly favourable	At risk
<i>Descriptive</i>	At risk	At risk	Highly favourable
Analysis			
<i>Statistical hypothesis testing</i>	Highly favourable	Highly favourable	At risk
<i>Content analysis</i>	At risk	At risk	Highly favourable

It is important to stress again that the “risks” listed in the table are not fixed and immutable consequences of a particular approach, but rather tendencies or possibilities. Thus, when the table reports that the internal validity of non-experimental designs is “at risk,” this should not be regarded as indicating that low internal validity is an unavoidable consequence of such designs. The statement draws attention to a *danger* or *threat* that is harder to avoid in non-experimental than in experimental designs. In a similar way, although experimental designs offer good prospects of high internal validity this is not guaranteed. As can be seen by combining information from Tables 2.1 and 2.2, neither “pure” quantitative nor “pure” qualitative approaches offer highly favourable conditions for all aspects of reliability and validity, although it would be possible in theory to use “mixed” approaches combining some aspects of both to enhance the likelihood of high reliability, high internal validity and high external validity. Indeed,

there is a literature on such approaches (e.g., Tashakkori & Teddlie, 2010).

Revised concept of validity: Kvale (1995) criticized what he saw as the excessive weight given to conventional notions of reliability and validity in discussions of the rigorousness of research, as did Niederberger and Dreieck (2020) a quarter of a century later. According to Kvale, the quantitative approach to reliability and validity rests on the assumption that there is (a) an objective reality and (b) knowledge about this reality in people's minds. The job of research is (c) to construct a "map" of this reality. The research is then judged to be valid if there is a high degree of correspondence between the map yielded by the research and the objective "facts" it investigated. However, Kvale did not reject validity out of hand. Rather, he drew attention to the fact that the concept takes on a special form in qualitative research. To anticipate later chapters, he argued for a revised definition of validity derived from classical philosophy of truth rather than the "mapping of reality" approach. According to Kvale, there are two criteria of validity:

- a) the "coherence" criterion that refers to the *consistency and internal logic* of a research study;
- b) the "pragmatic" criterion that refers to its *practical usefulness*.

In this book a similar position is adopted: The idea of validity is not simply abandoned, but is replaced by a revised concept that emphasizes consistency, logic and usefulness.

The balance of this book can be seen as a discussion of how to conduct reliable and valid (i.e., rigorous) qualitative research. The following chapters are concerned with questions such as:

1. What are the special characteristics of qualitative research that make it different from other approaches (i.e., its internal logic)?
2. How is it possible to conduct qualitative research that is rigorous (i.e., that has high reliability and validity)?
3. How are qualitative data analysed in reliable and valid

ways?

4. How can the results of qualitative studies be interpreted and reported (i.e., made useful)?

This chapter and the one that follows (Chapter 3) are concerned with background material, especially theory and history. Chapters 4-8 focus on techniques for doing qualitative research. It would be possible to read the book without taking any notice of Chapters 1-3, focusing instead only on practical issues. However, although the intention is to be practical and introductory, a certain minimal understanding of the theory of practice is necessary for informed and critical application, and this is preferred to blind following of rules. Tashakkori and Teddlie (2010) divided their discussions of the practice of qualitative research into four “phases,” and a structure similar to theirs will be adopted in this book. Chapter 4 is mainly concerned with *conceptualization* of a project: What is it really about and who are the people who know about the issues in question? Chapter 5 focuses on *methodology*: How can data be collected? Chapter 6 extends the discussion of data collection by introducing approaches which combine qualitative and quantitative methods (i.e., mixed methods). Chapters 7 and 8 discuss *data analysis*: How can the contents of the information obtained from the participants (or “constructed” by researcher and respondent in the course of their interaction, as some writers would put it) be teased out? Finally, Chapters 9 and 10 outline procedures for *generalizing and reporting*: How can researchers use the information they gather to enlarge understanding of social science phenomena, and work out useful practical advice?

Chapter 3: The special nature of qualitative research

Qualitative research emphasizes (a) *discovery*, (b) *broadly defined goals*, (c) *flexible research questions*, (d) *emergent contents*, (e) *generation* (as against testing) *of hypotheses* and (f) *credibility*. It can be understood in terms of its understanding of reality, and the kinds of question it examines (its epistemology) and the methods used to find answers to the questions (its heuristics). Qualitative research is largely inductive and “flexible” in nature, whereas quantitative research is mainly deductive and positivistic. Qualitative methodology raises particular problems for researchers, largely connected with rigour, and relies heavily on special criteria of validity such as soundness of reasoning or plausibility of conclusions. Some researchers turn to qualitative methods because they expect them to be easier than quantitative methods. However, qualitative research makes special personal demands on

researchers that are different from those of quantitative approaches; for example, the ability to elicit personal accounts from participants or the ability to think inductively. In addition, it raises special ethical issues, because it not infrequently asks participants to reveal how they construct personally significant or even painful events.

Chapter Glossary

<i>deductive reasoning:</i>	inferring unknown specific events from a known general law
<i>ethics:</i>	rules or principles which provide guidance on how to behave in a moral or proper manner
<i>ex-post-facto design:</i>	a research design in which the treatment has already occurred prior to commencement of the research
<i>flexible research question:</i>	a research question that is only loosely defined at the beginning of the investigation and is adapted as data are analysed.
<i>inductive reasoning:</i>	inferring a previously unknown general law from known specific events
<i>natural setting:</i>	the real-life situation in which behaviour being investigated occurs
<i>object of study:</i>	the issue, problem, or phenomenon with which a research project is concerned
<i>trustworthiness:</i>	the extent to which research findings

can be accepted with confidence that they accurately reflect reality

The Essence of the Qualitative Approach

As Hollis (1994) put it, the qualitative approach seeks to *understand* phenomena, whereas the quantitative approach seeks to *explain* them. The goal of qualitative research is emic (to describe and analyse the world *as it is experienced, interpreted and understood* by people *in the course of their everyday lives*). It is frequently micro-analytic and it is often (although not always) focused on a specific problem in a specific situation. By contrast, quantitative research is frequently (although not always) etic and macro-analytic and focused on discovering new general laws of behaviour or enlarging knowledge of existing laws.

The qualitative approach can be understood in terms of three dimensions: the way reality is understood, the kinds of question it regards as important, and the procedures it employs for obtaining answers to the questions just mentioned).

1. The qualitative approach involves the view that (a) each person actively “constructs” an individual “reality” out of his or her own particular experiences, (b) this *reality differs from person to person*, and (c) it is shaped by interactions with other people, i.e., reality is ‘socially constructed.’ This means that people are not “subjects” who passively receive whatever an objective external reality offers them, but that they are “actors” who “construct” the world they live in.
2. Qualitative procedures emphasize questions such as: (a) How do people make sense of the external world? (b) How do they understand how they themselves and others fit into this world? (c) How do they decide how to act in terms of their own reality? (d) How do they communicate their

understandings to other people?

3. The social sciences are empirical disciplines in which advances in knowledge are achieved *via observation of people's behaviour*. In this they differ from non-empirical disciplines that advance understanding either by following the rules of an internally coherent, closed system of axioms that are taken as indisputable truths (e.g., Marxism, theology) or by the application of logic (e.g., mathematics, philosophy). This use of observation of behaviour as the source of new knowledge is shared by both qualitative and quantitative research methods. However, the qualitative approach differs from the quantitative in the following way: It involves

- *observing* one or more people in a real life setting and recording what happens or *actively eliciting descriptions* of real life from people, for instance by means of interviews. These descriptions are almost always verbal, very frequently spoken, and are often referred to as “narratives”;
- *recording* what the people say or write, or how they behave. This often involves a written text, even when the original record consisted of audio or video tapes;
- *interpreting* what the people said. From the point of view of the present book, the purpose of this interpretation is to use the concepts of a particular discipline such as sociology or psychology to understand what was observed and recorded in the steps already outlined. The results of this step constitute the findings of the research;
- *generalizing* these interpretations by relating them to other people or settings, or to existing theory (or both). This step yields the *conclusions* of the study. Such conclusions may consist of a general description of a setting or a group of people such as a description of interpersonal relations in, let us say, a hospital (more typical of sociology or anthropology than of psychology or education), or of the

extension and clarification of existing theory regarding some phenomenon such as the development of intelligence (more typical of psychology and education).

Contrast with the quantitative approach: The special nature of the qualitative approach's ontology, epistemology and methodology can be made plain by contrasting them with the quantitative equivalents. (The latter are usually already well known to anybody who has completed an introduction to research methods.) At the heart of quantitative methods is *positivism*, a view of the human psyche that rests upon the following assumptions:

- When we conduct behavioural research, we are dealing with clearly defined properties of human beings.
- These properties are present in relatively stable, specifiable amounts in all individuals.
- The properties take the same form in all these people.
- Different people possess different amounts of the properties under investigation.
- Researchers possess instruments capable of measuring these amounts.
- The amount of the property in question can be expressed in numbers.
- It is not even necessary for participants to be aware of the property the researcher is studying – it is, so to speak, not “their” concept but “ours.” For instance, schizophrenic patients do not spontaneously use terms such as “over-inclusive thinking.” They say things like, “Ideas jumble themselves up in my head.” It is the researchers who interpret such everyday statements in terms of psychological concepts.
- The purpose of research is to demonstrate the applicability of existing “laws” in new situations and thus to extend the range of situations for which the laws are valid and/or to refine and perfect the laws.

Transferring existing “laws” to new situations and demonstrating

that they apply there too involves *deductive* thinking.

Qualitative methods: By contrast, the ideas at the core of the qualitative approach are, in a sense, the mirror image of the positivist principles just stated. As Wisdom and Creswell (2013, p. 3) put it, qualitative methods “give a voice to study participants and ensure that study findings are grounded in participants’ experiences.” In particular, qualitative approaches emphasize that:

- the task of research is *to discover how people see the world*, not how much they possess of certain universal characteristics;
- researchers and respondents must work together *to reconstruct the respondent’s conceptualization* of the world – research is a process of cooperation between researcher and respondent;
- the process of interpreting data frequently involves *inferring* general hypotheses from particular information (the hypotheses *emerge* at the end), whereas quantitative methods involve *confirming* particular hypotheses derived from existing general statements (the hypotheses already exist at the beginning). Thus, in qualitative approaches theorizing is often based on *inductive* reasoning.

Put briefly, the task of qualitative research is to find out how people “... *impose order* on the flow of experience to make sense of events and actions in their lives” (Riessman, 1993, p. 2) [emphasis added]. As a result, qualitative research designs in social science usually involve observing and recording people’s behaviour either in natural settings, for instance via participant observation, or as the people themselves recall such settings (e.g., by means of various kinds of narratives such as interviews). This can be contrasted with quantitative approaches which regard subjective properties of individual research subjects as disruptive elements with the potential to confuse the issue. Such factors are only considered in research to the extent that they can be measured and quantified, and the purpose of doing this is to eliminate them from the study

in question (i.e., to “control” them). Statistical analyses of data focus on aggregated scores and group statistics such as means or variances, and thus eliminate the peculiarities of individual people from the data. Qualitative researchers, by contrast, seek to adopt the individual perspectives of the participants and *to think themselves into* the system of experiences and constructions these people are asked to reveal in the study.

The adoption of the principles outlined above leads to a qualitative research orientation that emphasizes (see also Elliott, 1999, p. 252):

- *discovery* (as against confirmation of the expected)
- *broadly defined goals* (as against a set hypothesis)
- *flexible research questions* that may change in the course of the study
- *emergent* content themes (i.e., themes that develop in the course of the analysis)
- *generation of hypotheses* rather than confirmation/disconfirmation
- *credibility* as the criterion of scientific soundness (as against sampling and measurement technology)

The Special Properties of Qualitative Research

Non-experimental designs: As a rule, qualitative studies are not experimental but descriptive, and focus on the everyday life of the people involved. Qualitative designs focus on optimizing the *external* validity of findings, whereas quantitative approaches emphasize *internal* validity. This internal validity is assessed in terms of technical criteria such as control of confounding variables, construction of random samples, or satisfaction of the formal

criteria for application of statistical procedures. External validity, by contrast, derives from properties such as *authenticity*, *plausibility*, *insightfulness* or *usefulness*. The central question is not, “Is this design technically perfect?” (i.e., internal validity is paramount) but “Does this design help us to understand the phenomenon that is being investigated – as it is conceived by the participants – better?” (i.e., external validity is paramount).

Because the “treatment” that led to the construction of reality of a particular participant had already occurred prior to the research study currently being conducted (in the course of the participant’s life), qualitative research designs are usually *ex-post-facto* designs. It is conceivable (at least in theory) to use quasi-experimental designs in qualitative research: For instance, a group of participants who had had particular life experiences in the past could be compared with the members of a group who had had different experiences in their lives, although this is uncommon. An example would involve comparing the way people who have been in prison understand the purpose of penal confinement with the way people who have never been imprisoned understand such confinement. In qualitative research genuine experimental designs are also very uncommon, because such designs require random sampling, assignment to “treated” and “untreated” groups (experimental group vs. control group), control of confounding variables, and deliberate manipulation (treatment) of the experimental group (but not the control group). Because such conditions (a) require that the researcher possesses advance knowledge of the relevant causal factors so that instruments/techniques can be selected for measuring them, (b) treats additional factors as an annoying source of error and controls them through various techniques such as sampling or statistical control, and (c) deals with (a) and (b) by collecting data under closely controlled conditions or even in a laboratory, experimental designs usually contradict the principle of qualitative research that data should be collected under *everyday conditions that resemble day-to-day life*. For these reasons, using

“classical” quantitative criteria to classify and analyse qualitative designs hardly makes any sense.

Flexibility¹³: In traditional quantitative research, the research question is always derived from *existing theory*. Researchers study what is already known from earlier findings and formulate a formal hypothesis on the basis of this material. This hypothesis is then tested, using an appropriate design and objective measurement procedures (frequently: tests, scales or instruments of some kind) and statistical tests. The purpose is to test whether the hypothesis derived from earlier findings can be transferred to a new setting and confirmed as applying there too. Without existing knowledge and appropriate theory about the issue being investigated there would be no hypothesis and, as a consequence, no research study.

By contrast, qualitative studies are *flexible*. While it is true that they have to have a clearly-defined object of study, i.e., they have to investigate something specific, they can commence without any theory at all, for instance simply because of the researcher’s interest in some topic or issue. This interest can result from simple curiosity, from a hunch or an intuition, or from special involvement in an area. For example, a single parent could be active in counselling others in a similar situation and could be interested in obtaining information about how other single parents handle their lives, or a bank clerk who is active in the union movement could investigate how other employees conceive their relationship to their employer. This does not mean that qualitative research is not interested in theory: In order to understand any object of study researchers must collect, organize, analyse and reflect upon data, and must convert the data into scientific insight. Theory is essential for achieving this. However, in qualitative research the theorizing often *follows* the data analysis. In other words, theory can play its main role *at the*

¹³ In earlier writings I referred to this characteristic of qualitative research as “openness.” However, I now prefer to speak of “flexibility,” because openness in research has a new meaning related to the concept of “open” research (see Open Science Collaboration, 2015).

end rather than at the beginning.

An emergent topic: In the mid-2000s a large European airfreight firm was experiencing difficulties with its cargo handling division: Experienced and highly valued freight loaders were resigning in large numbers, and new recruits staying only a few weeks. In casting around for a possible explanation management decided that a new computerized loading system which had recently been taken into service must be the culprit and, in order to analyse the precise nature of the problem, a questionnaire was designed by in-house researchers with a psychological background. This involved items like: “I find it difficult to understand the loading instructions given by the computer,” “I object to being given orders by a machine,” or “Inputting information into the computer is too complicated and time-consuming.” The loaders responded to each item by selecting an alternative ranging from “agree completely” to “disagree completely” on a seven-point scale. To the surprise of the researchers, responses to the questionnaire quickly showed that, in fact, the workers expressed approval of the new computer-guided loading system. It was easy to use, gave sensible and easily understandable instructions on how to load freight, and did not insult the loaders or make them doubt the value of their pre-computer efforts. But despite this, the loaders continued to resign, and the questionnaire was adjudged to have been a failure, since it contradicted pre-existing beliefs about where the problem lay and gave no hint of its source. Fortunately, the questionnaire included a single open item where the workers could express their own views rather than responding to suggestions made by the designer of the questionnaire. The open item’s wording was as follows: “What is it like working as a loader?” Almost by chance the researchers noticed that many loaders had made use of the item to express their

dissatisfaction with the firm's senior management. A closer analysis of their comments showed that the problem was not the new loading system at all, but a strong feeling that management had no understanding of the difficulties of the work of cargo handlers and did not care about them. The result was that the loaders believed that their firm was being run by incompetents, and felt a sense of alienation from their employer.

The case study in the box above illustrates the way in which the exact topic of a qualitative study can *emerge* in the course of the research, rather than being defined in advance. In other words, a clearly formulated research question can develop during the investigation (especially in the course of the analysis of the data), rather than being worked out in detail prior to the research commencing, as is the case with quantitative studies. Or, in the words of Charmaz (2014, p.1), qualitative research involves “iterative strategies of going back and forth between data and analysis,” so that there is a constant interaction between data and evolving theory. A topic that arises through this interaction is referred to as “emergent” (see p. 195ff). In the case study, the employees’ alienation was a finding which came as a complete surprise for management. They thought the employees respected and trusted them and needed only some training with the new loading system in order to make them happy. The topic “distrust of management” was completely unexpected and had not been thought of at all prior to the commencement of the research – instead of being specified in advance, it “emerged” in the course of the investigation. As a result, no hypothesis about it had been formulated and no items about it had been included in the questionnaire. Only the lucky inclusion of the open, qualitative item revealed the true problem. The questionnaire, with its pre-formulated items constructed on the basis of the researchers’ expectations, was a typical etic quantitative

instrument, although by no means an exclusive example of how such research is done. The single open item involved a typical general and unstructured question of the kind often used, although not exclusively, in qualitative studies. The purpose of such questions is to provoke respondents to express their views, for instance in the course of an unstructured interview, a method of data collection which is very frequently used in qualitative research. The two kinds of item stand in stark contrast to one another, and this contrast will be sharpened in the following sections.

The basic principles of qualitative research: The ideas just outlined can be summarised in five basic principles. These are as follows:

1. *Research question:* In qualitative research it is important to investigate questions involving the intentions and purposes of people, the “Why?” “How?” and “What kind?” are important, when looking at social science phenomena, not just “How much?” The latter question (“How much?”) is the domain of quantitative methods.
2. *“Philosophy” of research:* In order to increase knowledge, research focuses on people’s direct experience with their immediate, everyday world, and proceeds through stages of (a) recording participants’ narratives, (b) description of what they say in terms of discipline-based concepts and (c) explanation of what this means for theory in the particular discipline. Since it proceeds from specific cases to general explanations, the process is *inductive*. The explanation is not infrequently stated in the form of a hypothesis, i.e., the research **ends** with a hypothesis, whereas quantitative research *begins* with one.
3. *Method of gathering data:* People’s experience of the world and their relationship to it can best be investigated by allowing them to describe these phenomena as they themselves experience them. In other words, some kind of narrative of the people being investigated is needed. The

dominant form of narrative is the interview, although the extent of its dominance varies from discipline to discipline.

4. *Roles of researcher and subject*: The people who are often referred to in quantitative studies as “subjects” then become *partners* with the researcher in the reconstruction of their own experiences and their ways of understanding them. These people are referred to in this book as “participants” or “respondents.”

5. *Data analysis*: The fundamental task of the analysis is to *interpret* the narratives, i.e., to work out their *meaning*. As a rule, the researcher makes a protocol of respondents’ narratives and uses these (a) to describe the situation of the participants and (b) to work out the implications of the narratives for relevant theory.

Examples of Qualitative Research Questions

Qualitative questions: A simple example of a quantitative research question would be “Is there a difference between the level of intelligence of men and women?” Earlier research suggesting that there might be a difference would be cited at length and a hypothesis based on this research stated. For the purposes of research design this is almost always expressed in the form of a null hypothesis (i.e., “There is NO difference”), despite the fact that it is hard to imagine anybody undertaking research with a genuine expectation that no differences would emerge. To answer this question it would be necessary to measure how much intelligence samples of men and women possessed, compare the two amounts with the help of appropriate analytic statistics in order to decide whether or not the null hypothesis could be rejected, and then draw

general conclusions. This would involve (a) knowing in advance what “intelligence” is; (b) possessing instruments for measuring it; (c) being able to express its amount in numerical form; (d) already having a hypotheses about its distribution across gender groups; (e) testing this hypothesis via statistical procedures (i.e., determining objectively whether any numerical differences revealed by the data were statistically significant); (f) generalizing the findings with the present samples to other groups of men and women apart from the ones involved in the study.

By contrast, a simple qualitative approach might ask, “Is there a difference in the nature of the intelligence of men and women?” The two groups might be equal in terms of amount of intelligence (quantitative approach), but might nonetheless display different *patterns* of intelligence, women for instance showing verbal, social, intuitively oriented ability, men revealing a mechanical, logical, mathematical pattern.¹⁴ The researcher might conclude that there is no quantitative difference but a substantial qualitative one – no difference in the amount of intelligence but differences in the kind. A more full-blooded qualitative approach would go further in rejecting the “How much?” approach. The researcher might ask, “How do men and women themselves experience their own interactions with situations requiring what is commonly called ‘intelligence’?” “What explanations do they use to explain what happens when they find themselves in such situations?” These questions define a more *phenomenological* approach.

Further examples of qualitative research questions are given below to help give a certain “feel” for this approach. The examples given here have no claim to representing the full range of possible topics that could be studied qualitatively. They are not arranged in any special order. Of interest here are not the actual contents of the questions but their qualitative style or form. It is important to

¹⁴ These gender stereotypes should not be seen as scientific summaries of relevant research. They have been introduced here purely as an artificial example in order to make clear the difference between quantitative and qualitative approaches.

remember that in looking at these questions researchers in different social sciences would emphasize concepts relevant to their own disciplines. A famous statement about the nature of sociology is that its aim *is to investigate sociological things in a sociological way*. Thus, whereas the question “Why do affluent, successful people emigrate from rich, highly developed countries?” (example number 1 below) might be examined by a psychologist in terms of motivation, personality, cognitive processes and similar factors, other social scientists (e.g. sociologists, anthropologists) might focus on for instance demographic, sociological, economic, or anthropological variables. The importance of this focus on discipline-based ways of looking at issues will become more apparent in Chapters 7-9, where interpreting and reporting research studies will be emphasized. It must be admitted, however, that the increasing use of interdisciplinary approaches means that the boundaries of what is “psychological,” “educational,” “sociological” or “anthropological” may be becoming blurred.

Some further examples of qualitative research questions are as follows:

1. Why do affluent, successful people emigrate from rich, highly developed countries?
2. How do drivers experience their interactions with other participants in road traffic?
3. How do parents of emotionally disturbed children conceptualize their situation?
4. What does it feel like to be a gifted child?
5. What factors in the workplace encourage or inhibit creativity?
6. What aspects of their work are especially stressful for social workers?

Qualitative answers: Questions such as these can be answered to some extent by formulating null hypotheses and testing them with appropriate analytic statistics. To continue with the example of why people emigrate, it might be expected intuitively

that they would do this to avoid unpleasant life conditions in the homeland (e.g., unemployment, poverty, harsh living conditions, lack of political freedom, repression, and the like). Supported by review of the relevant literature this might lead to the working hypothesis that people who earned less or were unemployed, experienced the lowest levels of social and political freedom, or suffered the greatest privations would be more likely to emigrate. This expectation could be formulated as a null hypothesis (e.g.: “There is no relationship between earnings level and interest in emigration”). This could be tested, for example by calculating the average income of a random sample of emigrants from a particular country and comparing this with the average income of a group of non-emigrants from the same country (thus *controlling* for variability deriving from country of origin by taking all participants from the same country). If the mean incomes of the two groups were statistically significantly different (demonstrated, for example by a t-test), the null-hypothesis would be rejected, and it would be argued that income really is a factor in emigration.

Despite the fact that numerous null hypotheses would have been rejected, none of this would have been discovered in a traditional quantitative study based on the opinion that has been dominant ever since Ravenstein (1885) formulated his laws of migration behaviour about 140 years ago – in a nutshell, Ravenstein’s laws say that people are driven away by privations and attracted by advantages (external factors). A quantitative study would have been *etic* (based on concepts imposed from the outside), macro-analytic and inductive (concerned with using specific people for the purpose of discovering general laws), and closed, and thus incapable of responding to the themes of separation, guilt, alienation and the like which *emerged* from the *emic*, micro-analytic and open study conducted by Cropley and Lüthke (1995).

In a study carried out in Germany with people who had applied for a visa to emigrate to Australia (Cropley & Lüthke, 1995)

almost all respondents nominated as decisive unfavourable economic conditions, fear of unemployment, shortage of accommodation, fear of war, desire for a better future for their children and desire for better weather. However, a more intensive investigation with the same people using group interviews, structured individual interviews and in-depth interviews found that under these conditions (as against when they filled out a questionnaire) Germans interested in migrating to Australia gave little weight to external, objective factors, such as those just mentioned, apart from wanting to go somewhere where life conditions were marked by prosperity and freedom (i.e., none of them wanted to migrate to the then Soviet Union). In in-depth interviews respondents revealed the “psychodynamic” meaning of emigration for them. Examples include: satisfying an urge felt since childhood to experience strange places; acting on a feeling of needing a challenge in a new environment; realizing a desire to escape from suffocating family ties. Many of them saw emigration as an aid to solving problems of intrapersonal development, focusing on processes such as separation, attachment, or individuation (e.g., Erikson, 1968).

A review of the contents of counselling for would-be migrants showed that it focused – as would logically be expected from the prevailing opinion just outlined – on providing them with information on unemployment rates in their target countries, average summer and winter temperatures, availability of affordable housing, and similar practical matters. Both interesting and useful, this information helped migrants formulate an apparently logical explanation for their interest in emigrating: As one man put it in an interview: “You’ve got to have something sensible to tell the neighbours about why you’re emigrating!” However, such counselling did nothing to help migrants identify and cope with

their uncertainty, guilt feelings, anxiety, or sense of loss once they were in the new country.¹⁵

Following the criterion of usefulness, Cropley and Lüthke (1995) worked out implications of their findings for practice, in their case psychological practice, and designed a one day workshop for in-service training of social workers involved in counselling both potential migrants in Germany before departure as well as actual migrants in the target country after arrival there. The researchers' recommendations included both guidelines for the content of the counselling, such as thematizing difficulties of attachment/detachment, or discussing how realistic it was to hope that migration would solve all personal problems, as well as for the organisation of the counselling sessions; for instance, in the research interviews those participants who took part in group interviews where other would-be emigrants were also present reported that these interviews were very helpful, as they felt that for the first time they were among people who understood their situation, which meant that the fact that tough-minded reasons for wanting to emigrate were largely a façade (something of which they had not previously been aware) could be identified and openly discussed. This was both an emergent, "accidental" finding which had not been part of the original research plan at all and is also an example of how qualitative research can sometimes benefit participants, on the side as it were.

In a methodologically similar study, Knapper and Cropley (1980) showed that in difficult traffic situations such as near accidents drivers interpreted the behaviour of other drivers on the basis of visible clues such as age, style and level of dress, age of the other person's vehicle, visible level of maintenance of the vehicle, or price group of the vehicle when it was new. Drivers then made inferences about invisible factors such as the intention of the other

¹⁵ A good 20% of Cropley and Lüthke's respondents returned to Germany within two years of emigrating. The record short stay involved two brothers, who never left the airport in Sydney, but booked seats on the next flight back to Germany!

driver (e.g., “The other driver is a respectable old gentleman who probably tried hard to avoid causing me any trouble but was trapped by external circumstances” vs. “The other driver is a careless young idiot who couldn’t care less about me”). Knapper and Cropley (1980) showed how their findings indicated that people “construct” traffic situations as miniature social interactions and regulate their own behaviour according to these subjective interpretations, rather than according to the objective facts of the situation. For instance, interpretations similar to the “respectable old gentleman” reaction produced less aggressive responses, as would be socially appropriate when dealing with such a person (respect for older people, protective attitude to the elderly), whereas the “young idiot” construction encouraged a desire to teach the other person how to do it properly or to punish him or her, as is more appropriate for children. Knapper and Cropley related this finding to social psychological processes such as attribution or selective perception. They also worked out practical implications of their findings for driver training and for measures aimed at encouraging more widespread use of seat belts.

In a “descriptive-phenomenological” study of how parents of autistic children understand their children’s disability, Probst (1997) showed that they use four main thinking categories¹⁶ to make sense of the situation: “categorization” (What is the nature of the disability?), “attribution” (Why did it occur? Who or what caused it?), “expectations about the progress of the problem” (What consequences will it have for the child? Is it treatable?), and “expectations of social support” (Who will help us and how will they do it?). These ways of looking at the situation were related by the researcher to existing theory in

16 In Chapter 8, I will discuss the term “category” in greater detail. In specific studies this term may be too abstract for my purposes in this book. For example, Probst’s labels for blocks of closely related ideas for understanding children’s disabilities (“categorization,” “attribution,” “expectations about the progress of the problem,” and “expectations of social support”) are probably better described as “themes.”

clinical and health psychology on “health schemata,” and to more general basic psychological theory on causal attribution, locus of control and cohesion. At a practical level, the findings were discussed in relation to the structure and function of parent support groups and to psychological processes in such groups, and based on his findings Probst designed a workshop for parents which was very positively received by participants. The kind of insights into the understandings and needs of families with an autistic child yielded by this qualitative analysis can be contrasted with those delivered by a later quantitative study by the same author in which he reported that there was an effect strength of .75 when children who received therapy (i.e., experimental group) were contrasted with children who received no therapy (i.e., control group). The findings of the latter (quantitative) study demonstrating the effects of therapy are by no means uninteresting, but such a finding is quite different in principle from those of the earlier qualitative study.

Plate (1997) investigated “the conditions that make it possible for people to make effective use of their own creativity in their work,” by asking employees what factors they experienced as encouraging their own creativity and what factors discouraged it. He showed that his respondents identified both “external” and also “internal” factors that played a role. External factors included level of demand for creativity within the job, degree of freedom in the work setting (e.g., independence, tolerance of errors) and presence or absence of time pressure. Factors within the person included “image” (the workers’ perception of how others saw them), “identification with an idea,” “motivation” or “demotivation,” “willingness to take risks,” or “willingness to risk conflict with other people.” These two sets of factors worked together to encourage or discourage creativity. Both sets led fairly directly to suggestions for

a creativity-facilitating management style.

Cropley (2002) conducted group discussions with a total of 70 German social workers working with asylum applicants, asking them about their motivation for work in this particular area and about the aspects of their work that caused them particular problems. Participants believed themselves to be particularly qualified because of their special skills, and consciously or unconsciously used work with refugees as a chance to relive and resolve problems from their own past lives. Among participants' cares and concerns about their work were existential issues such as finding a sense of purpose in it or finding ways of escaping from the isolation resulting from their total absorption in the day-to-day struggle. Other concerns were more professional (e.g., obtaining new perspectives, both theoretical and practical, seeing how to work effectively and objectively). In interpreting his data the author showed that in many ways the social workers in the study resembled people interested in emigrating, and connected this finding with the more general theory of personality and migration already mentioned above (e.g., separation, attachment, individuation). He also drew attention to the implications of the study for in-service training in the area.

The Relationship of Quantitative and Qualitative Methods

As was pointed out in Chapter 2 (see Table 2.1), quantitative approaches are strongly linked with experimental and quasi-experimental designs. It is thus tempting to regard quantitative approaches and experimental designs as synonymous, although this is not necessarily the case. In a similar way, quantitative methods almost always involve numerical data, often collected

using electronic or mechanical instruments, psychological tests, questionnaires, or some combination. They are also almost always associated with formal hypotheses and with the testing of such hypotheses (i.e. deductive reasoning), especially with the help of analytic statistics. By contrast, qualitative methods are strongly linked with collection of verbal data by means of, for instance, interviews. They are frequently conducted without hypotheses, at least in the initial stages (i.e. inductive reasoning), and the dominant approaches to data analysis are variants of content analysis (see Chapter 8). There is very frequently no statistical analysis, and the findings may be more in the form of hypotheses than of conclusions: Qualitative designs may have more to do with finding hypotheses than with testing them. Thus, there is *a correlation between research method and the kind of information that is delivered*. This fact is of particular interest in research involving “mixed” models (see Tashakkori & Teddlie, 2010), which seek to combine the two kinds of information delivered by the different approaches in order to gain a clearer picture of the phenomenon being investigated (for a more detailed discussion see Chapter 6).

Complementary relationship of the two approaches: The fact that qualitative and quantitative approaches do not compete is made clearer by a consideration of the fact that scientific inquiry typically starts with observation of a case or cases and then proceeds to the subsequent formulation of an explanation via inductive thinking, typical steps in qualitative research. The theory is eventually formalized and its applicability refined via testing of hypotheses and deductive thinking, processes that are typical of quantitative studies. Thus, the two approaches *supplement* each other.

A simple example is the anecdote about Isaac Newton and the discovery of the laws of gravity (this is useful to illustrate the point being made here, even if it is not historically accurate). Newton is said to have been sitting under an apple tree in his

garden one day when a ripe apple fell from the tree and struck him on the head. This event set Newton thinking about why objects fall to the ground. In other words, he started from an observation of his own experiences in everyday life. He constructed an explanation of his observations and this led to the emergence of what was then the novel theoretical concept of gravitational force. This sequence of events can be regarded as an extremely simple case study involving observation of real life, description, explanation, and generalization. The apple was observed to fall to the ground. Other objects do this too when they are not supported. Therefore, the earth must somehow pull them, i.e., there must be a force acting on everything.

The process just described involves using specific observations to formulate a general rule: i.e., it is *inductive*. Naturally, even prior to Newton people knew that objects fall to the ground when they are not supported. What was missing, however, was a mathematical-physical explanation of why this occurred (i.e., a theoretical explanation of observed facts formulated in terms of the system of concepts employed in physics). An explanation based on a different system of concepts already existed, and Newton's had to compete with this. According to the classical Aristotelian explanation, the universe was constructed according to a systematic order of things. An object would fall to the ground when its place in this order required it to do so. This explanation fosters a sense of living in a sensible universe, but does not permit prediction, only after-the-fact explanation: The apple fell because that was the right thing for it to do. What Newton did was to develop by induction a set of rules describing why objects behave as they do. This model yielded hypotheses about how new objects that were not part of the original observations (such as the moon) could be expected to behave. These hypotheses could be tested and, if necessary,

the original theory refined (i.e., deductive thinking). Newton's theoretical model served as the basis for subsequent quantitative research but was itself the product of an essentially qualitative approach.

Quantitative research, by contrast, commences with an existing body of knowledge that already includes generalizations and explanations, i.e., theories about relationships among specific events. A simple example would be a study based on existing findings about facilitation of the growth of intelligence by contact with a verbally enriched environment. A typical piece of quantitative research would find a way of measuring intelligence (IQ scores) and a measure of degree of contact with a rich verbal environment such as amount of formal schooling. In an *ex-post-facto* design two groups could be formed, one that had already experienced a high level of formal schooling (experimental group), a second with little such prior schooling (control group). If there were a statistically significant difference between the groups' mean IQs it would be concluded that the research hypothesis had been confirmed. The logic of this approach starts by stating a general rule (the hypothesis about a connection between verbal environment and intelligence) and using it to make a prediction about the specific samples: i.e., it is *deductive* in nature.

However, as was pointed out above, in empirical scientific thinking the very first general statement is often developed via observation of concrete experience – historically speaking, induction usually precedes deduction. What this means for the theory of research methodology is that qualitative and quantitative research are not rivals but complementary phases in the process of enlargement of knowledge. An intuitive model of their relationship would see a qualitative phase of induction as the first step in enlarging knowledge, followed by a quantitative phase of deduction. These two phases can be regarded as linked by an intermediate phase in which the knowledge gained via observation and description in the inductive, pre-experimental, pre-statistical phase becomes known

and stimulates hypothesis formulation, thus laying the foundation for quantitative research. This relationship is shown in Table 3.1. The table can be read by going down the columns “Purpose” “Action,” and “Result.” This shows how, to take the example of “Results,” expansion of knowledge moves from description and development of possible explanations (qualitative phase) to the formulation of specific hypotheses (intermediate phase) and finally to the statistical testing and subsequent rejection, acceptance or modification of these hypotheses (quantitative phase).

Table 3.1: The phases of scientific enlargement of knowledge

Phase	Characteristic			
	Process	Purpose	Action	Result
Qualitative Phase	Induction	Description and explanation	Observation of cases	-Description of general properties of cases -Explanation
Transfer Phase	Application	Switch from qualitative to quantitative mode	Working out implications of previous phase for the next	Formulation of specific hypotheses
Quantitative Phase	Deduction	Testing of hypotheses	Application of inferential statistics	-Confirmation -Disconfirmation -Revision of hypotheses

The reason why the two approaches may seem to be alternatives or even rivals is that quantitative researchers are often unaware of the qualitative phase that has preceded quanti-

tative work. This is because, from the point of view of quantitative researchers, the general theory developed in the qualitative phase can be taken for granted because it simply exists in the literature, as though it had always been known. This is often because the qualitative phase took place a long time ago or because there has been a mass of quantitative research since the original burst of inductive creativity. No physicist working today on a problem such as gravity waves would start by reviewing the case of Newton and the apple, although without his original insights the modern research could not exist. It is thus quite possible for researchers in the qualitative phase to be unaware of earlier phases. This is probably particularly true of graduate students in social science, who typically enter the research cycle in the intermediate phase at the earliest: They begin to work on a research problem when thinking about it has already progressed to a point where the overall outline of ideas is already known, i.e., despite occasional exceptions they are foot soldiers rather than generals. This does not mean that they are condemned to carry out unoriginal routine work: As Miller (2000) and Sternberg, Kaufman and Pretz (2002) pointed out, extending the already known is the commonest form of creativity in scientific settings (for a fuller discussion, see Chapter 9, p. 254).

One way to think about systematic qualitative research is thus to regard it as a formalisation of the inductive phase, even if researchers are unaware of this. The danger of this, however, is that it creates the impression that qualitative methods exist only as a preliminary phase for yielding ideas that can be investigated “properly” in quantitative investigations. It is true that qualitative approaches are indeed especially useful for investigating phenomena about which little is known, such as patients’ image of their therapist or the psychological trauma experienced by political refugees. However, it is important to bear in mind that they also offer a scientifically legitimate way of conducting research on familiar topics. In addition to their usefulness in investigating little known phenomena they are “a valid research genre” in their own

right, as Elliott (1999, p. 251) put it.

Core differences between the two: Nonetheless, there are key differences between the two approaches. These differences were brought out clearly by Kvale (1995) in terms of the concept of validity, as this concept was introduced in Chapter 2. He identified seven *phases* of the qualitative research process, and argued that each phase has its own special criterion of validity (see Table 3.2). These criteria differ markedly from those applying to quantitative methodology.

Table 3.2: Criteria of validity in qualitative studies

Phase	Criterion
Thematising	Sound reasoning in working out the research question
Designing	Clear explanation of the thinking behind the study
Collecting data	Quality of the questioning of participants and checking of their answers
Transcribing	Accuracy of the recording of participants' narratives
Interpreting	Logic of the researcher's interpretation of participants' narratives
Reporting	Accuracy of the researcher's account of what happened
Verifying	Credibility of conclusions. Usefulness of practical suggestions for application in real life

In the first phase of *thematizing*, validity rests upon the soundness of the reasoning that led to the study's research question (referred to above as "identifying the 'object of study'"). In the second phase of *designing*, validity involves the credibility of the methods used to pursue the object of study. The validity of the third phase, the process of *data collection* (in qualitative

studies this often involves interviews), derives from the quality of the questioning of participants and the carefulness with which the information they offer is checked by the interviewer – this is a matter of the “trustworthiness” of both interviewer and respondent. The validity of the phase of *transcribing* focuses on the question of whether the transformation of the respondents’ verbalizations into a written record truly represents what they said. The validity of the phase of *interpreting* derives from the logic of the researcher’s interpretations, while the phase of *reporting* is valid if the research report gives an accurate account of the study and its findings and derives practical implications. Finally, the validity of *verifying* centres on whether specialist readers find that the study reflects their experience in a credible way, whether the findings are surprising, and whether they are useful for specific communities of practitioners.¹⁷

Why do People Conduct Qualitative Research?

The dream: Many research students in psychology and education nowadays express strong interest in conducting qualitative studies and, indeed, “qualitative research methods” has become something of a catch cry. However, the feeling cannot be avoided that students’ motivation is sometimes a flight *from* quantitative methods rather than an attraction *to* qualitative ones. From a purely emotional point of view this is understandable. Quantitative procedures emphasize numerical data, complex mathematical procedures involving probability theory, and extensive calculations

¹⁷ Kivunja and Kuyini (2017, pp. 31ff.) provide a comprehensive summary of the differences between what I refer to here as “qualitative” and “quantitative” approaches to research.

by a machine (i.e., a computer) that uses procedures that are hard to comprehend. Qualitative methods are attractive because they seem to permit an escape from strict experimental designs, close attention to sampling, statistics, and the like, as well as a chance to avoid extensive calculations, i.e., they seem to be in a sense easier and to involve “labour saving.” On a more idealistic level, students can easily feel that the whole process of data collection, coding, and analysis is impersonal and, in a sense, out of their control. It seems to turn the people being studied into mere objects, and the researchers into scientific instruments. Qualitative methods thus offer a humanized alternative consistent with psychologists’ self-image as sensitive, intuitive, warm human beings rather than cold, logical, objective scientists. Qualitative methods are thus more “romantic.”

Indeed, this view of qualitative methods is not without foundation. There are two main reasons for this:

1. In the qualitative approach the *relationship between researcher and person being studied is different* from the quantitative approach. Both parties in qualitative research engage together in a joint process of production of knowledge. Qualitative research has more the nature of a cooperation between equals, whereas quantitative research involves an investigator examining an object (usually called the “subject”) under the microscope. In this sense qualitative methods are more “human.” They involve trusting relationships between people, willing cooperation, and the like.
2. In the qualitative approach *technical requirements of design are usually less stringent* (e.g., sampling, control of confounding variables, reliability and validity of instruments, technical quality of data such as parametric vs. nonparametric, and so on). These can be ignored or at least looked at in a “flexible” manner.

Harsh reality: However, the “romantic” and “labour-

saving” views of qualitative methods do not take account of several issues that need to be raised here. Briefly summarised, it can be said that qualitative methods too require hard work, as well as courage, trust in oneself and determination. They also require a high level of knowledge. There are a number of reasons for this.

1. Qualitative research methods derive from a complex and subtle body of theory on the nature of knowledge and ways of obtaining access to other people’s knowledge that grapples with difficult issues in the philosophy of science and knowledge (e.g., postmodernism).
2. Qualitative methods often require lengthy contact with respondents who must be convinced to sacrifice their time. Outside the researcher’s circle of family and friends, this may not be easy.
3. It may be necessary to spend a good deal of time in fairly intimate contact with people who are not particularly pleasant company. Participants do not always conform to the psychology student’s ideal of YARVIS (young, attractive, rich, verbal, intelligent, and social).
4. The respondents may be asked to reveal themselves freely to the researcher. This may include information that is painful to them. Such situations require special personal properties in the researcher – see below for a more detailed discussion of these properties.
5. Qualitative research raises important ethical questions. For instance, discussing suicidal ideation in an interview with a severely depressed respondent might trigger suicidal behaviour if the respondent were simply abandoned at the end of the interview. Ethical issues are discussed in more detail below and in Chapter 4 (pp. 117ff.).
6. Qualitative methods involve data collection and transcription procedures that require meticulous record keeping, are often extremely time consuming, and may

even become tedious. Much the same may be said of data analysis procedures. All these steps in the research process are of necessity usually carried out by researchers themselves, not for instance by typists or computer programs.

7. Qualitative approaches raise special problems for reporting both procedures and findings. In writing a thesis or dissertation students may have to forsake tried and trusted structures and devise their own.

8. There are thorny issues involving the reliability and validity of knowledge obtained through qualitative approaches. Usually no hypotheses are tested and nothing is “proved” in the traditional, quantitative sense. This raises the problem of how to draw conclusions that can be generalized to situations other than the one in which the data were collected – how to make findings *plausible* and *useful* in both scientific and also practical terms.

9. The outcomes of qualitative research are often unclear until a very large amount of effort has been expended. Researchers may spend long periods in a state of uncertainty about what, if anything, is going to come out of the many hours of work.

These issues may be summarised by saying that qualitative research methods are far from cut and dried and far from offering the “easy” alternative some research students hope for. Indeed, as Denzin and Lincoln (2017, p. 11) put it, they are often “messy” and “uncertain.” However, provided that this is accepted from the start, with sufficient effort all the problems just mentioned can be surmounted.

Who is qualified to do qualitative research?

Qualitative methods are not merely a matter of finding someone to talk to and writing up the results. It is necessary to recruit respondents and gain their help, as well as remaining aware of the possible consequences of participation for them. It is also necessary to be able to analyse qualitative data (see Chapters 7 and 8) and

to “make sense” of these data in psychological terms (Chapters 8 and 9). Consequently, there are certain personal qualifications for conducting qualitative research. These are found in specific areas.

- First comes *special knowledge of the field*, especially of the basic concepts psychologists have developed to understand it. In quantitative approaches the theoretical (conceptual) framework and hypotheses (i.e., the research questions) have usually been worked out and specified beforehand, largely by other people. In qualitative studies they must be worked out by researchers themselves.
- *Qualitative studies require inductive thinking* to generate new structures. These must be derived from clues and hints provided by the respondents, usually in everyday language, not social science terminology. This requires sharp observation, alertness to the “meaning” of clues, the ability to build concepts, and creative but disciplined fantasy. In quantitative research, by contrast, new material is incorporated into an existing framework.
- The *researcher’s own personality* is also important. Necessary properties include empathy and openness, as well as the ability to adapt oneself and to inspire trust in others. In addition, qualitative researchers need fantasy and creativity, as well as courage, perseverance, ability to tolerate uncertainty, and self-belief.
- *Good judgement and responsibility in dealing with other people* are necessary. This involves skilful selection of respondents, and a grasp of how intensely or how far to pursue material, as well as of when to stop, especially when probing procedures such as in-depth interviews are used.
- *An ethical attitude is important*. Great care must be taken to obtain informed Consent and avoid negative consequences of participation for respondents, for instance in dealing with intimate or painful material. The example of research with respondents who display suicidal

tendencies has already been given.

Of course, factors such as knowledge of the field, fantasy and creativity, sense of responsibility, and an ethical attitude are also important for researchers doing quantitative research. However, they assume special importance in qualitative studies.

Responsibility and ethicality may be particularly important in clinical research. To take an example from the past, in the early days of encounter groups it was not uncommon for beginners to encourage respondents to participate in group discussions of topics of extraordinary personal significance for particular members of the group, such as coping with the loss of a loved one, dealing with disordered thought processes, or fantasies of rape and murder, and then at the conclusion of the group simply send these people home to deal as well as they could with the emergence of such material into consciousness. Of course, clinical research is not alone in the danger of confronting respondents with material from their psyches that they subsequently have difficulty in handling. However, this is a particular danger in clinical studies since, by virtue of being clinical, they often concentrate on material that is causing difficulties for respondents. As a result, the personal qualifications for qualitative research may well be of particular importance for clinicians, although this should not be interpreted as meaning that the qualifications are not important in other areas or that clinicians are especially unethical.

Chapter 4: Designing a qualitative study

In qualitative studies, the initial steps in designing a research project – defining just what will be investigated, selecting an appropriate data-collection method and identifying appropriate participants – are different from quantitative investigations. Some authors even object to general discussions of how to carry out qualitative studies, on the grounds that such discussions imply the existence of a fixed qualitative paradigm that can be applied objectively and in a way that does not vary from researcher to researcher. Objections of this kind seem to imply that *reliable* and *valid* qualitative research is impossible. However, many researchers adopt a more moderate position, accepting for instance that there are “local” realities that transcend individual people and can be described and analysed in a systematic way. The present chapter outlines

the “phases” of qualitative research (attending, telling, transcribing, analysing, reading, and validating) that need to be carried out in an appropriate way in order to collect reliable and valid qualitative data. Some of these raise ethical issues that researchers need to address (e.g., preservation of confidentiality, informed consent, avoidance of risk for participants).

Chapter Glossary

<i>cognitive dissonance:</i>	a situation in which two mutually contradictory “facts” both seem to be true at the same time
<i>data shift:</i>	a change in what is perceived as a fact which results from a change in the context in which statements about the facts are being made (see also “social desirability” below) methodological limitations on the researcher’s freedom of action made necessary
<i>constraints:</i>	by the need to avoid the effects of factors like social desirability (see item below)
<i>meaning shift:</i>	the subjective meaning of objective events changes as people process, store and re-tell them
<i>narrative competence:</i>	the ability to provide useful information on a particular topic, often a topic related to the person’s own life

<i>placebo effect:</i>	a tendency for people to experience states and feelings simply because they have been led to believe that they will do so
<i>quantum of effort:</i>	an investment of time and effort into investigating a topic which is great enough to provide sufficient useful information about the topic
<i>social desirability:</i>	A psychological factor putting people under pressure to give answers which will please other people, regardless of what the person giving the answer actually thinks

The Phases of Qualitative Studies

According to Riessman (1993, p. 10), a qualitative research study has five “levels,” although since they describe steps in a process I refer to them as “phases.” These are as follows:

- *attending*: the people who later become participants pay attention to what is going on around them and build up an internal picture of the world;
- *telling*: these people tell the researcher about this internal picture (i.e. they provide a narrative). This often has the form of an interview, although other forms of telling are possible such as letters, texts of speeches, an autobiography, etc. These narratives are recorded in some form, usually by the researcher;
- *transcribing*: the narratives are transformed into a written text by the researcher;
- *analysing*: the transcriptions are analysed by the researcher, via for instance content analysis;

- *reading*: the results of the steps just outlined are communicated to appropriate readers, for instance in the form of a thesis or dissertation, or a research report.

A sixth phase can be added to these five

- *validating*: knowledgeable readers examine the contents of the report and decide whether it is stable, credible, useful and so on (i.e., whether it is reliable and valid).

As with any system of phases, these are arranged in order, starting with “attending” and proceeding through the intermediate phases until “reading” and “validation” are reached. The first three phases form a cluster focused on the *respondents* as they:

- a) experience life,
- b) develop a subjective understanding of their experiences, and
- c) select what they will pass on to a researcher.

Together with aspects of “transcribing,” the phases of “analysis” and “reading” form a second cluster focused on the *researcher* who:

- d) records what the participants say,
- e) develops an understanding of the general implications of these statements, and
- f) communicates these understandings to other people in a research report.

A third cluster focuses on the “*audience*” which consists of people who

- g) read the research report (phase of reading) and
- h) possibly accept and adopt its findings (validating).

1. Attending: *People organize and order their experiences and store them in their minds.* This is not simply a photographic record of everything, because attending is selective and interpretative (i.e., the people *construct* their records of their own lives). Thus, “meaning” *shifts* away from an objective, camera-like record to a subjective one;

2. Telling: The phase of telling corresponds to data

collection. However, telling looks at data collection from the point of view of what the participant does, not what the researcher does: the participant “tells,” the researcher “collects.” *When asked to narrate their experiences people select aspects that seem relevant or have special importance to them.* They relate these to the researcher. They do not simply regurgitate everything that has ever happened to them in connection with whatever the researcher is investigating. They usually present the aspects they select in a relatively organized and structured manner. They also shape what they tell according to the particular audience. As a result, meaning shifts a second time;

3. Transcribing: *The narrative is “fixed” by the researcher* by being written down or recorded in some other way such as tape recorded or videoed. Riessman (1993) made the point that videotaping preserves the richest record, although this raises substantial practical problems such as its demands on the researcher’s time. However, regardless of how the narrative is initially fixed, ultimately, a written record is almost always created. This record does not contain every piece of information provided by the respondent, but involves selection and organization on the part of the researcher – once again, meaning shifts;

4. Analysing: *The researcher makes sense out of the narrative(s)* by examining the content of one or a number of transcriptions. This phase involves “coding” what the participant said and is influenced by the researcher’s expectations, interests, attitudes, political orientation, and the like, so that a further shift takes place;

5. Communicating (Reporting and reading): The researcher writes a *report presenting the results of the analysis and makes it available to other people* (the first

step in the phase of communication), who read it (second step in communication). The contents of the report reflect the orientation of the researcher as just mentioned in 3 and 4. Once again there is a shift in meaning.

6. Validating: *Informed readers assess the report's "usefulness"* in the sense of Chapter 2 (i.e., with regard to credibility and so on). As in the earlier phases, the message the report contains for particular readers is modified by their expectations, interests, attitudes, etc. A final "shift" occurs, this time caused by the reader of the report.

Although Coughlan (2011) was writing about the special case of action research in organizations, his discussion can be applied to the more general task of defining central characteristics of qualitative research. He identified (p. 54) "critical themes" that help to do this:

- a) research involves an *emergent* inquiry process;
- b) data *shift* as a consequence of intervention;
- c) it is not possible to predict or control what takes place in the course of research or where it ultimately leads.

Coughlan's three themes emphasize that researcher and respondent are not subject and object, but are collaborators in a process of discovery, in which the object of study emerges in the course of the investigation, data shift (as just discussed above), and where the research will ultimately lead is not necessarily predetermined or even under the strict control of the participants (as in the airfreight case study). As broad principles these themes apply to all qualitative research, although to differing degrees according to the particular approach involved (see Chapter 5 for an overview of qualitative data collection methods).¹⁸

The relationship between the phases and actions of participants and researchers are summarised in Table 4.1. This phase model of the process that leads to a rigorous qualitative study

¹⁸ This point of view raises once again the issue of what I used to call "openness" but now call "flexibility."

can be compared with Cropley’s (2001) expanded phase model of the emergence of a *creative* product. This model of creativity also involves phases in which an internal picture of the external world is built up, this material is processed, the results of the processing are communicated to other people, and these people judge the results. Cropley also emphasized that without validation by knowledgeable observers neither a “creative” idea nor a qualitative research report satisfies the criterion of “usefulness.” A comparison with the creative process also emphasizes a further important criterion of the external validity of qualitative findings: It is important that they generate “surprise” (see Chapter 9 for a further discussion of this topic), i.e. that they lead to something unexpected. If they do not, the accusation of banality or triviality immediately comes to the fore.

Table 4.1: The phases of qualitative research

Focus on	Phase	Content
Respondents	Attending	Respondents organize and order their own experience and store a record in their minds
	Telling	Respondents select relevant aspects of their experience to narrate to the interviewer
	Transcribing (first step)	Respondents narrate their experiences to the researcher

Researcher	Transcribing (second step)	The researcher writes down or otherwise records each respondent's narrative
	Analysing	The researcher "makes sense" out of each narrative and discerns the general picture
	Communication (first step)	The researcher writes a report and makes it available to knowledgeable readers
Readers	Communication (second step)	The report is read by knowledgeable readers
	Validation	Readers form an impression of the report's contents and assess its usefulness

Factors Affecting Reliability and Validity in the Phases

Attending: Psychologists have paid considerable attention to the *attending* phase, for instance in research on perception and memory, as well as on cognitive and non-cognitive aspects of information processing (receiving, selecting, structuring, storing and recovering information). Research findings show, for instance, how processes such as selective attending or formation of harmonious “gestalts” lead to organizing and smoothing out of what is perceived in order to form “rounded” and closed images in the memory. For practical purposes, however, little if any influence can be exerted on this phase by the researcher, since – because qualitative designs are usually *ex-post-facto* – it normally occurs prior to the commencement of the investigation. Nonetheless, it is important to remember that this phase exists and to remain aware of its implications in the phases of analysing and communicating. By contrast, the threats to reliability and validity that arise in

the remaining phases of *telling*, *transcribing*, *analysing* and *communicating* can be reduced by appropriate planning and execution of research projects.

Telling: Holstein and Gubrium (2015) explained the consequences of the social construction approach for research design by contrasting it with the more traditional idea of research as a technique for “prospecting for true facts and feelings residing within (p. 2).” In the latter, the data function as “a pipeline for transmitting [objective] knowledge (p. 3)” – the data are, as it were, a mirror of an objective reality. Put briefly, the quantitative approach rests on the view that there is a reality “in there” that researchers must dig out or “mine” via appropriate techniques. By contrast, emphasis on social construction holds that “all knowledge is created from the action taken to obtain it (Holstein and Gubrium, 2015, p. 3).” In other words, researchers do not uncover objective knowledge that exists independently of the research process, but construct or create knowledge by the act of research. This means that research *creates* knowledge rather than discovering it. Although it has received increased attention recently, this point of view has been present in modern social science thinking for a considerable time, and was already the subject of a major review well over 30 years ago (Gergen, 1985). More recent psychologically-oriented discussions are to be found in Patton (1990) and Kvale (1996). The conceptualization of research as a process of creation of knowledge through the joint efforts of researcher and participant(s) is stated with particular emphasis by Coghlan (2011).

Social construction of information: Shifts in meaning result from the fact that information is recalled, narrated, recorded and interpreted in a social framework. A simple example is the way in which, let us say, “rich people” is differently understood (constructed) from different social viewpoints. To people who move in one social circle rich people are the carriers of high culture who are engaged in the noble enterprise of preserving the best traditions of society. To people from a different milieu they are

parasites who prey on the majority of society. Finally, to a Marxist they are the owners of the means of production but are doomed to be swept away by the inevitable forces of history. Feminists argue that much of what is regarded as knowledge about women has been socially constructed by the imposition on women of stereotypes and preconceptions stemming from male-dominated conceptions of women's experience. Finally, social construction is even influenced by a group of people who are not directly involved in the research at all – the readers of the report. They too bring with them opinions, expectations, attitudes and the like that influence their understanding of its contents. Thus, no fewer than three separate groups are involved in the social construction of a research project: the participants, the researcher and the readers of the report.

The role of the researcher in constructing information: It is important to distinguish between social construction of information that occurs during the actual collection of data (i.e., in the phases of telling, transcribing, analysing and communicating – see Table 4.1) and that which already existed prior to the commencement of the study in question (i.e., in the phase of attending). *Both respondents and researchers already come to the research setting as bearers of socially constructed information*, because in the past they have “attended” to aspects of their world that social factors tell them are important, significant, interesting, or even correct. It is easy to see that participants communicate socially constructed views of reality to the researcher, but less obvious that researchers' socially constructed reality influences the very questions that they ask, as well as what they regard as worth transcribing and how they interpret the transcribed material. *Social construction thus influences not only what the respondents say, but also the questions researchers ask and the answers they attend to.* Once the data have been collected, the respondents' work is over, but social construction in the form of researchers' pre-existing ideas and expectations continues to affect the development of the research findings and the emphases, form and style of the final report. These

aspects of social construction are of great importance, and influence the process of data collection and analysis, even where they are not actually part of the process itself.

Returning to the actual process of collecting qualitative data, it must be remembered that information is obtained by means of an *interaction* between researcher and respondent. Naturally, what emerges during, for instance, an interview is heavily dependent upon the respondent's views. However, it is also largely *shaped or formed by the researcher* (Briggs, 1986). In the first and more obvious instance this is done by directing the respondents' attention to particular contents that interest the researcher – possibly more than they do the respondent. It is also affected by factors such as the researcher's level of language or the level of abstraction of the discussion. However, as Cicourel (1974) pointed out, social construction of information goes further: Researchers are likely to impose on the respondent ways of conceptualizing experience that are more consistent with the interviewer's construction of these experiences than the respondent's. For instance, if a researcher studying motivation for adults to stay home and care for elderly parents conceptualized this as women's work that is atypical or even odd for men, this construction of the phenomenon being studied would influence the way the topic was presented to respondents for discussion, as well as the concepts and categories they would be encouraged to use in describing and evaluating their own behaviour as carers. Such socially derived preconceptions may well also influence the researcher's selective recording of data, as well as the interpretations placed on data.

The mechanisms of social construction: In addition to coming to the interview situation as bearers of socially constructed information, both researchers and respondents interact and *produce socially constructed information in the interview itself*. The mechanisms through which this occurs are already familiar to psychologists and educators. A concrete mechanism is the phenomenon of “shaping” of behaviour through what is known

as the “Greenspoon effect” (Verplanck, 1956). If interviewers say “Yes,” or “Good,” or display simple, possibly unconscious actions such as nodding when particular contents are emphasized or specific opinions expressed, or even merely look more alert and interested, this can lead respondents to concentrate on statements that elicit such obvious signs of approval. This is a subtle example of the effects of body language in the collection of qualitative data. The problem is particularly pronounced in the case of respondents with a strong external locus of control.

Focusing more on psychological processes within respondents, phenomena such as “Yea-saying” (Couch & Keniston, 1960), avoidance of “cognitive dissonance” (Festinger, 1957), conforming to “social desirability” (Crowne & Marlow, 1964), “model learning” (Bandura & Walters, 1963) and simple conformity (Asch, 1955) play a role in social construction of information during an interaction between researcher and respondent. The general thrust of findings on these phenomena is that there is a tendency for people to agree with views expressed by others whom they perceive as being powerful or as having high social status (yea-saying), or even to adopt these people’s views and copy their behaviour (imitation). Respondents also tend to give answers, state opinions and express attitudes that seem to them to make their own behaviour and points of view appear rational and logical to both themselves and other people (avoidance of cognitive dissonance). In addition, they try to give answers that will please the interviewer and that conform to dominant social norms (social desirability or what is nowadays often called “political correctness” or even “virtue signalling”). The effects of such processes can be very strong.

The two examples given in the box that follows illustrate in a simple and concrete way the social construction of responses in questionnaire studies. These seem to be particularly susceptible to this phenomenon in the sense that they have difficulty in getting behind the façade. However, the research process itself is part of the process of social construction. Although this is undoubtedly also

true of quantitative methods, it is particularly obvious in the case of qualitative approaches, where the research process usually involves a more or less intensive personal interaction between researcher and respondent. What emerges from this interaction is socially constructed information.

Researchers in an early study of the attitudes of members of the public to the wearing of seat belts (Knapper & Cropley, 1980) – at a time when their use was still controversial – found that about 5% of men who claimed in an interview that they *never* wore them had already been observed with their seat belts buckled up prior to being interviewed or were seen wearing them afterwards. Apparently, the people in question did not want to admit to using seat belts because they feared that the interviewer would regard this as unmanly, or not sufficiently “macho,” as we would say nowadays. In their study of Germans applying for a visa to emigrate to Australia mentioned earlier, Cropley and Lüthke (1995) found that nearly all of the respondents initially gave similar “reasons” for wanting to leave Germany: e.g., unemployment, difficulty of obtaining accommodation and high interest rates. However, closer inquiries showed that virtually none of the respondents was unemployed, homeless, or in debt. On the contrary they nearly all had good jobs, a pleasant apartment and few financial problems. In fact, the reasons they gave in initial interviews closely matched lists that had appeared in print and electronic media reports on migration. In the early interviews respondents simply echoed what they had read or heard about reasons for wanting to leave Germany, apparently on the grounds that such reasons must be sensible, rational and socially acceptable, since they had been in the newspaper or on TV, i.e., their “reasons” were socially constructed. By contrast, subsequent in-depth interviews with the same people

gained access to much more personal material and were thus able to break through the façade of social construction.

Dealing with social construction: One response to this phenomenon is to regard it as a source of contamination of the information elicited and to try to eliminate its effects via good technical practice such as sophisticated questionnaire construction or skilful interview techniques. Holstein and Gubrium (2015, p. 4) referred to these as “methodological constraints.” Examples would be avoiding questionnaire items whose meaning is ambivalent, conducting interviews in language that is meaningful to the respondent in terms of level, use of vernacular expressions, etc, or avoiding leading or tendentious questions that put words into the respondent’s mouth. Without denying the importance of best practice in data collection, however, it should be borne in mind that emphasizing these as a technical solution to the perceived problem of social construction still accepts at least to some degree the information-mining point of view, that there is an objective reality which can be obtained in a more or less “pure” form if the extraction procedures are refined enough, i.e., if social construction can be minimized.

The thoroughgoing social construction position, by contrast (e.g., Alasuutari, 1995), accepts that social construction of qualitative data is unavoidable. It does not try to avoid it but seeks to take account of it in data collection, analysis and interpretation. To take an example: A master’s candidate conducted a content analysis of 12 interviews with people who had experienced the loss of a loved one, in order to cast light on the process of grieving and reconciliation. One examiner pointed out (correctly) that the analyses were heavily dependent on the particular person who made them and suggested asking a judge with expert knowledge of the area of loss and grieving to analyse some of the interviews. The examiner’s line of argument was that if the student’s analyses agreed with those of the expert, it

could then be assumed that they were “correct.” However, from the point of view of social construction of information this argument is incorrect. It assumes that there is a single, objectively “correct” meaning of the information yielded by the interview and that all sufficiently skilful researchers would come to the same conclusions about the contents of the interview. In other words, the examiner was adopting an essentially positivist position. A theorist like Alasuutari, by contrast, would point out that differences between analyses would be a unique product of a particular interaction between an individual researcher and a particular respondent, and the interpretation of these contents would be specific to this interaction.

Key Aspects of Qualitative Research

Previous sections have shown that the qualitative approach is based on special ideas about the nature of data and the effects of collecting data on the data themselves. These have great significance for the way in which qualitative studies are conducted. In designing such studies the researcher must decide:

- just what is to be investigated
- whose constructions are to be examined
- what method to employ for gathering data
- how to record information
- how to transform these records into a narrative that can subsequently be analysed
- how to analyse the data
- how to report the results to interested people

These steps correspond to the steps of “telling,” “transcribing,” “analysing” and “communicating” in the list cited at the beginning of the chapter. Logically they follow each other in order, although – unlike quantitative designs – there is a tendency in qualitative

research for them to mix with each other, to overlap or to alternate: For instance, the phenomenon being investigated may be changed after some data have been collected, or additional participants may be recruited after initial results have been obtained. Thus, because of the *openness* of qualitative research earlier decisions on points in the list of necessary decisions just outlined may be changed in the course of the investigation

Levels of qualitative research: Qualitative studies can be conducted at three *levels*. These are listed below in a hierarchical order – it is possible to conduct research at the first level without taking any account of the issues involved in levels 2 and 3, while research at level 2 does not require familiarity with the issues at level 3.

1. *Quasi-qualitative studies* emphasize “What kind?” as against “How much?” but otherwise retain the properties of a quantitative study more or less intact (sampling, measurement, statistical tests, etc). They treat the phenomenon being investigated as a more or less real thing that respondents possess and that causes their behaviour. However, they focus more on the style or kind of way in which the phenomenon is expressed rather than on the amount or level. An example would be a study dividing respondents into “Type A” or “Type B” personalities (probably on the basis of a psychological test) and showing that Type A people react differently to certain stimuli from Type Bs. For instance, a fictitious study might show that As are likely to suffer psychosomatic illnesses, whereas Bs more frequently suffer mood disturbances. Partially qualitative studies are usually based on one or more hypotheses that are tested in the usual way with appropriate statistics such as Chi-square, i.e., they are deductive.

2. *Phenomenological studies* also assume that the object of study is more than a unique private construction

in the mind of a particular individual. However, the main interest of such studies is in discovering how the phenomenon being studied is experienced by the people in the investigation. A phenomenological study of, let us say, parents of children who suffered serious brain damage at birth might ask how it “feels” to be such a parent: How do these people see their situation (for instance as a curse or as an opportunity to serve another human being)? How do they explain it (e.g., as a punishment from God, as a result of their own foolishness, as a random accident that could have happened to anybody, or as the result of technical incompetence on the part of doctors and nurses)? How do they view the future (hopelessness, optimism)? Such studies are mainly descriptive, do not usually involve numerical data, and do not test hypotheses (except for mixed approaches – see Chapter 6).

3. *Radical qualitative studies* are based on a rejection of positivism and, indeed, of the idea that there are objective facts at all. Theorists adopting a radical position argue that the phenomena psychologists study do not exist as specifiable entities at all, but are invented by the people experiencing them or the researchers studying them (cf., postmodernism). The task of the researcher is to interact with the respondent – most frequently by means of an interview, although numerous other possibilities exist – and “build” or “reconstruct” the respondent’s ideas.

Discussions in this book emphasize the need to investigate social science phenomena *from the point of view of the people experiencing them* and accept the idea of the social construction of information. However, they are not radical in the sense just outlined in 3 above. This book does not reject the idea of reality or assume that it is impossible to make generalizations about it. On the contrary, it will show how to collect reliable data and draw scientific conclusions using qualitative methods. What follows is concerned

with the question of how to design studies that do this.

Defining the “Object of Study”

The first step is to identify the phenomenon that is going to be studied, i.e., to specify what the research is going to be about. There is a substantial and important difference between the ways this is done in qualitative and quantitative research. A traditional quantitative study is *always based on theory* – as a matter of necessity, the theory precedes the research project. This is because of the deductive nature of quantitative research: A particular study rests upon a pre-existing hypothesis, which is tested in order to ascertain the applicability of the theory in a new setting or under new conditions. Without the theory there would be no hypothesis and thus no research. More than 50 years ago Meehl (1967, p. 103) warned about “not relying on ad hoc explanations” (i.e., formulating a hypothesis only after the results are already known – HARKing¹⁹, as it is now called). Such hypotheses are of course irrefutable, since they are only formulated after the results are known, so that, in the case of quantitative methods, adjusting the null hypothesis after you know the results is highly reprehensible. By contrast, qualitative studies may commence without any theory at all, as in the example below, and it is normal to derive theory after you have data. Nonetheless, even where there is no formal theory, *it is essential that the research has a clearly recognizable focus and specifiable goals*. In other words, it must have an “object of interest” (Hamel, 1993, p. 41), or an “object of study,” as I prefer to call it. In a sense, the clear statement of the object of study in a

¹⁹ HARK = “hypothesis after results known.” *Open science* now insists that a hypothesis must be stated before the collection of data begins, thus eliminating “accidental” findings such as the discovery of penicillin by Alexander Fleming as a result of forgetting to put away bacteria samples before leaving the lab for a brief vacation.

qualitative research project has the same function as the research hypothesis in a quantitative project.

In a recent dissertation, Eidson (2021, p. 1) started by giving an example of a clearly stated object of study in a qualitative research project: “The purpose of this qualitative descriptive study was to *explore the experiences* of working mothers in the financial services industry in the United States ... after an extended child-rearing leave” [italics added]. By contrast, a quantitative study might have commenced with a statement similar to the following: “This study tested the hypothesis that a sample of women who were mentored upon their return to work after an extended child-rearing leave would display a significantly higher level of self-efficacy on the Cropley Self-Efficacy Test (CST)²⁰ than a control group whose members received no mentoring; the null hypothesis was $M_{\text{mentored}} = M_{\text{unmentored}}$.”

Kinds of object of study: The object of study can be what Taylor, Bogdan and DeVault (2015, p. 25) called “substantive,” or it can be “formal.” In the case of substantive research, the purpose is to gain *in-depth understanding* of a particular setting or to provide an exact and detailed description of special patterns of behaviour in a particular situation. To take an example, substantive research might investigate the patterns of friendships in a prison²¹: Who is friendly with whom and why? The object of study is an aspect of social interactions in the prison and the desired result is knowledge about it; research of this kind is often referred to as “action research.” By contrast, quantitative research seeks to develop *theoretical insights* into a behavioural science phenomenon.

In the example just given, a researcher might be interested in investigating the relationship between personality and friendship or between attitudes and friendship, or in looking at psychological factors that affect the process of forming friendships. These are

²⁰ This instrument is purely fictitious.

²¹ Again, this example is not based on an actual research study but has been constructed to aid understanding.

the object of study: The purpose of the research is to improve understanding of these issues, not to describe the prison setting more exactly, although such a description might emerge as a valuable by-product. The prison is not chosen as a setting because prisoners “represent” the general public, but because the alliances represented by friendships have a particular significance for people in prisons and such people can therefore be expected to know a lot about them, i.e., to possess “narrative competence.” Although, especially in the case of action research, understanding of the specific situation of prison life is more likely to be the goal of the research, the study might lead to suggestions for revision of some aspects of social psychological theory such as theory on formation of peer groups, on opinion formation and change, on the development of power relationships, etc.²²

Formal objects of study are not forbidden. Taylor, Bogdan and DeVault (2015) argued that whether to pursue a substantive or a formal version of the object of study can be decided once the researcher has had some experience with one or more interviews. The main thing is for researchers to “get their feet wet” (p. 27) by hearing what the first one or two cases (respondents) say. However, research in social science typically requires more than a simple descriptive study. Thus, in designing the study researchers would normally expect to collect data going beyond substantive issues and permitting at least rudimentary formal conclusions. In discussing qualitative methods, Strauss and Corbin (1998) referred several times to the special position of psychology and education and seemed to imply that the strong quantitative tradition of these disciplines makes development of theory especially important for qualitative research in them. Thus, in the case of educational and psychological research, especially that of master’s and doctoral candidates, there is typically more emphasis on formal than on substantive objects of study, i.e., researchers seek to go beyond description of a setting to

22 For example, Petersen’s study of hobby authors (see pp. 266ff.) led her to suggest a clarification of Maslow’s hierarchy of needs.

obtain understanding of underlying processes and to relate these understandings to existing theory, ultimately even expanding or refining that theory.

The examples of qualitative research questions on pp. 68ff. help to clarify what is meant by a *formal* object of study. For instance, in the first example it would be “motives for emigrating,” in the second “attitudes to seat belts,” in the third “understanding of disability.” The hoped for ultimate result of research would be not merely a list of motives or feelings but an understanding of how these relate to each other and to general psychological theory in these areas, as well as discussion of what consequences the findings in connection with migrants, drivers or parents respectively have (a) for psychological understanding of motives, feelings and the like and (b) for practical psychological/educational work with groups such as would be migrants, people refusing to wear seatbelts or parents of children with a disability – this could involve counselling, training, educational measures, etc.

The source of the object of study: An important issue here is that of *where researchers obtain ideas about an object of study*. The criterion of coherence already mentioned (e.g., p. 52) demands a consistent internal logic in working out the object of study. In quantitative research the usual source is existing knowledge, and the purpose of the study is to show that that knowledge also applies in a new situation. To take a concrete example, a master’s student of mine at a Canadian university was reading research reports on socioeconomic status and intelligence one day. These emphasized existing findings showing that contact with the dominant culture (whose norms, attitudes and values are reflected in intelligence tests) affected test scores. He realized that towns, small villages and remote settlements within the university’s region offered settings with different degrees of such contact. Thus, quantitative thinking (deductive thinking – reasoning from the general to the particular) suggested that mean IQ scores of samples of children should become progressively lower as the children’s place of residence

progressed from city to country town to small village to isolated bush settlement. Since it was also known that the cultural content effect is stronger in the case of verbal intelligence than non-verbal, it would also be expected via a process of deduction that mean verbal IQs would diminish more markedly than non-verbal.

Appropriate representative samples of children living in the various settings were obtained, intelligence tests administered and scored, and mean IQs calculated. The null hypothesis that there were no differences between group means was rejected: There was a steady reduction in mean IQs of the samples, starting with the city group and proceeding in order from the country town to the remote village to the children living in the bush. Furthermore, this reduction was greater in the case of verbal than non-verbal IQs. It was therefore concluded (Cropley & Cardey, 1975) that the theory about IQ and cultural contact had been confirmed in a novel setting, thus demonstrating its explanatory power.

In the case of qualitative research, the *coherence* criterion of validity is particularly important. As with quantitative research, the object of study can derive from existing research findings and can rest upon fairly specific expectations that may even take the form of clear hypotheses. In the example just given, for instance, it would have been possible to have investigated the idea, let us say, that people in remote rural areas value different knowledge and skills from those in the city (for instance, in a bush camp it is more important to know how to bait a fishing hook properly than to know what to do if somebody shouts “Fire!” in a crowded room). Thus, a more interesting research question might be that of what knowledge and skills children from different life settings regard as important, and why. Ultimately, the research might lead to a general conclusion that there is no difference in the *amount* of intelligence (mastery of socially relevant skills) displayed by the different groups of children, but substantial differences in the *kind*

of skills necessary to get along well in the various life settings. A general conclusion might then be that there are differences in the relevance of conventional intelligence tests to different life settings.

However, in qualitative approaches a particular object of study can also be the result of everyday hunches or intuitions, simple curiosity, or personal experience. For instance, an educational psychologist with a highly intelligent child might be moved to investigate how such children regard school, or what teachers think of clever pupils. Someone who has recently recovered from the experience of losing a loved one might wish to investigate how people cope with such a situation. To return to the example of German migrants to Australia, I became interested in this topic while working at the University of Hamburg. I received several phone calls in the space of two or three months from Germans who were contemplating emigration to Australia and had heard that I was an Australian. The callers were eager to obtain information about that country, as well as advice. These people were in every case successful individuals with a well-paid job. I contacted the Australian Immigration Centre, and they advised me that there had recently been a very large surge in applications for visas. This raised the question of why people who were experiencing economic prosperity wanted to emigrate to a country about which they knew so little, and the object of study became “Reasons for emigrating.” Subsequently, with the help of the Australian Immigration Centre, I was able to interview a number of visa applicants. Analysis of the interviews led to recognition of the fact that applicants’ motivation was often unclear, even to themselves, and hidden beneath a layer of rationalizations aimed at avoiding cognitive dissonance. This indicated the necessity of in-depth interviews to investigate such otherwise hidden material on *motivation*, which had now been identified as a more focused object of study as a result of the first interviews.

The data collection method was accordingly changed to in-depth interviews, and the sample was widened to include not just

visa applicants but also actual migrants and finally, people who had emigrated but had not been able to deal with the new surroundings and had returned to Germany. Ultimately, an understanding was developed of the structure and function of motives for emigration, and this led to challenges to existing theories of migration such as the cost-benefit approach (i.e., that people migrate to avoid things they do not like and to gain things they desire, seeking to maximize gains and minimize losses), which were labelled “external” in nature. Instead, an “internal” model was proposed that sees migration as simply one possible way of dealing with developmental tasks involving aspects of personality development such as individuation, autonomy and the like (Cropley & Lüthke, 1994).

In this example, personal involvement played a role in identifying the object of study, since I was myself a migrant, as well as being an Australian citizen, while curiosity piqued by everyday experience (the phone calls) was the trigger that set the research in motion. It was conducted without prior hypotheses, and the focus on dynamic motives for emigrating and the subsequent internal model of migration behaviour emerged in the course of the investigation, “surprising” the researchers (Folkert Lüthke and me), who had not set out to distinguish between internal and external motivation or to demonstrate the connection between migration and personality development. The study also took on some of the characteristics of action research, something which had not been intended when it commenced. The problem of disappointment, failure to settle in the new country and quite frequently relatively rapid return to the homeland, usually with a feeling of shame and failure (as well as financial loss) emerged and aroused considerable interest. Investigating this concrete and for “failed” migrants specific personal problem became an unforeseen aspect of the project: Folkert and I then worked with such migrants and with counselling agencies both in Australia and in Germany to establish the dimensions of the problem and to work out appropriate responses in pre-migration counselling in Germany and post-migration counselling

in Australia. To establish the coherence (validity) of the research, the steps in this evolution of the object of study were described in detail in the research report, in order to bring out their systematic and logical nature.

The Sample in Qualitative Designs

In quantitative research, a failure to sample adequately poses a serious threat to both internal and external validity. A thorough examination of sampling theory and the issues and problems it raises would go beyond the aims of this book: A brief basic discussion will suffice for present purposes. In quantitative research, the fundamental idea of sample construction is *random sampling* in which every member of the entire group being studied (the population²³) has an equal likelihood of being selected into the sample and only chance determines who is in and who is out. When the population is not homogeneous but consists of identifiable subgroups – such as, let us say, people aged over and also people under 18 years of age – participants can be selected by chance and added to the sample until the proportion of the subgroup in question (for example people over 18) is the same in the sample as in the population. After this, no further people over 18 are added to the sample. This approach results in a *structured sample*, which in the example just given is structured by age. A sample is only structured in terms of variables that are known to be likely to influence the matter being investigated. A similar underlying idea is involved in *probability sampling*, *stratified sampling*, *quota sampling*, and *cluster sampling*.²⁴

23 From now on, when the discussion touches on quantitative methods I will refer to the entire group as the “population.”

24 A helpful compact and down-to-earth discussion of sampling techniques is to be found at <https://www.questionpro.com/blog/types-of-sampling-for-social-research/>

Narrative competence: In qualitative research, the situation is different. Although it remains extremely important to make plain in a research report who provided the data in question, sampling matters are often given less prominence than in quantitative studies. Nonetheless, constructing an appropriate “sample” remains a matter of considerable importance. In effect, qualitative researchers select people who the researchers think are well placed to cast light on the particular issue that is of interest to the researcher. Knowledge of the object of study takes precedence over sampling theory. Participants may even be deliberately chosen not randomly but because it is already known that they belong to a sub-group whose views on the object of study are atypical (see for example the discussions of “theoretical sampling” and “key informants”).

At the heart of the qualitative approach (a participant is in the sample because he or she knows a lot about the matter being investigated) is an important idea: *narrative competence*. This is the view that all people are competent to describe their own life and say how they understand it.²⁵ As Breuer, Muckel and Dieris (2018, p. 76) put it, all individuals are:

basically able to reflect upon and talk about themselves, about their connections with their objective, social and spiritual-cultural environment, about their world-views and interpretations, their actions, their life stories and their social-historical links [my translation].

An example can be seen in a study of poverty and the ways in which poor people understand their situation and their relationship to the mainstream society. Instead of concentrating on experts on poverty, such as economists, social workers, unemployment counsellors, even workers in soup kitchens, the researchers concentrated on talking with actual poor people, who were found to possess a rich

Another helpful discussion is at <https://www.scribbr.com/methodology/simple-random-sampling/>

²⁵ Whether this understanding corresponds with what other people think is irrelevant.

and differentiated understanding of their situation and to be able to communicate this to researchers; *they possessed narrative competence*. However, although it obviously makes sense to talk with people who are intimately involved with the object of study, two further issues in deciding with whom to talk about the object of study are of particular concern. These are discussed below.

Representativeness: In a quantitative study the object is to broaden the applicability of an existing theory by showing that it holds up for a new population (external validity). The sample from which data are obtained must therefore “represent” the people in the population in whom the researcher is interested. To take an example, a study of the general theory that there is a connection between children’s socioeconomic status and their intelligence might test the null hypothesis, “There is no difference between the mean IQs of samples of children of high and low socioeconomic status.” In such a study, the children labelled “low socioeconomic status” would have to represent the population of children of low socioeconomic status (i.e., they would have to be typical of such children), while those labelled “high socioeconomic status” would have to *represent* children of high status (i.e. be typical of such children). If in a study in Canada the high socioeconomic-status children all came from Quebec they might not represent Canadian children in general, with the result that any difference that was found might be caused by differences in mother tongue, not by socio-economic status at all. As a result, language would be a *confounding variable* that would not only prevent generalizing to all Canadian children (external validity), but would also obscure the relationship between socioeconomic status and intelligence, thus also threatening internal validity.

However, the issue is rather different in qualitative research. Here, the crucial question of representativeness can be seen in a different light: It is not a matter of, “To what population can I validly generalize from this sample?” but “What is the object of study for which these cases are appropriate?” Put in a less formal

way, the key question is, “Are the members of this ‘sample’ (the people with whom I am talking) well qualified to tell me about my object of study?” “Do they know a lot about it?” “Do they possess *narrative competence*?” As Hamel (1994) put it, when looked at from this point of view the representativeness (or not) of the sample will immediately be apparent to anyone who is familiar with the object of study. For instance, it is obvious that, in principle at least, applicants for a visa to emigrate from Germany to Australia are in a good position to tell a researcher about reasons for applying for a visa to emigrate from Germany to Australia, that classroom teachers possess relevant information about the effects of class size, or that people in gaol know a lot about power structures in prison. Thus, visa applicants are a “representative” sample for a study of reasons for wanting to emigrate, while classroom teachers are a “representative” sample for a study of what goes on in the classroom and people in gaol know more about being in gaol than people who have never been incarcerated.

Key informants: This aspect of sampling can be discussed more formally under the heading “key informants.” The purpose of key informant methods is to collect information from people who have a high level of first-hand knowledge of the object of study: Put in a down-to-earth way, people who know what is going on. Such people are *key informants*: they are chosen not on a random basis but because of their specialized knowledge. The most obvious example would involve experts, specialists, or professionals in a particular domain. Possession of key information can thus be thought of as a particular form of narrative competence (see above). The special value of key informants is that their particular knowledge and understanding mean that they can provide insights into a situation which go beyond what is available to “everyday” participants. However, as Lokot (2021) emphasized, the views of experts are likely to be strongly influenced – focused or even narrowed – by their expertise in a specific field: For instance, when the COVID-19 pandemic was first identified, I (as a psychology

professor) was immediately interested in its effects on people's attitudes to work, whereas people with other areas of focus were interested in, for example, its economic effects on the travel industry. Thus, different key informants should either be recruited from highly diverse knowledge backgrounds in order to combat the effect of specific expertise, so that the concepts or themes they use to understand events can be applied broadly or, where key informants all come from a specific field (such as psychologists or economists), the specificity of their themes and concepts should be kept in mind when drawing conclusions.²⁶

Sample size: A practical question that graduate students engaged in qualitative research frequently ask is how large their sample needs to be. There is a kind of informal, seldom specifically articulated law that the number of interviews required is inversely related to the intensity of each contact: It is as though there were a kind of “quantum of effort” that must be made. This can sometimes be achieved by studying a single case, for instance by conducting 100 hours of systematically recorded and transcribed interviews with, let us say, a famous inventor. Sometimes it requires a substantial number of respondents (e.g., 30 or more two hour talks with individual respondents). Thus, the necessary sample size partly depends upon *the intensity of contact with the individual case or cases under study*.

It also depends to a considerable degree on the object of study – *what the investigation is about*. Obviously, a study of Picasso's views about his own creativity would require only a single respondent – Picasso is an appropriate “sample” for studying Picasso! However, according to the quantum theory mentioned above, there would need to be an increase in the intensity of the contact, since $N = 1$. Shaughnessy, Zechmeister and Zechmeister (2014) present an introduction to the nature of single-case studies

26 Although in this paragraph I have discussed the key-informants approach in the context of sampling, it is sometimes discussed as a special method of data collection.

that is very helpful for those wishing to conduct such a study.

A third criterion for deciding on the necessary size of the sample is the *breadth of the issue being studied*. Sometimes a single case is sufficient to construct a convincing explanation of a very broad issue: Newton probably did not need to wait for several apples to fall on his head before developing the idea of gravity! Paradoxically, a broadly applicable explanation may require few cases, whereas a narrow explanation may require more. Had Newton theorized more narrowly, let us say, that gravity only acts on fruit, he would have required more cases. An interesting concept in this context is that of *theoretical sampling* (Glaser & Strauss, 1967). Although this idea will be discussed more fully on p. 215, because it is more closely linked with data analysis than with sampling (despite the name), it can be summarised here in a special sense as indicating that the sample is big enough when, in the analysis phase, the addition of further cases brings no change in the conclusions being drawn (i.e., “convergence” is reached – see p. 205). The word “theoretical” refers here not to a theory of sampling but to whether or not additional cases make it necessary to gain clearer understanding of a theory that has emerged from the earlier cases already analysed.

Obtaining access to participants: Often, researchers are not free to pick and choose the sample by strictly random methods or in a systematically structured way, for example because they do not have access to every single member of the population. Several procedures exist that can be applied in order to cope with this difficulty. For example, the researcher may make use of *convenience sampling* (whatever members of the population are available for whatever reason are included in the sample). A variant of this is *self-selected sampling*, which frequently involves simply calling for volunteers. A further variant is *snowball sampling*, where people already in the sample bring the researcher into contact with friends, acquaintances, fellow members of organizations, etc., who are willing to participate.²⁷

²⁷ Nowadays, it is possible to recruit participants online and to then construct

Ethical Issues in Research Design

The kinds of research on which this book concentrates always involve human beings. In addition to issues such as how to obtain data in ways that enhance their reliability and validity, working with people in order to investigate their subjective constructions of the world raises *ethical* issues such as avoiding harm – not only physical but also psychological – to the people involved. These issues are present in quantitative research too, but they take on particular importance and assume special forms in qualitative studies because these involve revelation by participants of their personal constructions of the world and may thus touch upon intimate material in a particularly personal way. The example of research with patients at risk of suicide has already been mentioned, and this example could be added to without effort. Among a number of organizations, such as the Professional Association of German Psychologists and the German Society for Psychology (BDP) or the Australian Psychological Society (APS), the American Psychological Association (APA) has published guidelines for ethical research.²⁸

Confidentiality: Participants have a right to expect that data obtained with them will not be published or otherwise revealed in a way that makes it possible to identify individual respondents. It is usual to regard it as ethical to publish findings that refer only to general tendencies in a group. However, the matter is more complex than at first seems to be the case. For example, in a study of attitudes of police officers in Canada a university lecturer analysed essays about First-nation Canadians submitted as a class

the sample by focusing on the people who know a lot about the object of study (i.e. who possess narrative competence): useful links include Positly, Prolific, CloudResearch, Respondent, and FindParticipants.

²⁸ A focused discussion of these issues is to be found in the *Ethical Principles of Psychologists and Code of Conduct* of the American Psychological Association, which can be assessed at https://www.apa.org/images/ethics-code-2017_tcm7-218783.pdf

assignment by experienced officers undertaking in-service training in criminology and concluded (rightly or wrongly) that many of the officers were profoundly racist. She then openly discussed this impression in other classes and even in public. No individual officer was identified, but considerable harm was done to the public image of the police force, while all officers in the seminar in question were stigmatized as racist. Furthermore, the policemen felt betrayed by their instructor, and their willingness to improve their qualifications as well as their acceptance of behavioural scientists' findings on crime and crime prevention were greatly diminished. The research probably did great harm. Obviously, more is required than simply confidentiality in a formal or technical sense.

Deception: In the past it was common for participants in research studies to receive incomplete information about the purpose of the research or their role in it, or even for them to be actively deceived. Many famous studies of the past in, for instance, social psychology were based on misinforming the participants about the situation, an example being studies in which participants were led to believe that other people present in a group situation were also students who had been recruited as they themselves had for a research study, whereas the other group members were really actors whose job was to make the genuine participants doubt their own senses. In the example in the previous section there was also deception, since the researcher lied to the police officers about the reason for asking them to write the essays and about what she intended to do with their narratives.

Consent: Great emphasis is placed nowadays on obtaining research participants' consent. However, consent is meaningless unless it is *informed*, which, among other things, precludes deception. Although this seems obvious, being informed turns out to have many facets: For example, (a) it is impossible to inform people about unforeseen risks; (b) lay people are not in a position to estimate the severity of risks that they do not understand; (c) risks may be direct or indirect. Detailed discussion of risks may also

frighten potential participants and convince them that events that are extremely unlikely to occur are in fact probable. (Readers of this book who have been warned by a doctor prior to an operation that there is, let us say, a 1 in 1000 chance of dying under the anaesthetic and have asked themselves if they really need the operation will recognize this phenomenon.) Psychologists also face the problem that informing a participant that psychological consequences could follow can, in some cases, ensure that the consequences do in fact occur, for instance via the workings of suggestibility and *placebo* effects. This seems to suggest that deceiving participants may sometimes be helpful to the researcher or even necessary, despite the fact that deception is now regarded as unethical. Thus, consent too has come to be seen to be a complex concept.

Some of these issues take on particular significance when applied specifically to qualitative research. Focusing on the issue of harm to participants, it is clear that while discussing their own construction of reality with an interviewer may not cause immediate physical damage to respondents (whereas receiving an electric shock may), such a discussion may set in motion processes such as self-questioning. These may continue to have an effect long after the interview has ended, and may have unforeseen, perhaps unforeseeable consequences. This means that it is impossible for the researcher to be absolutely certain that no harm can occur, or for the participant's consent to be fully informed. This issue becomes particularly acute in the case of in-depth interviews, whose purpose is often to bring into awareness material that has been hidden under a layer of rationalizations, repressions, and similar psychological mechanisms, possibly because of its highly unpleasant nature, its painfulness, or the threat it poses to respondents' sense of worth or wellbeing. The respondent may have a strong interest in **not** uncovering such material! Fontana and Frey (1998) went so far as to conclude that most in-depth interviewing is unethical. If the position of Fontana and Frey were accepted, much qualitative research would have to cease. However, some practical steps are

available to researchers to reduce the risk of harm.

Avoiding harm: In selecting participants, appropriate attention should be paid to the stability and robustness of their personality, avoiding recruiting respondents who are obviously disturbed or fearful. Interviews should not be pushed to the point where they cause distress to respondents. Researchers who investigate contents that are particularly affect-laden should offer participants the opportunity of appropriate *debriefing*, although without creating the impression that the interview is primarily for therapeutic purposes or seeming to be engaged in recruiting clients for therapy, or indeed actually doing this.

Fortunately, practical experience shows that participating in a qualitative study can have positive effects for participants, even if these are not actually part of the intention of the researcher. One way such consequences can arise is through the opportunity participating in an investigation can offer for dealing with otherwise seldom-confronted material. The investigation involving Germans considering emigration that has already been mentioned several times provides a concrete example. Participants emphasized that participation had significantly assisted them in coping with their situation. Again, when such unintended beneficial side-effects occur it is important to distinguish between a research study and a treatment. Beyond the benefit for individual participants, such effects provide a source of insights into potential applications in practical situations, i.e., they increase the usefulness of research. For example, Cropley and LÜthke (project: German emigrants) developed counselling strategies for people struggling to make a decision, and conducted training seminars for counsellors in several German cities, although such practical application was not part of the original research plan

Assessing the cost-benefit ratio: Shaughnessy, Zechmeister and Zechmeister (2014) summarised the relationship between possible risks to participants and potential value of a project in a useful way. In principle, every study involves dangers

for the participants, even if in most cases the risks are infinitesimally small. The existence of risk means that a research project not only should, but also must offer some prospect of usefulness – it must be externally valid. This means that ethical considerations require not only considering and avoiding foreseeable harm, but also specifying what possible *benefits* could flow from a research project. The researcher should work out the possible benefits and weigh these against imaginable risks, thus calculating the risk-benefit ratio. Only where this ratio is high (i.e., potential benefits are very high in comparison with risk) can a study be regarded as ethical. If there are no foreseeable benefits, the research has no external validity, and the question must seriously be asked if it can ethically be conducted.

Chapter 5: Collecting data in qualitative research

In collecting qualitative data, the fundamental unit of observation is the “case.” As a rule, a case is a human being providing information about how he or she understands the world, and the information provided is referred to as a “narrative.” Data most commonly consist of spoken or written words, although other forms of communication exist (including but not confined to non-verbal data such as body language, artwork, etc). By far the most common method for collecting such information is the interview, which may be unstructured, semi-structured, or structured, while other ways of collecting data such as participant or non-participant observation, case studies, or shadowing exist. Furthermore, researchers do not always interact directly with participants, but sometimes study already existing sources such as autobiographies, diaries, newspaper articles, films, videos, advertisements, or even

artworks, “borrow” material from other researchers, or even collect virtual data. The main thing is that narratives reveal the “lived details” of people’s lives. This means that in face-face-data collection (usually an interview) researchers must take pains to avoid leading interview partners by sending verbal or non-verbal signals that elicit particular answers.

Chapter Glossary

<i>autobiographical narratives:</i>	accounts of personal experience in some specific setting that is of interest to the researcher
<i>case studies:</i>	“intrinsic,” “collective,” “multiple,” and “instrumental” case studies involve a single individual or more than one individual who is observed under specified conditions
<i>commonplace evidence:</i>	evidence derived from people’s direct experience with life
<i>covert observation:</i>	observation in which the people being observed do not know that this is the case; covertness raises ethical questions such as the issue of informed consent
<i>indirect data:</i>	data which were not provided by the respondent specifically for the particular research project
<i>leading questions:</i>	questions whose wording already tells the person being questioned what answer is expected or required
<i>multiple-case study:</i>	an issue is studied by combining

	information from more than a single case.
<i>prompts:</i>	questions or requests by an interviewer which help the person being interviewed get started or direct attention back to the object of study without leading the respondent (see leading questions)
<i>second-hand data:</i>	data obtained by somebody else and “inherited” by the researcher
<i>shadowing:</i>	the researcher accompanies a single participant everywhere the person goes in the course of the activity that is being investigated (such as a normal working day)
<i>virtual data:</i>	data which are specifically elicited for a particular research project but without direct contact between researcher and respondent (e.g., via the internet)

The Nature of Data in Qualitative Research

The fundamental task of qualitative researchers is to *describe* the everyday life of people as it is constructed in the minds of those people, and to increase understanding of this construction using the scientific concepts of a particular discipline or group of disciplines (Crang & Cook, 2007, p. 4). “Action” research goes further and applies qualitative methods to the solving of specific concrete problems in real-life behaviour. Qualitative research thus

concentrates on “experiential” or “practical” knowledge (Heron, 1992). What is needed is what Hamel, Dufour and Fortin (1993, p. 31) called “commonplace evidence” derived from people’s direct experience with life. Most frequently this requires accounts from people actually experiencing or possessing experience of the aspects of the world the researcher is interested in, and the accounts must be true to the context in question. Collecting data on such knowledge involves methods for obtaining “rational verbal reports of experience.” Appropriate data collection methods must be in a position to reveal the “lived details” of people’s lives (Scheppele, 2004, p. 401).

What do such data look like, and how can researchers collect them? Although some approaches such as action research emphasize that this interaction is a two-way street (i.e., information flows in both directions), with researcher and respondent cooperating to achieve a mutually worked out understanding of the situation being studied, in its simplest conceptualization the interaction involves a researcher obtaining information about a situation in which he or she is interested from a respondent who is well informed about the situation. The information flow may simply involve the researcher using information that is already available such as for example, letters or diaries, and it may include non-verbal data such as artworks, but in the most obvious situation the researcher *elicits* information from the respondent, very frequently, but not always, by means of an interview (see the next section). This information constitutes the data in the research study.

The physical form of data: Although the fact that data can take other forms (such as film or photographs, videos, artwork or even dance) must be emphasized again, the practical fact is that in most studies the information is provided in *verbal* form (i.e., in the form of a narrative), although narratives may be:

- spoken (as in the case of interviews)
- written (e.g., biographies, diaries, letters)
- published in diverse media (e.g., official documents,

newspapers or magazines, books)

- in electronic form (e.g., films, tapes, videos, or digital media).

When verbal statements are recorded specifically for the purposes of a research project, the data may be written down by the researcher or by the respondents themselves. Where respondents speak their accounts (as in interviews), these are very frequently (although not always) recorded in the form of written notes made by the researcher or by electronic means (e.g., tape recorder, video). Prior to being analysed, records are almost always transcribed into written form, although written data may be supplemented by tapes or similar material. Indeed, some writers give considerable emphasis to the importance of flanking data – such as video recordings – in the analysis of data. To sum up, no matter what form the original narrative takes (interview, written document, video or audiotape, digital), the data in qualitative studies most commonly, although admittedly not always, consist of passages of text – at least eventually, even if they do not start that way.²⁹

Using existing data: Although it represents a relatively unusual approach, it is possible to carry out qualitative research without actually collecting any new data at all! This is done by *taking over narratives that already exist*. One of the most obvious ways of doing this is to make use of written personal statements such as diaries, autobiographies, or letters, in which people have already recorded their impressions or their experiences prior to the commencement of the research without any knowledge or intention connected with the research. It is also possible to use as data other highly personal statements such as speeches, even sermons and the like, that are less private than the sources mentioned above because they were meant for public consumption. Public material like government decrees or policy statements and similar public

²⁹ The process of turning observations into text and text into generalizations is largely guided by a *codebook* specifically constructed for a particular project, even when the initial material is not verbal in nature (see pp. 201ff. for a more detailed discussion of codebooks).

documents that may reflect not the views of an individual but those of an organization or agency such as a political party, a firm, a club or a society can also be analysed. Related sources are official documents such as school, hospital or police records: An example reported on p. 142 involves an analysis of the place of creativity in 11 national school curricula. Using the published form of each national curriculum as an individual “case” or “narrative,” the researchers identified content units related to creativity, and combined these to form seven creativity-related “concepts” or “themes” which occurred in at least several of the narratives (curricula). Other researchers work with novels, films or soap operas, even artworks. Advertising material can also be treated as data, as can newspaper reports, radio or TV broadcasts. In theory, all of these can be analysed via variants of the procedures to be described in Chapter 8.

Second-hand data: It is also possible to use “second hand” data (e.g., data collected by someone else, usually for some other purpose). An example is to be seen in Gray and Kunkel’s (2001) study of ballet dancers in the USA shortly after a bitter dispute between dancers and employers. The data were journalistic reports of interviews with dancers published nearly 20 years before Gray and Kunkel became interested in the topic in question (Gordon, 1983). This study is discussed in greater detail on pages 282ff. Such data are particularly useful in psycho-historical research, where the collection of fresh data is impossible: An example would be a study of speeches and other public statements of world leaders in the period immediately preceding the outbreak of the Second World War in 1939, in order to ascertain their state of mind at that time. The evaluation of existing material proves to be particularly useful for historical investigations in which the collection of personal data would be completely excluded, for example because the people being studied have long since died. Freud’s (1910) study of Leonardo provides a good example.

“Virtual” data: A recent addition to the possibilities for collecting data involves using the internet. This approach is

neither fully “direct” nor fully “indirect.” There is no physical contact between the researcher and participants, as in for instance interviews, but, unlike genuinely indirect methods such as analyses of diaries, the researcher deliberately elicits or provokes the responses of the participants, something which is typical of direct contact methods such as interviews. Using the terminology of modern information technology, it thus involves what Andersone (2002) called a “virtual” discussion or a “virtual” focus group (see below for a brief introduction to focus groups).

In a study of the social effects of computer games Andersone (2002) posted several rather confrontational and essentially negative statements about such games in an internet chat room and invited other participants to respond. She received 216 written replies in which the people reacted directly to her message and expressed their views on the games, despite the fact that she had no physical contact with them – hence the term “virtual” discussion. These responses were subjected to a form of content analysis (see Chapters 7 and 8) which identified a number of general dimensions that showed how players of computer games and bystanders construct the games: (a) “computer games as a problem” (e.g., the games cause conflict between husbands who are dedicated players and wives who feel neglected); (b) “values” (e.g., much of real life is futile, but there is something decisive about the games); (c) “fatigue and depression” (e.g., games offer an escape from a feeling of being burned out by the stresses and strains of real life); (d) “hedonism” (e.g., games offer a chance to have some fun), (e) “recreation” (e.g., games offer the chance of a rest from intellectual challenge or the mirror image: certain games offer an intellectual challenge); and (f) “communication” (e.g., through the games it is possible to avoid having to communicate with other people or, in the case of certain games only, games foster communication).

Nonverbal data: I have just stated that the data almost always consist of verbalizations. There is, however, a further kind of data that is easily overlooked – not language but *body language*. Respondents' narratives, especially when they take the form of an interview, are often accompanied by facial expressions such as smiles (sometimes even audible laughter), or the opposite, grimaces, scowls, even tears. There may also be gestures such as waving of the hands, pointing of the fingers, balling of the fists, etc., or body movements and attitudes (e.g., tensing of the muscles, refusing to look at the interviewer, sitting stiffly in the chair). Interviewers can make notes on such behaviour as it occurs. Of course, data of this kind cannot be gathered where there is no direct contact between researcher and respondent, such as when existing documents are analysed, since observational data require face-to-face contact between researcher and respondent, as well as a researcher who is not totally preoccupied with recording the respondent's remarks. Thus, recording respondents' behaviour on videotape is particularly valuable, because it makes it possible to collect nonverbal data without the necessity of making extensive notes at the time of the interview. Such data can greatly enrich descriptions and interpretations resulting from analysis of qualitative data.

Methods for Collecting Data

Although a number of alternative ways of collecting qualitative data have just been mentioned, the remainder of this chapter focuses on collecting new data via direct observations of people in relevant settings. (Where data already exist, as is the case with archive material, literary or artworks for instance, the main problem is not collecting them, but analysing them, and this issue will be taken up in the next chapter.) A variety of methods exists for doing this, and some of these will now be discussed.

Participant observation: Better known and more commonly applied outside psychology and education, for instance in sociology, anthropology or ethnography, is the method of *participant observation*. This is (Jorgensen, 1989, p. 9) “... a strategy for gaining access to phenomena that commonly are obscured from the standpoint of a nonparticipant.” The key to participant observation (as against non-participant observation), is that the researcher collects information within a social milieu by taking on a role in that milieu. This usually occurs in such a way that it is not apparent to the people being observed that it is happening (i.e., the observation is *covert*), except perhaps for a small number of people who are “in the know.”

Buckley (1975) spent a year working as a staff member in a German *gymnasium* (academic secondary school) at a time when schools in Germany were involved in a process of transformation from authoritarian to more “democratic” methods of self-government. This process involved weakening of the power of the school principal and senior staff and strengthening of the role of students and non-academic staff such as laboratory assistants and even janitors, the latter comprising a group that had not previously participated in the decision-making process. During the year Buckley kept written notes of behaviour and statements at staff meetings, in committees, in the classroom and in recreational activities such as sport, even relaxation in the pub after meetings or other events. These written notes were his data. Subsequently, he wrote these up in a chronologically organized account using pseudonyms and other mechanisms to protect participants’ identities. This entertaining account can be read almost as a biographical novel, but its scientific value is that it gives many concrete and “human” examples of the tensions and contradictions, alliances and enmities, fears and other

feelings, as well as the coping strategies that accompanied the change process.

Buckley's study demonstrates several aspects of participant observation emphasized by Jorgensen (1989). He had a definite role in the system he was observing (he was a teacher at the school in question), while, in addition to notes he made on his observations of people's behaviour, much of the data he gathered were collected in the course of informal conversations or via casual questioning that took place in work or work-related settings where the respondents were not aware that they were providing data. The need to fulfil a real role means that participant observation requires skill in using the everyday interactions required by the role in two ways: (a) as "normal" interactions with colleagues, fellow group members and the like (in the role of participant); (b) as sources of data (in the role of researcher).

Thus, ordinary conversations become, in a sense, interviews, without ceasing to be conversations. The actual data usually take the form of written notes made by the observer, although documents can also be analysed (e.g., letters, memoranda, minutes of committee meetings, policy statements, etc.), while tape-recordings and videotapes are also possible, probably covertly. It is thus necessary to routinize writing up of notes or other forms of information recording, for instance by writing them up every evening. A major problem, however, is that nowadays, participant observation of this kind (i.e., covert) raises obvious ethical issues, especially the question of *informed consent* (see p. 118), which appears to be partially or completely lacking in many studies, as well as that of *deception* (see pp. 118), which seems to be a necessary part of such research, especially covert research.

Non-participant observation: Non-participant observation is well known in empirical classroom research and has often been used in the past in, for example, studies in the style of

Flanders (1970). Much of what was said in the previous paragraphs applies in non-participant observation too (e.g., use of flanking material, disciplined regular writing up of protocols or other kinds of notes). In many such studies the frequency with which certain specific behaviours specified in advance by the researcher (such as a child seeking to attract the teacher's attention in order to answer a question) is *counted*. Such studies have many of the characteristics of quantitative research (e.g., they are *etic* – the researcher specifies in advance what data will be recorded, probably on the basis of existing theory – and data are recorded in numerical form and null-hypotheses tested). For this reason, they will not be discussed in detail here.

Shadowing: A special variant of the observation approach which is more participatory than observation as described in the previous paragraph but less participatory than the approach in the Buckley study is *shadowing*; this is especially (although not exclusively) appropriate for use in studies of organizations. During shadowing the researcher accompanies a single participant everywhere the person goes in the course of the activity that is being investigated (for example, a normal working day). The shadower is present, for instance, in meetings or presentations where the participant takes part or is even in charge, observes conversations or interviews with managers or colleagues, listens to telephone conversations, even sits nearby when the participant is working at his or her desk or workplace. The researcher writes up a protocol which contains both notes made in the course of the shadowing and also other sources of information (such as documents or audio- or videotapes). A memory protocol can also be written up each evening.

In principle, two kinds of data are collected in the course of shadowing: (a) data pertaining to what the participant or other people did or said, and (b) data on the reasons why the participant behaved in the way he or she did. This latter kind of data consists as a rule of statements made by the person being shadowed, either

spontaneous remarks (e.g.: “That’ll show him who’s boss around here!”) or else answers to specific questions posed by the researcher. Thus, shadowing is *holistic* and *context specific*. It focuses not only on concrete behaviours in a specific situation (context), but also on the subjective reasons for those behaviours (e.g., wishes, feelings, opinions, attitudes, values, ambitions, and the like). These are directly observed both in the specific context in which they occur and in the exact way they actually occur there, not for instance through reports made by people discussing a situation as they recall it, and certainly not through ticking items on checklists.

Autobiographical narratives: Autobiographical narratives are accounts of personal experience. The study of autobiographical narratives makes it possible to understand better how people construct the meaning of their lives (Bruner, 1990) and develop “narrative identity” (Ricoeur, 1991, p. 73). A central tenet of this approach is that during the course of actual experience – regardless of whether it involves ordinary routines, great adventures, or traumatic events – at the moment of its occurrence the experience usually has uncertain or ambivalent meaning, or even none at all (Gusdorf, 1980). It is the subsequent narration of the experiences that gives them meaning. A structured narrative “assigns a cohesive meaning” to the events that are being recounted (Vikis-Freibergs, 1991). An autobiographical narrative can be brief and focused on a specific event, or it can be a lengthy narration about extended sequences of events spanning decades or even an entire lifetime. The narrative can be spoken or it can be written, but by nature it is a story with a beginning, a middle and an end. This does not mean that the narrative needs to follow a classical plot structure, but rather that it begins and finishes at specific moments in time. The contents can be presented as fiction or as fact; if they involve the narrator’s own “lived experiences” we speak of an *auto*-biographical narrative, also commonly referred to as a “life story” or a “life history.”

The etymological roots of the term “autobiography” hold

several keys to its function and significance. It has three elements: “*auto*” (concerning oneself), “*bio*” (involving the person’s life) and “*graph*” (presented in written form). Literary theorists such as Gusdorf (1980) and Olney (1980) noted that prior to the past several decades, autobiographical narrations, particularly in the form of written texts, were considered to be relatively accurate or veridical reflections of a person’s life (i.e., of the “*bios*”). Following in the footsteps of Aristotle, memories were considered to be direct copies of personal experience “isomorphically imprinted in the soul” (Casey, 1987). Nowadays, however, the relationship between an autobiographical narrative and a person’s actual concrete experiences is conceptualized differently, although there are differences according to various philosophical positions.

This book accepts the hermeneutic position (see page 225ff. for a more detailed discussion), which asserts that actual experience “elicits” or “constrains” the story, to be sure (i.e., the narrative is derived from lived experience), but that the story also “articulates and modifies the experience” (Widdershoven, 1993) – people’s narratives of their experiences provide those experiences with meaning and also affect how subsequent events are understood. Discussion during the past several decades has thus emphasized the “*autos*.” Autobiographical narration is not regarded as a photographic re-presentation of what has been experienced in the past, but rather as a process of construction focused on the self, a “creation of the self by the self” (Gusdorf, 1980). Olney (1980) made the same point in a slightly different way by concluding that an autobiography “half discovers, half creates” the self. French post-structuralists such as Roland Barthes have contended that the most important aspect is the text – the *grapheme* – and that “... the subject is merely an effect of language.” However, the view taken in this book is that the construction of an autobiographical narrative entails a weaving together of the *bios*, the *autos*, and the *grapheme*.

Sebre (1992) provides a helpful example of an autobiographical study. A total of 10 sibling pairs now living in the United States who fled Latvia for Germany in 1944 were asked to describe the experience of flight. The interviews were structured on the basis of specific questions. Examples include: "What can you remember about the Soviet occupation of Latvia?" Or "How did you react when you realized that your family had to flee from home?" The written transcriptions of the interviews were subjected to a kind of content analysis. An initial unexpected finding was that there were often large differences between the narratives of the individual members of a particular sibling pair, even though both participants had experienced the same events side by side. These differences related to different emphases on specific aspects of events as well as differences in the emotional significance of the events described. For example, one member of one pair experienced the singing of the Latvian national anthem as the ship left Riga as a truly tragic moment, whereas the other sibling saw it as hypocrisy. Sebre found that in some respects there were greater similarities among participants belonging to the same age than between siblings belonging to the same family and concluded that the way such events are subjectively experienced is strongly age-dependent. According to Sebre, the "distortion" of interview statements through the effects of social construction poses a particular problem in analyzing biographical narratives. She recommended a hermeneutic approach that constantly seeks not only to sink deep into the data and be absorbed by them, but also to constantly ask, "What does the text mean overall?" Although the text is deconstructed into content units – as in most qualitative investigations – the hermeneutic approach requires that the overall meaning of the entire text be kept in mind when interpreting deconstructed fragments of it.

Case Studies

As Creswell (2013, p. 97) put it, case studies examine “a real-life ... bounded system”. This system *stretches over time and place*. As a result, case studies emphasize – more strongly than most other qualitative approaches – the temporal aspect (a case study could involve data stretching back over a number of years) as well as the location aspect (data could involve a number of settings or places). According to Merriam (1998) they are experiential and evocative because they are so closely linked to real-life events and people’s feelings about them. Case studies were particularly prominent in sociology in its earlier years and were widely used to develop general theory (see Hamel, Dufour & Fortin, 1993). In sociology or anthropology the “bounded system” that is studied (the case) is often (a) a broad, informal social group such as drug users or football fans; (b) a more clearly defined group such as nurses, ballet dancers, or police officers, or (c) a formal, narrowly defined group (e.g., a particular school class). Stake (1998, p. 88) distinguished among three kinds of case study:

1. *the “intrinsic” case study*, which focuses on a single case (i.e., $N = 1$), because of the overpowering interest of the case in question. Picasso, Einstein, DaVinci or Freud would be examples of “overpowering” cases;
2. *the “collective” case study* which aims at enlarging understanding of the collective of cases (as against an individual person), as in the example of a study of friendships in prison already discussed. What special processes determine interactions among members of the collective and what factors guide these processes? Can the social psychology of the collective of prison inmates be compared with that of other collectives such as professional football players?

3. the “instrumental” case study, whose aim is to gain insight into a phenomenon or to refine theory, rather than to study a particular person or group. The case or cases studied are selected because they are particularly useful for refining theory, not because of their overwhelming intrinsic interest as people. The “classic” example is Ebbinghaus’s (1885) study of memory and the processes of forgetting. Ebbinghaus – who was simultaneously researcher and “case” – observed himself over a period of several months, during which he learned lists of nonsense syllables, forgot them, then relearned them. Ericson was not studying the properties of an individual person (in this case, himself), but the nature of memory. Almost 100 years later Ericsson, Chase and Faloon (1980) conducted a similar study, also with a single case. In these two examples, it was not the particular person who was the object of study but the psychological phenomenon of memory. In a certain sense the researchers took as the case any convenient person who was available.

A further example of an instrumental case study is to be seen in the study of Riddoch and Humphreys (1987) on the steering of visual recognition of stimuli by stored information. The case they studied was a man who had suffered a brain injury and, as a result, experienced great difficulty in recognizing people, despite the fact that he knew them well (e.g., members of his family). This study is particularly interesting because it demonstrates the use of a clinical case to cast light on “normal” processes. The particular person involved (the case) was not chosen by chance, because not every potential participant would have been equally informative, but despite this it was not the particular person who was of interest, but a general psychological phenomenon (i.e., the case was an “instrument” for investigating a phenomenon, not the actual object of study).

In psychology and education the focus is frequently on individual people. Thus, to use the terminology of quantitative methodology, in case studies in psychology and education it is common that $N=1$. The question of sample size is, of course, automatically answered when $N=1$ and, remembering the “quantum of effort” theory (there is an inverse relationship between sample size and intensity of treatment of the individual cases, i.e., the smaller the number of cases the greater the intensity with which each case is examined), it is obvious that the depth of treatment must be substantial when a single case is investigated.

The data obtained from a case study may be in the form of a narrative provided by the person in question (such as an interview protocol) or notes yielded by observation of the “case” (see the discussion above of observational approaches). In some studies, the data have been “second hand” in the form of diaries and letters, autobiographies or biographies, or similar documents. In one famous study Allport (1965) investigated the mother-son relationship by analysing 301 letters that a mother had written to her son. The data can also be works of the person being studied (e.g., the paintings of Picasso), which could be analysed for instance for the presence of certain themes such as violence, joy, despair, and the like, or for other kinds of content such as original scientific ideas or prophetic statements (e.g., in the writings and technical drawings left by Leonardo da Vinci).

An example of a qualitative case study involving an analysis of existing works can be seen in an analysis at the University of Hamburg of the poems of the German poet, Friedrich Leopold von Hardenburg, known as “Novalis,” whose work greatly influenced 19th Century Romanticism. He wrote many short poems whose contents were analysed using content analysis in order to identify themes related to death. Because the dates when the poems were written is known, it was possible to relate the appearance of the death theme in the poems to real

events in Novalis's life, such as the early death of his fiancée, Sophie von Kühn, and to show that the kind of poetry he wrote was closely connected with such events. Thus, his creativity was shaped by things that happened in his immediate, everyday life.

Multiple case studies: Gunersel (2009) described a case study project which combines several aspects of the qualitative case study approach just described and also introduces an extension of this method – the *multiple case study* approach, in which a single issue is studied by combining information from more than one case.

Four prominent Turkish authors – one of whom has already been nominated for the Nobel Prize on several occasions – were studied in order to gain insights into the most important social factors that influence the development of people as authors. The particular participants were selected because they all possessed intensive “intrinsic” knowledge (see above) of the object of study, i.e., they possessed “narrative competence,” not on the basis of some statistical procedure such as random sampling. Semi-structured interviews were conducted based on an extremely compact list containing only three fairly general questions, one of which was exemplary for open interviews: “What influenced the development of your creativity?”³⁰ The interviewer also had “prompts” handy, in case a respondent began to answer in too general a way (e.g., “What do you understand by “creativity”?). These interviews, which lasted 1-2 hours (i.e., not a great deal of time – see discussion of the “quantum of effort”), were recorded by tape recorder and transcribed later. In

³⁰ This question is exemplary because it unequivocally directed the attention of the respondents to the object of study but remained open and left the construction and organisation of answers to each particular respondent.

addition, further data from other existing sources were also collected: among other things audio - or video interviews, biographies, book reviews, newspaper articles or material from the internet. In other words, the rather brief interviews were fleshed out with additional data from existing sources, and in this way the quantum of effort fulfilled.

A detailed discussion of Gunersel's (2009) analysis of her data and the results would be premature at this point. However, it can be mentioned that she carried out the analysis in two phases. In the first phase each case was analysed separately. "Content units" (see pp. 199) were identified and clarified using the method of constant comparisons (p. 222), content units with similar content were then combined to form "categories" (p. 220), and finally "concepts" were built up (p. 222). In the second phase a cross-case analysis of the results of the first phase was carried out in order to draw more general conclusions. Three broad kinds of influence were identified: environmental catalysts, support networks, and self-efficacy, with specific catalysts and networks emerging for different authors.³¹

Weaknesses of the case study: The case study approach is beset by a number of problems. The data may be highly idiosyncratic, i.e., specific or peculiar to a single respondent or even to the interaction between a particular respondent and a particular researcher. A case study of the development of Picasso's creativity would obviously require only a single case (Picasso himself) and would obviously be extremely interesting and informative.

³¹ This study is also an example of mixed-model research (see Chapter 6). Gunersel's descriptions of her data collection and analysis provide examples of how to deal with a number of practical issues which are discussed in a clear and concrete way and provide almost a case study of several issues in qualitative methodology, especially as it is conceptualized in this book. Readers are advised to read the article – which is readily available at [http://dtserv2.compsy.uni-jena.de/ss2012/ndlger/73307830/content.nsf/Pages/OAA95935E87BoE9AC1257A320041EDEF/\\$FILE/aca-3-4-222.pdf](http://dtserv2.compsy.uni-jena.de/ss2012/ndlger/73307830/content.nsf/Pages/OAA95935E87BoE9AC1257A320041EDEF/$FILE/aca-3-4-222.pdf) – after they have finished this book.

However, such a study demonstrates two problems: (a) The extent to which creative processes in Picasso were peculiar to him alone (*the problem of idiosyncrasy*); (b) The extent to which their nature and development in him was typical of creative people (*the problem of generalizability*). Kamyliis and Valtanen (2010) also made the point that acknowledged towering personages sometimes simply parrot public stereotypes of themselves that enhance their special status, such as the idea that they are a kind of conduit for messages from higher powers, that their ideas come out of the blue, that they have sacrificed all for their work, or that they are above conventional standards and norms. Silvia, Kaufman, Reiter-Palmon and Wigert (2011) concluded, however, that this does not involve deliberate deception, but reflects the fact that public figures can easily become simply pretentious.

The reliability and validity of case studies: These issues raise questions of bias and rigour that ultimately cast doubt on the reliability and validity of case studies. As a result, as Hamel, Dufour and Fortin (1993, p. 23) pointed out, the case study came to be seen as involving a “prescientific heuristic,” most useful for exploratory studies. For example, the fictitious case study of Picasso just mentioned could easily function as a rich source of hypotheses about creativity and the person. It is well known, for instance, that women experienced severe difficulties in intense relationships with him. He seems to have been uncaring, selfish, cruel and exploitative of them. This raises questions such as whether these are typical characteristics of men hailed as creative geniuses. More generally, without prejudging the nature of the relationships, it could be asked whether there is a systematic connection between interpersonal relations and creativity.

According to Kromrey (2009, p. 507) case studies are mainly useful as exploratory studies. They are valuable in order to (a) “help develop grounded theoretical concepts, theories and hypotheses,” (b) to “add to the plausibility of such theories or hypotheses,” or (c) to “provide concrete examples of results of

quantitative research.” However, case studies offer possibilities that go beyond their “prescientific” usefulness as a “breeding ground for hypotheses” (Shaughnessy, Zechmeister & Zechmeister, 2014, p. 311). For instance, a specific concrete case can (d) demonstrate a rare phenomenon or a unique combination of circumstances that are so uncommon that they can only be investigated by looking at the specific case or cases: A once in a hundred years case like Picasso is an example, as are the rare cases of “feral children” such as Victor the Wild Boy of Aveyron, or Amela and Shamela, who are thought to have grown up in India without contact with other human beings. (e) cast light on the distribution of a phenomenon under everyday conditions or (f) provide examples of how the phenomenon departs from the normal under extraordinary conditions. A series of linked case studies conducted over time with either the same person or several people can also (g) show how a phenomenon develops with the passage of time, in something the same way as was done in the Novalis study mentioned above.

Interviews

Although a number of other ways of collecting qualitative data have just been outlined, Briggs (1986) estimated that 90% of all social science investigations make use of interviews. The proportion may be lower in psychology and education, but Briggs’s estimate leaves no doubt about the importance of interviews as a research tool. The interview can be regarded as “a device for inciting narrative production” (Holstein & Gubrium, 2015, p. 25). What is required in a narrative is what Geertz (2000, p. 3) called a “thick description” of the interview partner’s construction of reality. Typically, this is not a continuous, unbroken monologue but a series of responses to “prompts” from the researcher such as, “Tell me more.” As will be shown in more detail below, there are, in fact, different kinds

of interview. The earlier discussion of problems of the social construction of knowledge (pp. 95) applies particularly in interview situations. One of the responses to this problem is to develop principles of “good practice” in interviewing – rules on how to do it in ways that allow the position of the respondent to be most truly represented in the material that results. These may be divided into rules of good practice representing the “traditional” view and those deriving from a more “radical” view.

As summarised by Douglas (1985), the traditional view is to regard the person being interviewed (the respondent) as a “repository of information” or a “basket full of answers.” The interviewer’s task is then seen as obtaining unrestricted access to what the respondent knows or thinks, in a way that produces accurate information. There are two fundamental tasks for the interviewer: (a) to create an appropriate atmosphere, (b) to formulate questions in such a way that respondents can give answers that do not distort their views. The former means being friendly but maintaining a certain personal distance, making respondents feel that they are being taken seriously, avoiding being judgmental with them about their views (i.e., not obviously agreeing or disagreeing, showing signs of approval or disapproval, or indicating belief or disbelief), and avoiding challenging or “grilling” respondents. The interviewer’s behaviour and body language should also be respectful and alert (i.e., displaying interest) but non-directive, in order to avoid phenomena such as the Greenspoon effect described on pp. 98.

The interviewer is expected to lead the interview, although to different degrees depending upon how structured it is, and to keep it from wandering away from the topic under discussion, but not to assume too active a role, for instance by behaving as though the interview were a casual conversation between two acquaintances. These considerations mean that interviewers should:

1. use appropriate language (e.g., they should not use specialized psychological jargon with laypersons);
2. avoid leading or tendentious questions that put words into respondents' mouths;
3. avoid going into details of the research;
4. avoid entering into private conversation with respondents during the interview;
5. ask occasional "check" questions to be certain that an answer has been understood and that the respondent is being consistent in answering;
6. check their own understanding of respondents' meaning by asking appropriate follow-up questions.

These recommendations may seem not to fit well with the traditional image of the researcher as a neutral observer who hides his or her own personality in the interest of scientific objectivity. However, interviewers should not under any circumstances try to behave like a machine or turn themselves into a walking test instrument. In qualitative research the purpose of interviews is to encourage interview partners to reveal their construction of reality (their internal picture of the world) as openly and fully as possible. To achieve this, interviewers should make use of their full armoury of skills and abilities to lead interviews and persuade people to reveal their thoughts. This means that tact, empathy, insight and the like should be employed. However, interviewers should remember the purpose of research interviews: to obtain information about the interview partner's construction of reality *in order to learn about the object of study*, not in order to gain insight into the problems of the interview partner or to establish a therapeutic relationship with him or her.

More radically qualitative writers give more emphasis to the interview as an exchange between two people with a mutual interest. In the case of atmosphere, it is important that respondents

feel that they are not only being taken seriously but are being approached as partners. This can be facilitated by a certain degree of *self-disclosure* on the part of interviewers: What is meant here is that they reveal something of their own opinions (without forcing this on the respondent), on the one hand, and that they show something of their own feelings (without appearing excessively nervous, showing impatience or seeming to laugh at the respondent), on the other. Interviewers can also show awareness of respondents' world view and point of view, as well as of their feelings and concerns, not only responding appropriately to these but also allowing them to be thematized in the interview. It can be seen that this way of conceptualizing interviews does not support the ideal of the detached, uninvolved interviewer who stands above both the respondent and the material being discussed.

Non-verbal factors: Fontana and Frey (1998) emphasized the importance of nonverbal factors in interviews such as *dress*. An amusing example of problems that can arise from this, involves an attempt to engage habitués of a seedy bar in Hamburg in conversation about their lives as casual labourers in the Hamburg harbour. The interviewer made the mistake of wearing a suit, and was immediately assumed to be a plain clothes police officer and advised in a very firm manner to go somewhere else! This seems to imply that interviewers should wear casual clothing. However, a mirror-image experience contradicts this view. An adolescent boy in an institution for young offenders spoke very disparagingly about a social worker, whom the boy refused to take seriously. This was surprising, as the young male social worker was regarded as a very open, friendly and engaging person who wore very "cool" clothing such as stone-washed jeans and sandals. When conversation with an older man was turned to the subject of this social worker the boy commented: "Oh you mean the hippy. I'm not interested in talking with a hippy!" It turned out that the boy expected his social worker to be a grave and formal older person in a suit (like me!). In this case, a suit would have helped. Thus, there is no simple rule on

clothing.

Fontana and Frey (1998) also mentioned body *posture*, to which *facial expression* can be added. These can create the impression of boredom, nervousness and the like. Fontana and Frey drew attention to possibly unintentional *use of personal space* to communicate attitudes that may or may not be desired (e.g., sitting very close to the respondent, which might imply an assumption of intimacy), and body movements or postures such as leaning towards or away from the respondent. Further important aspects of communication in an interview are related to *language* although not part of the actual words used (Fontana and Frey). These include (a) pacing of speech and use of pauses (for instance, long pauses could suggest lack of confidence in the respondent's ability to comprehend) and (b) variations in speed, volume, pitch and quality of language such as shouting, which might imply impatience or surprise. As Fontana and Frey emphasized, it is important to adapt all aspects of communication in interviews to the special characteristics of the people being interviewed. For instance, the desirable level and style of language, degree of informality, kind of clothing, use of personal space and other aspects would differ dramatically with children from what would be appropriate when interviewing recently bereaved widows.

Questions in Interviews: As Jorgensen (1989) pointed out, the first kind of question used in interviews is the simple, informal question typical of normal conversation. Such questions may occur on an impromptu basis, even in a fairly formal interview situation, for instance when the interviewer does not understand a statement or when curiosity is piqued. However, systematic pursuit of a specific subject of interest requires in addition to such impromptu questions deliberate, purposeful questions. It should be noted that the word "question" is used loosely here to mean any request for information made by the interviewer, even if it is not couched strictly in the grammatical form of a question. Spradley (1979) described several kinds of systematic interview questions,

including:

1. questions that ask for an overview of some object of study such as “Please tell me about your interest in emigrating”;
2. questions asking for more detail on some point already introduced (e.g., “What is the most important thing that led you to name Australia as the country you would most like to emigrate to?”);
3. questions asking for an example of something introduced by the respondent to make it easier for the interviewer to understand (e.g., “Could you give me an example of intolerance of initiative in Germany?”);
4. questions asking for a concrete example from the respondent’s own life (e.g., “When have you had an initiative blocked by intolerance in Germany?”);
5. questions asking for clarification of idiosyncratic language (e.g., “Would you give me an example of what you mean by “the tall poppy syndrome” in Australia?”).

It is generally agreed that questions commencing with either “Why?” or “What do you mean?” should be avoided as much as possible, because they imply disagreement or impatience on the part of the interviewer, whereas “What?”, “When?”, “Where?” and “How?” encourage respondents to express their views.

It is well known that the way in which a question is stated can exercise considerable influence on respondents’ answers. Phenomena such as *yea saying* and *social desirability* (see earlier discussions on p. 98) mean that if a question seems to imply that there is a “correct” answer, many respondents will give this answer, regardless of what they really think. A simple example would be a highly suggestive question like: “Surely you wouldn’t drive without doing up your seatbelt?”

A real example can be seen in an opinion survey on beer preferences in which respondents were asked: “Do you prefer

a light, sparkling lager or a dark, heavy ale?" Most people said that they preferred the "light, sparkling" beer, despite the fact that at that time in the location in question far more "dark, heavy" ale was sold than lager. It is obvious that "light and sparkling" is better than "dark and heavy," and it would require a certain amount of courage to admit to a preference for the latter in an interview, regardless of what kind of beer the person in question actually preferred. Questions that through the choice of words suggest to the respondent (rightly or wrongly) that a certain answer is expected, are called "leading" questions. It is important to note that respondents may be "led" not only by the wording of questions but also by nonverbal communications such as nodding or smiling versus frowning, leaning forward or away, or making a note of some answers but not others.

At first glance it seems that interviewers should avoid leading questions, and an obvious reaction to the problem is to try to eliminate it by careful wording of questions and control of the interviewer's body language. However, leading questions have their place. As Kvale (1996) pointed out, they can be used to *check the authenticity of answers*. For instance, if a respondent makes a particular statement but then contradicts this view in response to a leading question, this casts doubt on the view originally expressed. On the other hand, if the respondent resists being led, this adds greatly to the credibility of the answer. Drawing attention to self-contradictory statements can also help to *gain access to hidden ideas*, especially in in-depth interviews. Kvale did not call for elimination of leading questions. On the contrary, he argued that they are not used enough. However, he did emphasize that researchers should be conscious of their own use of such questions and draw attention to them in their research report, so that readers can take account of the questions in interpreting the findings. As

was said earlier, it is not a matter of methodological purity but of transparency. The importance of clarifying in a scholarly report what questions or prompts were used is made clear by Kvale's assertion that in a certain sense *all* questions asked in interviews are leading questions, as well as by his point that in analysing a text the researcher is in effect asking leading questions of the narrative!

Kinds of Interview: Because of the importance of the interview in qualitative research in psychology and education, this form of data collection will be dealt with in greater detail than the other approaches outlined above. For our purposes, five kinds of interview can be discerned, and these are discussed below.

- *Unstructured interviews:* Such interviews offer respondents only a general field or issue about which they are asked to talk (e.g., "I am interested in the use of seat belts. Would you tell me what you think about this, please"). The interview may seem almost like a conversation, except that the interviewer asks numbers of questions that give the interaction a certain degree of structure (Knapper and Cropley refer to such questions as "prompts" – see below). Although always referring in some way to the object of study, these questions may differ substantially from interview to interview, according to the responses of the particular participant. The order in which they are asked may also vary from respondent to respondent, or of course a question may not be asked at all if the respondent has already dealt with the relevant material without any prompting. The crucial element of open/unstructured interviews is that the researcher attempts to establish how respondents structure the area in question in their own thinking about it, and what categories they themselves choose when trying to communicate their views to another person. In the seat belt study one respondent spontaneously began to talk about why some people refuse to wear a seat belt, another about parents who risk their children's lives by not buckling them up, a third about the benefits of seat belts in avoiding injury, and so on, thus indicating differing priorities in the minds of different participants. The interviewer may offer encouragement

such as “Go on, please,” or even the well-known “Uh, huh,” but the respondent is not directed in a narrow way to specific issues that are of interest to the researcher, since the researcher’s goal is to identify the issues the respondents themselves introduce.

The data yielded by unstructured interviews are the statements made by the respondents. These must eventually be converted to the form of a written record. However, this is often difficult in unstructured interviews because the statements are frequently long and complex. Writing them down verbatim as participants utter them is often more or less impossible, for instance because it would mean interrupting respondents part way through an answer while the interviewer hurriedly wrote it down, not only breaking the respondent’s train of thought but also causing the interview to last several hours. Thus, if possible such interviews should be tape-recorded or videoed. Written notes can be made during the interview or afterwards, for instance in the form of a memory protocol, to remind researchers of special observations or ideas that occurred to them in the actual interview. Ultimately, a written record must be developed for later analysis (see Chapter 5), and this may require a great deal of time and effort.

• *Semi-structured interviews*: As the name implies, these are more structured. The researcher will already have some knowledge about the key issues and some expectations of what might emerge, or even wish to focus the interview to some extent. As with unstructured interviews the discussion is initiated with a fairly broad opening question or request that none the less guides the discussion in a particular direction: “I am interested in why so many people do not wear seat belts. What is your view on this?” A small number of general questions or at least *direction-giving cues* or “prompts” (Knapper & Cropley, 1980) will be used to see to it that respondents stick, at least in a broad way, to issues that interest the researcher (e.g., “Is it anything to do with personality, do you think?”). These can be written down for the interviewer on a protocol or recording sheet. However, the interviewer does not have

a list of specific concrete questions that are put to all respondents in a set order. Semi-structured interviews offer better prospects of recording data in writing during the actual interview. For some answers it may not be necessary to note the complete statement but only a relevant word or phrase, because the researcher already knows more or less what is interesting for the purposes of the particular study. Brief notes of this kind can be expanded afterwards via a memory protocol. However, electronic recording is still extremely valuable, as otherwise much of the spontaneity and richness of respondents' answers may be lost.

- *Structured interviews* involve a highly structured question-and-answer exchange in which the researcher enters the research with specific questions, to which concrete answers are desired. These questions are usually derived from existing theory in the area, from earlier research findings, from a pilot study, or from hunches or intuitions of the researcher, although they may also simply be areas of special interest to the researcher. The questions are frequently read aloud from a protocol or recording sheet that the researcher works from, and are asked of all respondents using the same wording and in a set order, unless a respondent offer relevant answers spontaneously (i.e., before the researcher gets to the question at issue). Sometimes questions are open ended (i.e., the respondents answer in their own words), but in some structured interviews the interviewer may even offer specific alternatives from which the respondent can choose. Frequently, the respondents' statements need not be written down verbatim by the interviewer, since because of the structured nature of the questions only certain "key" words or phrases are interesting to the researcher and only these need be noted. The interview recording sheet may include spaces for noting the key contents of the respondents' answers, some of which may be no more than an indication of agreement or disagreement (e.g., "Do you agree that requiring people to wear a seat belt is an infringement of their basic freedom?").

Such interviews offer certain advantages: They yield pre-

structured information, are less time-consuming, and are easier to analyse. Indeed, it is possible to conduct highly structured interviews without direct contact between researcher and respondent, for instance by telephone or even by mail, and the approach is often used for surveys in areas such as political polling or market research. However, there is a danger that they may become little more than spoken questionnaires. According to Jorgensen (1989), the data are less rich, more superficial, and more difficult to interpret. They are also prone to the effects of factors like yea-saying (agreeing with whatever the interviewer asks) or social desirability (giving answers that seem to the respondent to be what the right sort of person would say).

- *A phase approach to structure in interviews:* Kremer (2010, p. 12) offered a helpful overview of the kinds of interview just discussed: In effect he identified four levels of structure, which I conceptualize as a *hierarchy* that can be organized in terms of the degree of specificity of prompts. The four levels are outlined in the box that follows. The hierarchical nature of the differing levels of structure clarifies an important point: The levels are not necessarily mutually exclusive but form a dynamic hierarchy – for example, a research study could start with an unstructured interaction between a researcher and interview partners which identified a key aspect of the respondents’ understanding of the broad area the researcher is interested in (in the box below the broad area is “major life incidents”). After this unstructured interaction with a number of interview partners (for example, people with a migration background), the researcher could become aware of a more specific issue (e.g., in the box, “separation from the homeland”). The investigation could then move into a minimally-structured phase in which the prompts offered prompts referring directly to this separation. Statements made by respondents might then reveal the importance of “feelings” in regard to the separation, which would in turn lead to a semi-structured phase focusing on feelings and finally to a highly-structured phase investigating the particular feelings

respondents experienced (such as sadness). This phase approach emphasizes the emergent nature of much qualitative research, while it also provides an example of how an initially general understanding of an object of interest may close in on a specific understanding: At some point in this dynamic focusing of the investigation a switch may occur from qualitative to quantitative methods: For example, the highly-structured phase may be based not on interviews (qualitative method) but on a questionnaire (quantitative method). Combination of qualitative and quantitative methods is discussed in detail in the next chapter.

- Level 1 (unstructured): The prompts can be almost completely neutral (e.g. “Please describe a major incident in your life”); the interview partner is free to choose any aspect of his or her life, although further more specific prompts (see following sentences) are likely to follow at some point.
- Level 2 (minimally structured): The key situation can be specified (e.g. “What was it like to leave your homeland?”); the aspect of the interview partner’s life is specified (leaving the homeland) but details are left unspecified.
- Level 3 (semi-structured): Not only the key situation but also a specific aspect of it can be specified (e.g., “How did you *feel* as you left your homeland?”); the interview partner’s attention is directed to feelings, but without specifying what feelings
- Level 4 (highly structured): The key situation and specific aspect (in this case feelings) and also a specific element of this aspect are specified (e.g., “Did you experience a feeling of sadness as you left your homeland?”); the general area (the interview partner’s life), the specific life experience (leaving the homeland), the general area of experience (feelings), and the specific aspect of feelings (sadness) are all specified.

• *In-depth interviews*: Such interviews are characterized by the interviewer’s attempt to get below the surface and identify the

underlying emotions and feelings experienced by respondents, or recognize the rationalizations and defences they use when dealing with the issues being discussed. In-depth interviews may thus become almost clinical, even making use of clinical procedures such as free association, encounter techniques, and the like. For instance, an in-depth discussion of why a particular person refuses to use seatbelts might reveal that the whole topic is surrounded by hostility and anger ultimately directed at an over-controlling and authoritarian parent. In an in-depth interview a respondent might reveal that refusing to buckle up is a symbolic act of defiance and a confirmation of the cutting of the psychological umbilical cord. Material of this kind may be disturbing for the respondent, especially one who is sensitive and insightful, and needs to be treated with care. This example confirms the importance in qualitative research of personal properties in the researcher such as empathy, but also responsibility.

- *Group interviews:* As the name implies, these are conducted with more than one respondent at a time, i.e., in a group setting. In theory, all the variants listed above (structured, semi-structured, unstructured, in-depth) are also possible in groups. In the study of German migrants, for instance, a number of group interviews were conducted with people who had applied for an Australian immigrant visa. It was found that rather different themes were emphasized in the group setting than in individual interviews. Once they realized that they were alone with other visa applicants the respondents became much more open in their answers. In other words, mechanisms such as yea-saying, social desirability and avoidance of cognitive dissonance became far less obvious. Fontana and Frey (1998) developed an interesting overview of group interview approaches ranging from the “focus group” – where specific questions based on highly developed expectations based, for instance, on earlier research are put to a group – through “brainstorming” and “Delphi” groups to groups interviewed in the field. Such interviews range from unstructured to structured, and

may be used for exploratory purposes (hypothesis finding), for pre-tests and for phenomenological studies.

Although they were writing about focus groups, the principles for conducting group discussions outlined by Stewart and Shamdasani (2015, p. 75) are generally instructive: group interviews are conversations, but conversations among groups of people purposely chosen because they possess narrative competence regarding a specified research topic. The exchange of views in the group interview is conducted by a moderator who is interested in this topic and who takes care that the conversations do not become too lengthy, e.g. by means of an interview guide. Because they take place in the language of the individuals involved (not in the language of tests, for example), such interviews are very productive for the analysis of perceptions, attitudes, motives, and the like. Group interviews offer additional technical challenges for interviewers: Not only do the usual rules of good practice apply (openness, friendliness, sympathy, ability to encourage reflection and communication, willingness and ability to reveal something of oneself and thus engage in a genuine interaction), but in addition new skills are needed. The interviewer has the further task of moderating a group, for instance discouraging a situation where one or two members dominate, preventing mobbing, or persuading retiring members of the group to participate. As experience in the migrant and social worker studies already mentioned shows, group interviews stimulate respondents and can greatly facilitate openness. They also yield a large amount of information relative to the time they take.

Chapter 6: Mixed approaches to collecting data

Mixed approaches combine qualitative and quantitative designs to explore a single research topic. The justification for this combined procedure is that it provides more authentic findings (i.e., mixed approaches enhance external validity). Such approaches are not based on a fixed paradigm: they include a number of variants and hybrid forms, which can be summarized as involving either a *mixture of methods* (combination of separate qualitative and quantitative datasets, each of which could stand alone) or a *mixture of models* (combination of qualitative and quantitative data-collection methods to obtain a single dataset, which could not exist without the combination). There is a broad and multi-faceted – also in part highly critical – literature on the epistemological (i.e. what constitutes knowledge) and heuristic (i.e. how can

knowledge be obtained?) foundations of mixed approaches, but in this chapter I focus on the issue of what is most productive for casting light on a given problem, especially in practical terms. Thus, my purpose is to explain how approaches can be mixed in ways that (a) improve understanding of the object of study, and (b) facilitate application of findings in real-world settings.

Chapter Glossary

<i>explanatory mixture:</i>	a secondary data collection takes place <i>after</i> the main survey
<i>exploratory mixture:</i>	a preparatory data collection takes place <i>before</i> the main survey
<i>multi-method studies:</i>	a mixture that leads to two data sets, which could also exist separately
<i>mixed models:</i>	a mixture that leads to a single dataset in which each datapoint results from the combination of qualitative and quantitative methods
<i>nested mixed design:</i>	a secondary data collection takes place <i>during</i> the main data collection (see <i>sequential</i> designs below)
<i>pragmatism in research:</i>	willingness to use both qualitative and quantitative methods, according to which is more profitable for a specific subject of investigation
<i>reconstruction logic:</i>	the idea that research should focus on understanding how people make sense of the world

<i>sequential mixed designs:</i>	the two data collections (qualitative and quantitative) occur one after the other
<i>triangulation:</i>	both data collections have equal status. They take place separately, and are only combined at the level of data interpretation

The Core of the Mixed Approach

In the heuristic sense (e.g. Tashakkori & Teddlie, 2010), mixed approaches involve the simultaneous adoption by researchers of the perspectives of both sides in the paradigm wars (see the discussion of these wars in Chapter 1). Johnson and Onwuegbuzie (2004, p. 17) defined mixed research in an all-encompassing way as “the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts, or languages into a single study.” The use of mixed approaches is based on a “reconstruction logic” (Herborth 2011, pp. 138-42), according to which methodologies are tools for reconstructing the way participants in a study (i.e. the people whose views are being sought by the researcher) understand a particular issue (object of study) about which they are well informed. According to this view, research methodology is not a corset or straitjacket, but should be adapted as required for the purpose of casting light on a specific topic (e.g., Herborth, 2011, p.146): It should focus on the solution of a problem and not on the technically perfect application of a particular paradigm (Katzenstein and Sil, 2008, p. 110). Achieving an accurate understanding of how participants in a study understand the world in which they live involves enhancing the study’s *external* validity.

Although the box that follows contains an exaggeration

and oversimplification of complex issues, the difference in the underlying philosophies of qualitative and quantitative methods can be summarized in a starkly contrasting way as follows:

- The purpose of the quantitative approach is to apply correct methods for acquiring knowledge about an object of study, *regardless of whether it yields any new knowledge or not!*
- The purpose of the qualitative approach is to enlarge understanding of the object of study, *regardless of how you achieve this understanding* (as long as you report fully and accurately how you did it, so that your research is open and readers can make a judgment on internal validity).
- The purpose of mixed methods is to combine qualitative and quantitative approaches to produce the best enlargement of knowledge.

According to Denzin (2010, p. 419), such “mixed, multiple and emergent methods are everywhere today.” Indeed, Hussy, Schreier and Echterhoff (2013, p. 286) remarked that they have become “a kind of fad.” There is even a new journal (*Journal of Mixed Methods Research*) that has a solid impact factor. Kyne (2021) added an interesting dimension to the discussion of the increasing practical importance of mixed research by emphasizing its growing use as a decision-making tool in the commercial world of product-led companies. As he put it in a down-to-earth way, “product leaders love mixed method researchers,” whom he described as “full stack” researchers. He attributed the commercial importance of mixed research to the fact that, whereas quantitative data answer questions about the “what?” of behaviour, qualitative data provide information on the “why?” He gave user interviews as an example of a procedure focused on the “Why?” The discussion below on explanatory and exploratory mixed approaches helps to make sense of the what?-versus-why? distinction.

From what might be called the “strict” quantitative perspective the overriding criterion of rigour is that the methodology

is *internally* valid (e.g. p. 60). This perspective requires:

- application of a fixed paradigm based on deductive thinking (e.g., starting with a clear expectation of what is likely to happen – a hypothesis)
- formal sampling techniques (such as random, systematic, stratified, or cluster sampling)
- a numerical data-collection toolkit (e.g., tests, questionnaires, instruments)
- routinized data analysis (e.g., correlation-based procedures, analysis of variance)
- statistical criteria of the strength of effects (e.g., significance tests, tests of power or effect strength, confidence intervals).

As McManus (2016, p. 157) pointed out, close adherence to this perspective means that only those phenomena that are accessible to the one true research approach are investigated: The result is that quantitative methods really do become the only course of action that is worth adopting, since anything else is declared *a priori* to be worthless. Modern ideas of science and especially scientific publishing – as encapsulated in the “open science” movement (e.g., Open Science Collaboration, 2015) – are increasingly regarding qualitative research reports as unpublishable, because data are not collected specifically to test a pre-formulated and pre-registered hypothesis, there is no control group and no random, double-blind assignment of participants to experimental and control groups, no objective scoring, and so on.

Putting findings to work: The perspective just stated can be contrasted with approaches which emphasize that the measure of “correctness” of a procedure is its ability to provide understandable, plausible, usable answers to questions about real-life behaviour, even if these findings were not anticipated by the researcher (i.e., there was no formal hypothesis). It emphasizes the importance of a second step in research: *putting findings to work* by drawing practical conclusions. Morgan (1998) pointed out that a mixed

approach can be particularly helpful in doing this. For example, quantitative data could be used in health research to test whether a particular treatment had resulted in statistically significant changes with respect to objective variables that can be measured in numbers, such as blood pressure, blood sugar, or kidney function. Qualitative data could then *explain* how the people involved subjectively experienced their improved state of health, i.e. such data could clarify the *subjective* meaning of *objective* quantitative relationships: Did the people feel differently about their lives, and in what way? Short descriptive anecdotes or direct quotations from interviews, narratives or letters, among other things, could be used to understand what the statistically-ascertained improved state of health actually “felt like” (phenomenologically). The following box gives an example of a study in which the mixed evaluation of a “treatment” both quantitatively (by means of test scores) and qualitatively (by means of interview comments) went even further, and prevented a finding not detected by quantitative methods going unnoticed.³²

Schmid (2016) conducted creativity training with a group of visitors to a day-care centre for the elderly. Participants were very consistent in their attendance and praised the positive impact of the program so much that groups from other centres approached her and asked if the program could be extended to their centres too. After several weeks, participants’ scores on a creativity test and on a symptom checker questionnaire were compared with pre-training scores in the hope of revealing significant health-related changes, but the comparison of mean scores on the tests did not reveal any statistically significant differences. In a strict quantitative sense, the treatment had

³² This study involves what I call an “explanatory” design (see p. 162). It is also an example of what I have called “multi-*method*” research. Finally, it is also an example of triangulation, as both datasets are capable of standing alone, although individually they lead to contradictory conclusions.

had no effect. However: In interviews the participants drew attention to clearly noticeable positive effects related to their quality of life. Among other things, they were able to see the funny side of old age, could identify their own strengths as well as those of their partners, and felt good about their own contribution to domestic tasks. Thus, the explanatory, pragmatic mixture of approaches showed that there were effects, but they were qualitative rather than quantitative.

“Exploratory” versus “explanatory” approaches:

The core of the mixed approach is that a combination of qualitative and quantitative designs can lead to findings that go beyond what is possible with a single approach, and thus promote *external* validity. As a tool for collecting data in a way that enhances their external validity this combination of approaches can take several different forms. Examples of these include:

- A qualitative precursor study may reveal naive, diffusely-defined concepts of interview partners that can be followed up quantitatively.
- A qualitative precursor study can provide formal hypotheses that can then be formally examined in a quantitative phase.
- A qualitative intermediate step can provide data that are then converted into quantitative data (qualitative -> quantitative).
- A qualitative follow-up can clarify how states revealed in a quantitative study are experienced in the real world (quantitative -> qualitative); see the example in the box below.

Kyne distinguished between “exploratory” mixed research, where supplementary quantitative data clarify initial qualitative findings, and “explanatory” mixed research, where the reverse occurs: supplementary qualitative data cast light on initial

quantitative findings. The box on pp. 161-162 involves a study in which the initial quantitative data were unable to detect changes in behaviour, whereas the changes were readily explained by subsequent qualitative data (i.e., an explanatory mixture).³³ The case study of young women who had lost contact with their fathers, also presented below (see p. 168-169), is an example of qualitative data leading into a quantitative study (i.e., an exploratory approach).

Simplifying rules of thumb:

- Exploratory mixed design = follow-up quantitative data explore initial qualitative findings.
- Explanatory mixed design = follow-up qualitative data explore initial quantitative findings.

Multi-Method vs. Mixed-Model Designs: The investigation in the box above involved a mixed situation in which all participants experienced the same “treatment” and both the quantitative and the qualitative data were collected by the same researcher with the same participants, within the same organizational framework and in relation to the same “treatment” (participation in the creativity training). That may all sound understandable in principle, but in scholarly circles there is a controversial, sometimes even embittered discussion, involving both aspects of epistemology and the idea of what is “scientific,” as well as questions of what constitutes scientific evidence. In addition, topics such as the effects on data collection and analysis of researchers’ worldviews, the power relationship between researchers and respondents, and the treatment of knowledge as the private property of researchers are often discussed controversially (e.g. Denzin, 2010; Johnson & Onwuegbuzie, 2004). As Denzin makes clear, in the last 50 years there have actually been not one but three paradigm wars (see p. 23). These have included academic struggles among critical theorists,

³³ The study described in the box on the next page on the effects of creativity training on elderly participants is an example of an explanatory mixture.

anti-naturalists, feminists, post-positivists and post-structuralists, as well as others, all of whom believe they have discovered the only true paradigm (see Tashakkori & Teddlie, 2010). The disagreements have been discussed on “axiological, ontological, epistemic, epistemological,³⁴ and methodological levels” (Denzin, 2010, p. 421). Without questioning their importance and with full respect for them, in this book I do not address such issues at all but focus on the way in which research practice can combine quantitative and qualitative approaches in the same study to make research findings more useful.

Tashakkori and Teddlie (2010) concluded that the paradigm wars are over and that the time had come to become “pragmatic” – in other words, to use both qualitative and quantitative approaches, whichever of the two is more profitable for a specific subject of investigation or for a specific phase of a complex investigation. They went even further, proposing a *combination* of approaches to improve both the effectiveness and efficiency³⁵ of research. In principle, according to Tashakkori and Teddlie, there are two good options for an effective combination of approaches: mixed *models* on the one hand and mixed *methods* on the other. However, Hussy, Schreier and Echterhoff (2013, p. 292) concluded that the distinction between mixed models and mixed methods is no longer up-to-date due to the diversity of mixed approaches that now exists. Consequently, they use the term “mixed methods” as a generic term that includes both “multi-method” and “mixed-model” designs. For the sake of simplicity, however, I still distinguish between two idealized types of mixed approach, which in the following are presented rather plainly and baldly:

a) *Mixed models* lead to a single dataset where each datapoint

34 For example, can you really trust the findings of mixed research?

35 “Effectiveness” refers to the ability of a method to detect *all* genuine effects. However, this may be done at the cost of mistaking some spurious effects for genuine ones. “Efficiency” refers to the ability of a method to identify *only* genuine effects (i.e., to avoid mistaking spurious effects for genuine ones), but involves the risk of missing some genuine effects.

results from the combination of qualitative and quantitative survey methods. Without the combination, the final dataset would not exist. Often the combination is process-like: for example, a qualitative study leads to the creation of an inductive category system, which then serves as the basis for a deductive quantitative survey. Both approaches, however, are indispensable for the survey.

b) *Multi-method* studies involve the combination of at least two datasets which can be clearly distinguished from each other (in the simplest case one qualitative and one quantitative), which could also exist separately and could even be, and occasionally have been, collected without any deliberate coordination or combination. The study of the effects of creativity training presented in the box above is an example of a multi-method study.

Simplifying rules of thumb:

Mixed models = a single, combined data set

Multi-methods = two (or more) data sets, which could also exist separately

Mixed Models

Ideally, mixed models combine qualitative and quantitative survey techniques to collect and/or analyse a single compound dataset. The point is that without the combination of approaches it would be impossible to collect the specific data. For example, qualitative methods can be used to formulate hypotheses inductively, which are then tested deductively using quantitative methods. Strictly speaking, this is always the case. Familiarity with a phenomenon is indispensable for formulating hypotheses that can be tested in a specific environment. In hypothesis-driven quantitative research, the hypothesis is derived from existing findings. In a typical

quantitative research report – such as a master’s thesis – there is usually at least one chapter on the origin and the emergence of the hypothesis/hypotheses. By contrast, not only can qualitative surveys begin without such theory, but because of their inductive nature they offer good opportunities for formulating valid hypotheses after the fact. These emerge from the narratives of the participants and depend upon the participants’ narrative competence. As Corbin and Strauss (2015, p. 7) put it, in qualitative research, theory is “derived from data collected in the course of the research process, not established before the study begins.”³⁶

Derivation and adoption of terms: Mixed models can use qualitative methods to understand how a particular concept is understood in the words of people in everyday life, as against how it is defined by researchers using scientific concepts and terms. For example, an instrument for the quantification of the concept can be developed using terms used by the participants in the qualitative phase. Practical experience confirms that this procedure can yield questionnaire items that participants at a later stage immediately recognize as an authentic reflection of their opinions (e.g., Knapper & Cropley, 1980, 1982). In this case, the development of terms by means of qualitative methods increases the face validity of questionnaires for a quantitative study. The case study example presented in the following box provides an example of such a combination of approaches. It involves a very simple variant of the combination of models. In the first step, the naive concepts by means of which members of the public made sense of the object of research (property insurance) were determined (by asking the participants in interviews in a precursor qualitative study what they thought about insurance). On the basis of this step, concrete contents for a typical questionnaire investigation with a new group of participants were derived from the utterances of the participants in the precursor study, and the questionnaire administered in the

36 Rejected out of hand as HARKing by strictly quantitative researchers.

follow-up quantitative study.³⁷ The second phase of the investigation (an exploratory quantitative study) yielded findings that eventually led to recommendations for practice.

At the time of the investigation (Knapper & Cropley, 1982), many insurance companies were facing the serious problem of rising numbers of minor damage claims. The study examined possible reasons for this. In the first step, semi-structured interviews were conducted with insurance brokers and policy holders. The design of this part of the project, the construction of the initial sample and the first phase of the data collection were qualitative in nature. On the basis of a content analysis of the interviews, several decisive motives were identified: Participants wanted to get back more from the insurance than they had put in (“profit motive”), they wanted to be protected from every conceivable type of risk (“protection motive”), and they wanted to be compensated for the smallest loss (“full-coverage motive”). Using verbatim comments made by the people interviewed, the researchers then designed a traditional questionnaire to measure the strength of expression of these motives. By administering this instrument to a large new sample of policy holders a link was established between the three motives and quantitative factors such as the number of successful recent claims, age, socioeconomic status, educational attainment, or level of familiarity with television advertising related to property insurance. In the end, the researchers came to the conclusion that the insurers’ TV advertising campaigns were simply too successful in activating profit and full-coverage motives. The concrete proposal for a practical “defensive action” was a redesign of the advertising message to move its focus away

³⁷ The questionnaire – a *quantitative* instrument – could not have existed without the interviews (a qualitative data-collection method), so that the quantitative data depended on the qualitative study for their existence.

from the claims mentality – according to which suffering a loss seems to be almost a blessing – and instead emphasizing the protection motive – a loss is a bad thing, but you have support. In terms of methodology, the questionnaire would not have existed without the precursor interview study.

Deriving quantitative data from qualitative:

An even stronger intertwining of quantitative and qualitative models is presented in the next case study. Both approaches were indispensable for this piece of research. The same data were analysed both qualitatively and quantitatively, but both the quantitative data and the quantitative analysis were derived from the qualitative findings. The study provided information regarding both the type (qualitative statements) and the frequency, intensity and severity of the participants' reactions (quantitative statements) from the same data. Consequently, it was possible to address quantitative questions, such as whether extra-punitive reactions were more frequent in the group than intra-punitive ones, or whether those women with the most intense responses had an impoverished affect spectrum, as the Weber-Fechner law would predict.

The psychological reactions of young women whose parents had divorced during their adolescence were investigated. Since the divorce the women had never, or had only rarely, had contact with their fathers. The protocols of interviews with the women were subjected to a content analysis to determine how they experienced the separation. The following types of reaction were observed: 1. feelings of loss, 2. yearning, 3. doubt and helplessness, 4. anger (extra-punitive reaction directed against the parents), 5. guilt feelings (self-directed intra-punitive reaction because they believed they themselves were responsible for the separation of the parents). These findings

identified the *kind* of reaction experienced by the participants in the study (qualitative). Subsequently, the proportion of women who mentioned each feeling was calculated. These data allowed a statement about the *extent* to which the various reactions were present in the whole group (quantitative). The intensity of each response was also calculated for each participant by counting the frequency with which she mentioned the feeling of loss, longing, helplessness, anger or guilt. In addition, the strength of feeling was calculated by assigning numerical values to each relevant statement: For example, “From time to time, I miss my father,” was classified as a weak feeling and received one point; “Often I have the feeling that I cannot live without my father,” on the other hand, was classified as a very strong feeling and received 5 points. Intermediate values received 2 to 4 points. It then became possible to test null hypotheses such as: “There is no difference in guilt feelings between women who were 12 or over when their parents split up and those who were under 12.” As in the insurance study described in the previous box, the qualitative study could not have been carried out without the interview data.

Multi-Methods

At least in theory, multi-methods can mix qualitative and quantitative approaches in all phases of an investigation: research design, setting, data collection method, data type, data evaluation, and generalization strategies. For example, the same set of data could be subjected to two evaluation methods, thus constructing two sets of results that could be compared. Another example would be a study in which data were collected using not just a laboratory

experiment but also a series of interviews with the same participants. The data resulting from the experiment (stereotypical quantitative data, see Table 2.1) could be subjected to a standard statistical analysis. The interview data, on the other hand, could be subjected to a content analysis (a stereotypical qualitative method, see Table 2.1) and the two results compared. Nowadays, four types of mixed methods are frequently distinguished: exploratory, explanatory and nested designs, and triangulation (see Creswell, Klassen, Plano Clark & Clegg Smith, 2011). In this section, I will slightly revise this classification scheme and distinguish between two bipolar dimensions: exploratory versus explanatory designs on the one hand, and nested versus sequential designs on the other. However, it must be kept in mind that these categories are not exclusive. A simple example: an explanatory design (see below) can be nested or sequential. Explanatory sequential designs are often discussed under the heading “triangulation” (Creswell, Klassen, Plano Clark, & Clegg Smith, 2011).

- *Exploratory designs*: such designs involve an initial investigation that functions as a precursor to a partner investigation that is usually regarded as the “real” investigation. The initial, exploratory investigation leads to the discovery of structures or contexts in situations where the researcher cannot or does not want to resort directly to existing knowledge, models and concepts, for example by studying the literature. This happens, for example, because little or nothing is known about the research subject or because the researcher wants to deal with this subject without preconceptions. Initially, such designs are not based on hypotheses – otherwise they would not be exploratory. Most frequently, but not always, the initial investigation takes the form of a qualitative study that provides material that is then used for the purposes of a

subsequent quantitative study. Particularly suitable for exploratory purposes are, among other things, unstructured interviews with individuals or experts – as was done in the insurance case study in the box above – focus-group interviews, expert discussions or unstructured observations, although other data sources (such as diaries, observation procedures, etc.) are also conceivable.

- *Explanatory designs:* as the label suggests, explanatory investigations are usually carried out *after* the main investigation. Their purpose is to improve understanding of the findings of the main investigation. The qualitative assessment of the quality of life of some visitors to the day-care centre (see the box on page 161), which followed the quantitative investigation based on scores on tests and rating scales, made it possible to further differentiate the numerical results of the quantitative analysis by taking into account the subjective experience of the participants once the quasi-experimental main procedure had been completed. Because explanatory designs are mostly done after the main study (it is not possible to explain what has not yet been discovered), explanatory designs are mostly sequential, usually quantitative followed by qualitative.

- *Nested designs:* In nested designs, one type of survey plays the key role; the results of this survey could stand alone and would be enlightening even without data from the other method, although incomplete. The methodological approach (opportunity sampling versus structured sampling, the use of narratives versus testing to collect data, meaning versus statistical, hypothesis-driven

data analysis) is consistent with the norms of the dominant method. The data of the secondary method are collected more or less incidentally during the main study; without the primary data, they would not make sense. Nested designs are often nested in a quantitative design to make the research set-up easier to understand, to develop questionnaires, or to identify participants who possess narrative competence.

- *Sequential designs*: In a typical sequential design, a researcher first collects quantitative, mostly numerical data - although the reverse order is conceivable. These data are analysed and then data of the other kind are collected in a second step. In a process of “mutual elaboration” the new data supplement the results of the first phase. The mutual elaboration involves combination of approaches *at the level of interpretation*, but not at the level of data collection or analysis; otherwise the design would involve nesting. Sequential designs involve a kind of cross-validation: for example, quantitative data could provide a general understanding of the research subject, but subsequent qualitative data could refine these insights by also taking into account the opinions and perspectives of the participants (see for example Tashakkori and Teddlie, 2010).

Simplifying rules of thumb:

- Sequential mixed designs: the two data collections (qualitative and quantitative) occur one after the other.
- Nested mixed designs: a secondary data collection takes place *during* the main survey.

- Exploratory mixed designs: a preparatory data collection takes place *before* the main survey.
- Explanatory mixed designs: a secondary data collection takes place after the main survey. (Both exploratory and explanatory designs are necessarily sequential.)

Triangulation

The term “triangulation” derives from a method for locating, for example, an electronic signal. In my youth triangulation was used, for example, to catch people who were using a television set – in those days a rare novelty – without a TV licence. In order to determine the origin of the signal, it is possible to determine the compass direction from which the signal is coming at two locations. When a straight line in the direction of the signal is marked on a map from each of the two measuring points, the two lines intersect. This locates the source of the signal. Together with the two locations from which the signal direction was detected this location forms the third point of a triangle and so we speak of “triangulation.” In the simplest sense, “triangulation” – when applied to the combination of qualitative and quantitative methods – means the use of two (or more) *separate* research studies (such as a qualitative and a quantitative study), both of which involve the same object of study, in order to see where the two results intersect. The core of triangulation as a research method is that the two pieces of research are independent of each other. As Nastasi, Hitchcock, and Brown (2010, p. 307) put it, such “complementary” designs are used by researchers to work out the overarching meaning of data. Triangulation thus refers to a particular kind of *multi-method* approach.

In what might be called “full blown” triangulation, quantitative and qualitative data are collected and analysed quite

separately and without any cross-methodological influence, the only link being that both data sets relate to the same research topic (the object of study). The respective results can be compared with each other in order to understand the research subject authentically, credibly and usefully (see p. 44ff – “external” validity). In order to obtain the strongest possible triangulation effect, the two surveys should relate to the common research subject but otherwise be as independent of each other as possible and each provide its own dataset, which is analysed according to the standards of the respective approach.³⁸ Any research that is based on two datasets is sometimes regarded as involving “triangulation.” In my view, however, research such as the study of the daughters of broken marriages described on p. 161 does not involve triangulation but is an example of a sequential mixed design: I recommend using the term “triangulation” as a label for the kinds of design described below (see the next two pages).³⁹

The following box shows a study (Kvale, 1995) referring to school grades in a Danish school. Two separate datasets were collected in order to be able to draw better informed conclusions about an aspect of grading in school classrooms. As illustrated in the next section, this example is a very simple variant of triangulation. In fact, my own remarks in the previous paragraph represent an over-simplification of triangulation. More recent discussions of triangulation are more differentiated (see the discussion on the next page).

38 This means, for example, that the case study of research with young women whose parents had divorced when they were children (see above, p. 168) did not involve triangulation, because the quantitative data were simply an elaboration of the original qualitative data, or, to put it slightly differently, the qualitative data were simply a preliminary step leading to the quantitative data.

39 The crucial point, however, is that both approaches have equal status: the quantitative study is carried out quantitatively, the qualitative one qualitatively and neither is the servant of the other. Only in the overarching interpretation are they combined. In terms of the process of drawing conclusions, triangulation means that the results of the two approaches are combined to better “locate” the object of study.

In interviews, many of the children complained that grades were “unfair.” They argued that grades mirror verbal participation in class activities to the same extent or even more strongly than knowledge of the material. This was unfair. Kvale triangulated the object of study by conducting a quantitative study *with a new group of children*, examining the null hypothesis that there was no connection between verbal participation and grades. From a sample of 30 schoolchildren, he collected data on how talkative the children were in class – he saw this as an indicator of the probable extent of their verbal participation in lessons – and correlated these values with the children’s grades. This resulted in a statistically significant correlation coefficient of 0.65, which justified the rejection of the null hypothesis. Thus, the children were correct in the claim they made in the separate qualitative study.

Recent triangulation approaches: According to Hussy, Schreier and Echterhoff (2013) there are five different types of triangulation:

1. *Methodological triangulation:* Separate data sets relating to the same object of study but collected using different designs (for example: qualitative design for one set, quantitative for the other) are collected. The two sets are analysed separately and the results are compared in order to obtain a higher-level interpretation. This form of triangulation is the “classical” approach most frequently referred to in discussions of triangulation in qualitative methods. The same researcher may (but is not obliged to) collect both data sets and the same participants may be involved in both studies, although separate groups of participants are possible. The Kvale study in the box above corresponds well with this design model.
2. *Data triangulation:* data sets on the same object of

study, but using different designs, are collected *with the same group of participants*. The data are analysed separately and the results compared with each other to obtain an overarching interpretation. The same researcher may (but is not obliged to) collect both data sets and the same participants may be involved in both studies.

3. *Researcher triangulation*: data sets on the same object of study are collected by *different researchers* using the same design. The data are analysed separately, but in the same way, and the results are compared to obtain an overall interpretation.

4. *Theory triangulation*: A single data set (i.e., same participants, same design, same researcher) is interpreted on the basis of different theories.

5. *Environmental triangulation*: The same data are collected in different settings, preferably with the same participants. “Setting” includes among other factors, time (e.g. at different times of day, on different dates or in different seasons of the year) and space (e.g. at two or more different places).

Simplifying rule of thumb:

Triangulation = Both data collections have equal status. They take place separately and the combination of information only takes place in the course of a higher meta-interpretation.

Chapter 7: Basic issues in analysing qualitative data

Qualitative analysis is a process of *discovering meaning* and involves deconstruction and reconstruction of narratives in order to tease out their more abstract and general meaning in a parsimonious form. Doing this is a subjective and dynamic process, so that the reliability and validity of conclusions about meaning are of particular interest and are of key importance for credibility and usefulness. Computer-assisted analysis offers promise of providing more rigorous findings, but software cannot replace insight or intuition, for example in developing theory. The process of analysing a narrative often starts by identifying and making sense of (clarifying) specific pieces of information (content units), followed by grouping of content units that contain the same idea to form “categories” or “sub-themes” and ultimately grouping of categories to

form “concepts” or “themes.” This iterative process, which yields increasingly abstract and general understandings of the meaning of a narrative, is referred to as “coding.” It is important that the rules for forming categories [sub-themes] and concepts [themes] are understandable, are recorded in a concrete form, and are consistently applied from case to case: Otherwise codings may be no more than the fantasies of an individual coder (i.e. unreliable) and may refer only to special ways of looking at the material (often statements in narratives) being coded (i.e. lacking in validity). Achieving consistency is usually done by constructing a *codebook*.

Chapter Glossary

<i>coding:</i>	identifying specific statements in a narrative that are understandable, concrete, specific, and relevant to the particular object of study (see <i>deconstruction</i> below)
<i>codebook:</i>	a written list of groups of content units (see next entry) that are judged to belong together because they refer to a single idea (see <i>reconstruction</i> below)
<i>content unit:</i>	a single word or small group of words in a narrative containing a single, specific idea relevant to the object of study
<i>convergence:</i>	the point in the analysis of a string of narratives where coding a new narrative yields no additional

	content units (i.e. no new ideas concerning the underlying meaning of the narratives)
<i>deconstruction:</i>	the breaking down of a narrative into a number of small highly specific statements which each mean something (see <i>coding</i> and <i>content units</i>)
<i>emergent issue:</i>	a topic introduced by respondents themselves which may not have been thought of at all by the researcher
<i>gist:</i>	the essence of the contents of a text that transcends the exact words
<i>meaning-based analysis:</i>	analysis aimed at working out the meaning of a narrative
<i>parsimony:</i>	reduction of a mass of data to the smallest possible number of broad dimensions with the least possible loss of information
<i>reconstruction:</i>	the re-combining of the specific statements obtained by deconstruction (see above) into broad general statements
<i>tabula rasa approach:</i>	commencing an analysis without imposing any structure <i>a priori</i>

Basic Principles of Qualitative Analysis

Although he was writing specifically about content analysis (see next chapter) Kromrey (2009, p. 298) made a point that can be

applied more generally when he stated that “... conclusions can be drawn on the basis of *any kind of data* by means of *systematic and objective analysis* of its constituent elements [emphasis added].” Nonetheless, there is no single, universal paradigm of qualitative research, so that it is not possible to describe set procedures that can be applied more or less mechanically in a fixed sequence in every qualitative analysis.⁴⁰ As a result, it is necessary to focus initially on underlying *principles of analysis* and broad guidelines derived from these.

Qualitative analyses usually begin with a written document, for instance the protocol of an interview or the records of a period of participant observation, notes derived from observation of a specific case, a diary, a biography or, as in the example of Gray and Kunkel (2001), a book that had already been published by a completely different author. Particularly in the case of interviews, the document to be analysed can also be the transcript of a video or audio tape. In the remaining sections of this chapter, I will focus on interview transcripts and other texts such as letters or biographies – written records of individual people’s personal constructions of reality – because the practical application of the underlying principles is easiest to grasp in this case. Despite this focus on interviews and the analysis of interview data, the basic principles spelled out below can be applied to other kinds of data. This means that the following sections can be regarded as a general introduction to deconstructing and later reconstructing narratives (see also Chapter 8).

Discovering meaning: The purpose of qualitative data analysis is to discover the *general* meaning underlying a group of sometimes quite differently worded specific statements about a specific topic. Despite differences in their specific content, at a more abstract or overarching level the statements are all related to the

⁴⁰ Such set procedures would require a single, universal, qualitative research *paradigm*, i.e., a single unified “conceptual lens through which the researcher examines the methodological aspects of their [sic] research project to determine the research methods that will be used and how the data will be analysed” (Kivunja & Kuyini, 2017, p. 26).

same broad idea – this idea is the *gist* of what the person or persons making the statements was getting at. Identifying the gist of a group of related statements is the task of qualitative data analysis. In the case of a group of people who provide narratives (the “sample”), this involves working out the gist (*essential meaning*) of groups of related statements in each narrative uttered by each specific respondent in a particular concrete situation such as an interview with a specific researcher, and then identifying the *overarching meaning* of related groups of “gists.” “Meaning” in this sense goes beyond the exact content of concrete individual statements.⁴¹ It is inferred from overlaps among individual statements either made by an individual – possibly at different places in an interview – or made by different people referring to the same object of study.

Qualitative data analysis is essentially a procedure for discovering latent structures of meaning in overlapping statements. It answers the question: “What more general – but not directly observable – factors in the mind of a person or persons (in a psychological study these factors might be attitudes, values, feelings, wishes, judgments, or understandings of the world, to give a few examples) generated this group of related specific statements?” This procedure has two phases:

1. breaking a single jumbled mass of tangled pieces of information (i.e. a narrative) into a number of separate units (content units) each with its own core meaning (*deconstruction*) and
2. putting the separate pieces back together to form an organized whole (*reconstruction*).

The process can be thought of as like taking a sponge cake and uncooking it back into eggs, flour, milk, butter, sugar,

⁴¹ A “statement” frequently consists of words uttered or written by a particular person, but the term could be used to refer to special characteristics of, let us say, an artwork or a piece of music. In the latter case, the idea of “deconstruction” and “reconstruction” still applies.

baking powder, icing sugar, whipped cream, and vanilla essence (*deconstructing* it into content units), then putting these back together in a new way: for instance, “plant products,” “animal products,” and “chemical products”⁴² and now describing the cake in terms of these concepts rather than highly specific ingredients. Using the terminology of qualitative research we would say that the individual content units (the specific ingredients of the cake) had been “coded” in a way that made it possible to understand their underlying nature or, as I put it above, their gist. The process of coding a larger mass of specific and concrete content units into a smaller number of general and more abstract principles (in the case of the sponge cake, coding for example flour, sugar and vanilla essence as “plant products” and coding milk, butter and whipped cream as “animal products”) involves a loss of detail (for example referring to eggs and milk as “animal products” is less specific than naming the individual contents of the cake). However, “animal products” is more *parsimonious*: it reduces the list of ingredients into a smaller number of broader dimensions. Nonetheless, forming these broader dimensions needs to be carried out with the least possible loss of relevant information: vegans still need to be able to recognize that they should not eat sponge cake.

The quantitative analogy to qualitative analysis is factor analysis. This procedure also seeks to identify a smaller number of more complex “factors” that underlie correlations among specific statements (usually test or questionnaire items). Although it is mathematically not correct, the logical way to understand a factor is to view it as a broad psychological dimension that “causes” specific test items to correlate. The reduction of a person’s 50 individual scores on a 50-item questionnaire to five more general factor scores is an example of the achievement of parsimony in quantitative research.

⁴² “Milk” would be a content unit, “animal products” a category and “sponge cake” a concept (see Table 8.1 for a more scientific example).

Thus, analysing data yielded by procedures such as interviews has the same purpose as multivariate procedures that are used in quantitative analyses – such as factor analysis – even if the actual techniques are quite different. Factor analysis has the advantage, however, that it is based on numerical data and mathematical procedures for discovering the underlying structures, whereas content analysis largely depends on systematic application of the logical reasoning, insights and intuitions of the researcher. As a result, factor analysis is more obviously reliable and objective according to strict definitions of those properties.

Achieving reliability and validity: As Bryman and Burgess (1994) pointed out, many writers on research methods regard the findings yielded by the analysis of narratives as much more interesting than those obtained by means of quantitative methods. However, conducting the analysis is a “problem,” and qualitative data are “an attractive mess” (Bryman and Burgess citing Miles, 1979)! These authors focused on analysing qualitative data in ways which enhance the reliability and validity of analyses. Remember that in the sense of this book, “reliability” refers to the likelihood that other researchers would come to the same conclusions as you if they analysed your data (i.e. it is not all a fairy tale you have made up), and “validity” refers to extent to which your conclusions can legitimately be applied in the real world. The desire for research fulfilling both these conditions is common to qualitative and quantitative studies. However, what constitutes reliability and validity differs in the two approaches. The first step in this chapter is to develop appropriate criteria of these two qualities that can be applied in qualitative research.

Altheide and Johnson (1998) equated reliability of qualitative research with *stability*. Data must be collected and analysed in a systematic and organized fashion that could be reproduced by another researcher if such a person wanted to do so.

Lincoln and Guba (1985) referred to this property as *confirmability*. Turning to validity, Hammersley (1992) argued that what he called “an account” (a qualitative analysis) is valid if it displays *accuracy*. It must truly represent those features of the object of study that it set out to describe, interpret and explain. Altheide and Johnson made a similar point by adding *truthfulness* to accuracy. Turning to what I call “external” validity, Lincoln and Guba (1985) emphasized *credibility*, while Hannes et al. (2011, p. 4) stressed *transferability*, and Kvale (1995) added *usefulness*.

Table 7.1: Criteria of reliability and validity
in qualitative studies

Property	Characteristic	Description
Reliability	Stability	The procedure could be reproduced by another researcher
	Confirmability	Another researcher would obtain similar results
Internal Validity	Accuracy	The representations of the object of study and the findings are not distorted
	Truthfulness	The project is presented without being biased by the researcher’s preconceptions
External Validity	Credibility	The findings make sense to knowledgeable readers
	Transferability	The findings can be transferred to the real world
	Usefulness	The practical implications of the findings are spelled out in a realistic way

Subjective nature of the analysis: As was pointed out above, in the case of quantitative analysis achieving parsimony (identifying the smallest number of broad dimensions that summarise the general meaning of a large number of separate, narrow statements with the least possible loss of detail) is guided by mathematical principles. In qualitative research, by contrast, the analysis is “meaning based” (i.e., the process of reconstruction is based on similarities in the subjective *meaning of narratives* not on numerical properties such as common variance). Recognizing these similarities is guided by the general experience of the researcher, personal hunches, beliefs about the object of study, insights derived from knowledge of the field, and similar factors (see the discussion in Chapter 3 of the qualities needed by qualitative researchers). Because of this, the researcher is the principal instrument for conducting the analysis. As a result, it is common for the entire process of collecting and analysing the data to be carried out by a single researcher.⁴³ Data analysis is thus more subjective in qualitative than in quantitative approaches. None the less, as will be emphasized below, the fact that qualitative analyses are subjective does not mean that they are arbitrary or simply reflect the whims of the researcher. On the contrary, they are systematic and reproducible (i.e., they too are “reliable,” although in a somewhat different way from quantitative studies).

Dynamic nature of the process of analysis: The next thing that must be emphasized is that qualitative data *analysis* has a dynamic or fluid nature similar to that of qualitative data collection. What the researcher does with and to the data often changes in the course of a project, as earlier phases of the analysis yield insights into the phenomenon being studied (the “object of study” – see earlier discussions). There is *a continuing interaction between researcher and data*. This means among other things that

⁴³ However, it is also possible (and increasingly common) for rules for reconstructing narratives to be worked out by a *team* of raters, for instance by means of “objective hermeneutics” (see next chapter). The rules are then recorded in a *codebook* and applied to all narratives in a particular study.

later stages of the data analysis may be altered as a result of earlier stages, while even selection of participants (sampling) in later stages of the project may be modified as a result of the initial stages of the data analysis.⁴⁴

Focus on general dimensions: In social science studies employing qualitative research methods the data often consist of narratives by people discussing their lives, not infrequently problematic aspects of these lives. Examples already given include parents of autistic children, people struggling with the decision to emigrate, drivers experiencing annoyance at being required to use a seat belt, and so on. As has already been emphasized, the purpose of the analysis is to identify the general concepts or dimensions that such people use to make sense out of their situation and to relate these to both the object of study and to existing theory. To return to Kromrey (2009, p. 298), conclusions “go beyond the specific content of the particular document being analysed.” Unfortunately, students of psychology and education sometimes forget this when they analyse qualitative data, especially data focusing on problems or difficulties in life. Sometimes they find it difficult to differentiate between a narrative and a diagnostic interview. Instead of analysing with the purpose of expanding knowledge of a psychological or educational phenomenon, they may become bogged down in a clinical analysis of the particular case, concentrating on diagnosing and even making suggestions for necessary treatment. This is possibly a particular problem for clinically oriented researchers.

It is thus important to remember the following:

1. Analysis of qualitative data is not a clinical procedure.

The purpose of the analysis is not to make a diagnosis of the particular case in order to ascertain whether the person needs therapy, or to recommend appropriate treatment.

What the researcher is interested in are the subjective dimensions respondents use to describe their situation – how they structure their own internal picture of the world

44 In quantitative research, actions such as these are regarded as close to heresy.

as they have experienced it.

2. *The task is to look for general dimensions.* The researcher's interest is the categories and concepts individual respondents use to structure their descriptions of their own situation. However, the ultimate task is not usually that of describing and diagnosing a particular case but of placing the respondent's descriptions into a more general context. Thus, the dimensions identified should be general enough to be applied to other respondents (new "cases"): In other words, a dimension like "Self-Esteem" would be better than "Low Self-Esteem," or "Inferiority Complex." "Low Self-Esteem" and "Inferiority Complex" are specific descriptions of a particular case – by contrast, "Self-Esteem" is a general dimension on which it is possible for different people to be high, low, or in between. If a second case were analysed it might be found that the second person displayed high self-esteem. This could be accommodated by the dimension "Self-Esteem," but not by "Low Self-Esteem" or "Inferiority Complex." In any case, "low" and "high" are not dimensions along which people differ from each other, but positions on such a dimension.
3. *Specialists in a particular discipline conduct discipline-related analyses.* This means that researchers should use labels for the categories and concepts they identify when analysing narratives that come from their discipline. A label like, let us say, "Beautiful Appearance" is not a psychological or educational category but "Body Image" is (although the latter term is rather journalistic – see the comment below on avoiding everyday language).
4. *Research is a scientific endeavour.* Researchers should use formal, scholarly, scientific terms. To take an example, "Feelings" is psychological, but it is informal or everyday, more characteristic of journalistic discussions in popular publications than of scholarly reports. "Affect" or "Mood"

would be better. In the same way, “Social Contacts” is too informal – “Interpersonal Relationships” would be better.

5. *Analyses should focus on the essential.* The purpose of the analysis is not to try to interpret every single word in a narrative, but only to identify expressions that seem to say something of specific interest for the object of study. Indeed, a qualitative analysis may begin with a phase of summarizing and clarifying (see below) in which much of the narrative is discarded because it is meaningless or irrelevant to the particular object of study being investigated.

Despite the principles just spelled out, interviewers should not try to behave like a machine or turn themselves into a walking test instrument. In qualitative research, the purpose of interviews is to encourage interview partners to reveal their construction of reality (their internal picture of the world) *as openly and fully as possible*. To achieve this, interviewers should make use of their full armoury of skills and abilities to lead interviews and persuade people to reveal their thoughts. This means that tact, empathy, insight and the like should be employed. However, interviewers should remember the purpose of research interviews: to obtain information about the interview partner’s construction of reality *in order to learn about the object of study*, not to establish a therapeutic relationship with the interview partner and insight into his or her specific problem.

Computer-Assisted Qualitative Analysis

The complexity of qualitative data, together with the desire to analyse narratives objectively, raises the question of whether it is possible to develop automated procedures for analysing narratives, much the same as is already the case with quantitative data (e.g., SPSS and similar software packages). Such routinization or automatization of

data analysis would add to the convincingness of results by making them more objective and would reassure researchers and readers that findings are not just figments of the researcher's imagination. This is already possible with the help of paper and pencil and the sacrifice of enough time, but as Kuckartz (2007) put it, insistence on maintaining the paper-and-pencil approach for analysing data is rather like insisting on writing a report on an old-fashioned manual typewriter (with the assistance of an eraser) despite the existence of word-processing programs such as WORD. Wilson (1996) gave an amusing example of what manual analysis involves when she described a researcher sitting on the floor surrounded by hundreds of content units written on file cards and marking relevant ones with a felt pen. This is an accurate description of the way I worked when analyzing the data for my PhD 60 years ago – not even my wife was permitted to enter the room, for fear that she might walk on and damage a vital card!

In fact, computer-assisted analysis procedures do exist. These hold out promise of

- a) faster data coding,
- b) a more detailed overview of internal relationships within the mass of data,
- c) more differentiated insights when building categories and concepts,
- d) a formal structure for formulating the rules guiding the various steps in the analysis, and
- e) more comprehensive and differentiated interpretation of the data.

Unfortunately, existing computer software is not (yet) in a position to take over the key tasks of qualitative data analysis such as gaining insights, formulating provisional hypotheses or recognising emergent themes and grasping their significance. Current software cannot replace insight and intuition, the ability to interpret and accurately understand subjective or private figures of speech, the understanding of special concepts from specific disciplines, or the

ability to recognize associations – especially unexpected or remote ones. Computer-assisted analysis of qualitative data also contains some hidden traps for the unwary:

- a) researchers can cease “immersing themselves” in the data,
- b) researchers can begin to favour numerical data over genuinely qualitative data (because they can then hand over the work requiring strict discipline to a computer),
- c) data analyses can be directed less and less by intuitions and insights of researchers and as a result be led further and further away from the “real” life of human beings and thus lose authenticity.

It might be asked whether a computer-assisted data analysis package would have identified the theme “Loss of confidence in management” in the airfreight case study (see pp. 62-63) or have recognized the key importance of this theme. Several of the chapters in Bryman and Burgess (1994) even took the pessimistic position that the special quality of qualitative data and the dynamic character of their analysis make the development of genuine automation of qualitative analysis impossible. Thus, it is important to remember that computer software *cannot do the thinking for you*, just as WORD on its own cannot guarantee that what you write makes sense.

Existing data-analysis software: Nonetheless, according to Bryman and Burgess (1994), already about 30 years ago a good dozen researchers had developed software which sought to automate qualitative data analysis (QDA). Weitzman (2000) listed numerous such software tools, including “Ethnograph” (software especially intended for use in ethnographic research but also applied in anthropology and sociology) and the amusingly named “NUD*IST 6” (*Non-numerical Unstructured Data. Indexing, Searching and Theorizing*).⁴⁵ Other well-known software available

⁴⁵ In 1997 this software was renamed “NVivo”. It now exists in various versions, including NVivo 12.

on the internet includes “Atlas/ti” (www.atlasti.com), “NVivo 2” (www.qrsinternational.com), and “HyperRESEARCH” (<http://www.researchware.com/downloads/downloads-hyperresearch.html>). A well-organized review of several different software tools is to be found on the internet at the website http://onlineqda.hud.ac.uk/Which_software/index.php. However, such programs are essentially systems for tabulating and recording the intermediate steps in a data analysis such as labelling and grouping content units. To put it simply, the programs replace the file cards Wilson (1996) mentioned and I confessed to using (see p. 189).

A more recent and comprehensive tool for analysing qualitative data is AQUAD 7 (*Analysis of QUALitative Data*), developed by Huber (Huber & Gürtler, 2012). With this package it is possible to analyse not only written texts such as interview protocols, but also videotapes, audio recordings and even pictures, as long as the data are available in electronic form (e.g., avi, wav, mp3 or jpg). Data of this kind can be analysed directly on the monitor screen and files deconstructed and interpreted. Systematic links among content units are indicated by means of mouse clicks, and “logical minimization” (reduction of the contents of the narratives to the smallest possible set of concepts which still accurately reflect the participants’ construction of reality)⁴⁶ is carried out by a comparison of cases based on Boolean logic. An English-language version of AQUAD can be downloaded at www.aquad.de/en/download/

Another procedure for computer-assisted analysis of qualitative data is “GABEK” an acronym derived from its German name (*GANzheitliche BEwältigung von Komplexität* [holistic processing of complexity]), which was developed by Zelger (see for instance Zelger, Raich, & Schober, 2008). This is a procedure for the analysis, processing and presentation of texts in normal language which is supported by the PC software tool “WinRelan” (Windows Relational Analysis). The procedure pulls together elements of related information widely distributed throughout a

⁴⁶ I refer to this process as achieving “parsimony.”

text to identify, for example, generalizations, concepts, and value systems. These are output in the form of hierarchically ordered text groups, graphical presentations of associations, cause and effect structures and similar forms of presentation. Information on this procedure can be found at <https://www.gabek.com/en/>. The software program MAXQDA is discussed in some detail by Kuckartz and Kuckartz (2002). This program is available for download as an app for Windows and Mac OS at <http://www.maxqda.com>. It can be used to analyse interview and focus group transcripts, reports, tables, surveys, literature, webpages, bibliographic data and web pages. Data can be input in the form of text, audio and video files, images, and twitter tweets. A feature of this software is that it also offers a facility for analyzing mixed methods data. This permits direct connection of standardized quantitative data with qualitative data.

Possible future developments: As D. H. Cropley and Cropley (2021) pointed out, some writers nowadays are concerned that computers are threatening to take over thinking, whereas in this book up until now I have placed great emphasis on the role of the knowledge, experience, insightfulness, empathy, and common sense of the human being(s) carrying out an analysis of qualitative data. Nonetheless, recent developments in artificial intelligence (AI) seem to contradict what has just been said: Computers may be in the process of becoming capable of exploring data (in the context of the present discussion, narratives), finding interrelated blocks of content (i.e., content units – see above) and identifying patterns (such as categories and concepts). AI systems are beginning to emerge that go beyond identifying, recording and organizing content units already specified by human users and identify content units in texts and visual “narratives” and link them into categories and concepts (see above) on their own initiative. They may even spot units or define categories and concepts that a human might miss. As bizarre as it may seem, one area in which concepts and methods that have implications for computer analysis of qualitative

data are emerging is surveillance, especially large-scale surveillance such as in arms control.

I am not referring here to giving an AI system a list of, for example, already known content units for which the program has to search in new narratives. Genuinely computerized qualitative data-analysis systems would have to *learn* to recognize content units and work out how to group them, just as computerized surveillance systems have to learn what to look for and how to interpret what they see in order to detect, for instance, illicit weapons. In fact, systems that are capable of such learning are now emerging through machine learning (ML). It is not necessary to give the AI detailed, step-by-step concrete instructions which it can slavishly follow on how to recognize what the human user is interested in. It is sufficient to give it several examples of narratives already analysed by a human user and global statements about the result of the human's analysis in each case (e.g., "This protocol is highly creative," or "This protocol is very uncreative"). In a process of "*unsupervised* deep learning" the AI then works out for itself what the highly creative protocols have in common that is absent from the uncreative protocols and uses this information to examine new narratives for which the human rating is not available. This process can involve texts, but also other forms of narrative such as pictures, videos, and similar material).⁴⁷ This occurs through the construction of "deep neural networks" which, from the point of view of the human user, remain a "black box," since the human user does not know how the AI is recognizing content units, but only whether its decisions match those of a human rater. Once software has gained experience with narratives related to a particular object of study and taught itself how to recognize relevant material it can analyse new narratives on the same topic at very high speed, instead of possibly in hours.

⁴⁷ The kind of approach, where the computer is given concrete instructions on what to look for, is referred to as "supervised" AI. Procedures in which the computer learns for itself are "unsupervised."

A discussion of how AI can learn, how deep neural networks are formed, or what is in the black box would go well beyond the limits of this book and this discussion of the application of AI to the analysis of qualitative data is intended only as a speculation concerning possible future developments, not as a specific plan for action.

In a concrete practical discussion of AI analysis of qualitative data, Mahto, Hogan, Hatfield, Coppola, and Kulkarni (2021) described the application of AI to analyzing responses of workers who were asked for their views on what keeps them motivated in the workplace and what needs to change. Mahto et al. (2021, no page number) referred to the conventional approach I have outlined in earlier sections and will elaborate upon below as “the old-school method” which is based on “coding,” “categorizing,” “comparing” and “synthesizing,” and asked whether “using machines to analyze data could yield time and cost efficiencies and enhance the value of insights derived from the data.” They referred specifically to improving the *validity* of findings (see the discussion of validity earlier in this chapter and summarized in Table 7.1) and recommended “looping in your sidekick” (i.e., making use of AI) as the “better way.” Ultimately, however, these authors came to the conclusion (Mahto et al., 2021, no page number) that current AI qualitative analysis is markedly inferior to analysis by human experts: “... while machines can do a good first pass, human oversight and review is still often necessary.” They concluded that in its present state of development AI has not advanced beyond being a *useful tool* for analyzing qualitative data. Thus, it seems that computerized analysis of narratives – as against using computers to organize and store material identified by humans – is still some way away. However, Cropley and Marrone’s (2021) research on using AI to evaluate figural creativity tests suggests that substantial progress is being made. After being shown examples of highly creative

and extremely uncreative drawings, with no additional information, the AI worked out the difference for itself and subsequently scored drawings it had not previously seen with a high level of agreement with human raters (except that the AI required only a tiny fraction of the time the humans needed).

The Role of Theory in Qualitative Analysis

Strauss and Corbin (1998, p. 168) cited the following dictionary definition of theory: “A coherent group of general propositions used provisionally as a principle of explanation for a class of phenomena.” I will adopt this definition in discussions both below and in the following chapter. In the present chapter, the focus is neither the use of *existing* theory as the starting point for setting up a research study, nor generation of *new* theory on the basis of the study’s findings (this is discussed in Chapter 9). What is of interest in the present section is the role of *existing* theory in guiding or pre-structuring the analysis of a particular set of data.

Finding emergent issues: In fact, analysis can proceed without any reference to theory at all, via what might be called the *tabula rasa* approach. This involves simply seeing what emerges from the respondents’ statements without imposing any structure *a priori*, letting the chips fall where they may, so to speak. The advantage of *tabula rasa* approaches is that they do not specify in advance which contents will receive the main focus. As a result they make it possible for the analysis to take account of “emergent issues” (Ritchie, & Spencer, 1994, p. 180). These are introduced by the respondents themselves and may not have been thought of at all by the researcher. Such issues are very important because,

as Elliott (1999, p. 252) put it, “qualitative researchers hope *to discover something new and unexpected* [emphasis added].” The incorporation of emergent issues into the analysis is an example of the dynamic nature of qualitative analysis that was described above.

One way of looking for emergent issues has been suggested by Ritchie and Spencer (1994). The researcher starts the analysis without any pre-formulated expectations of what will emerge, or at least by behaving as though there are no pre-existing expectations, since it is hard to believe that a typical researcher would not have given any thought whatsoever to possible results. The first step is an initial phase of “familiarization,” in which the researcher gains an overview of the range and diversity of the contents of the narrative via a process of “immersion in the data” (p. 217). This may be done one case at a time or with several or all cases. During familiarization the narrative or the narratives are read through several times, in order to develop a feel for the contents. On the basis of this familiarization an understanding of the ideas contained in the narrative(s) is developed. Those ideas that are relevant to the object of study are identified and then used to structure the content analysis that follows. Which ideas are selected as the framework for the more structured analysis that follows familiarization depends upon the special interests of the researcher.

Theory guided analysis: Although as has just been pointed out, an analysis can be conducted without an initial theory, qualitative analyses can also be guided and structured by existing theory. One way of doing this is to define in advance the ideas that are expected to emerge and check in the analysis whether they did actually appear, or in what form they appeared. When this is done, the analysis takes on some of the characteristics of hypothesis testing. Restricting the analysis to issues that theory suggests ought to emerge provides a preliminary structure and goes some way towards routinizing the analysis. However, it has the disadvantage that emergent issues (see above) may be overlooked or ignored.

Theory-guided analyses need not be as strict as the form

just outlined. As Strauss and Corbin (1998) pointed out, the initial theory may be not existing formal theory as it is found in the relevant literature but the researcher's own intuitive or naïve theory emerging from hunches or personal experience with the phenomenon in question. Well-qualified researchers who possess deep knowledge and substantial experience of the phenomenon under investigation are particularly well placed to generate their own theory because of

- their training,
- their familiarity with the relevant literature,
- their knowledge of prior research or even their own earlier investigations,
- their specialized disciplinary and professional knowledge.

They are “theoretically sensitized” (Strauss & Corbin, 1998, p. 166) and are thus likely to notice relationships and connections that might not be apparent to people without the necessary expertise.

Choosing a theoretical orientation: Researchers may choose a particular theoretical approach (e.g., psychoanalytic, social psychological, learning theory, critical events theory) because they have a high level of familiarity with it or even special expertise, or because it fits well with their own interests. An approach may also seem to the researcher to be especially useful for studying a particular object of study in a particular way. To take an example, in the study of German migrants to Australia already mentioned a psychoanalytic orientation was very useful because (a) psychoanalysis emphasizes “hidden” motivation, precisely the issue on which the study came to focus in the course of the analysis; (b) psychoanalysis includes a developmental model and thus offers leads on how people interested in migrating got to be like that; (c) psychoanalysis has its own approach to therapy and thus suggests procedures for counselling people experiencing psychological distress in connection with migration. The freedom to choose the theoretical orientation raises the problem of apparent arbitrariness and thus lack of rigour. For this reason, it is important to show that

there were systematic, scientific reasons for adopting a particular approach, i.e., that a particular tool was chosen because it was precisely the one needed for the job in hand.

In the dissertation already mentioned on p. 105, Eidson (2021, p. v) stated the object of study at the very beginning:

The purpose of this qualitative descriptive study was to explore the experiences of working mothers in the financial services industry in the U.S. and the impact of mentorship on feelings of identity and self-efficacy after an extended child-rearing leave.

She then went on to explain that “seasons of life” theory (Levinson, 1996) offered an appropriate theoretical framework for the research questions, and identified the core elements of this theory for the purposes of her research (she was “sensitized” to these areas by her interest in and knowledge of seasons of life theory): These core elements included identity, self-efficacy, feelings of personal worth, competency, goal setting, challenges, and connection. Ultimately she identified five overarching themes related to these dimensions in the narratives of her participants: (a) feelings of identity and self-efficacy, sense of personal worth, personal versus professional challenges, and pressure when returning to work; (b) merging personal and professional identities when returning to work; (c) reinventing professional competencies and self-efficacy; organizational intent to provide connection, goal-setting, and path development through mentorship; the value of mothers in a male-dominated industry and the importance of mentorship to support them.

Coding Content Units

In quantitative studies the “meaning” of participants’ responses to researchers’ inquiries is relatively obvious and objective, since the responses very frequently consist of scores on instruments (e.g., reaction time, galvanic skin response time, etc) or standardized answers to items in a questionnaire such as “What was the highest level of schooling you reached?” or “To what extent do you agree with the statement, ‘Children should always obey their parents’?” Analyses usually begin with the entry of highly structured responses on an answer sheet (i.e., construction of a data matrix). This may require converting answers into numerical form. For example, the answers to the question about level of schooling might be coded as follows: incomplete elementary school = 1, completed elementary school = 2, incomplete secondary education = 3, completed secondary education = 4, incomplete tertiary education = 5, completed tertiary education = 6, postgraduate education = 7. Responses to the question about children’s obedience might be coded in the following way: “I disagree completely” = 1, “I disagree to some extent” = 2, “I cannot make up my mind” = 3, “I agree somewhat” = 5, “I agree strongly” = 5. The purpose of coding in this sense is usually to convert relatively subjective information to objective forms (as a rule, numbers).

Earlier remarks in this chapter may have created the impression that the rigour of quantitative methods is abandoned in qualitative research: i.e., that anything goes as long as it is surprising. Indeed, in qualitative analyses it is almost always necessary to tease out relevant ideas contained in sources such as interview transcripts, participant-observation protocols, field notes, documents, websites, video tapes, or photographs. As Saldana (2009, p. 3) succinctly put it, qualitative analysis usually involves finding elements in a flow of information that identify a “salient, essence-capturing, and/or

evocative attribute” of language-based or visual data. I refer to such words, short phrases or images as “content units,” although they are often labelled simply “codes.” This process of “teasing out” by identifying “salient” or “evocative” content units raises the issue already referred to of ensuring and demonstrating that there really is a difference between qualitative research and fiction. Qualitative analysis must be at pains to avoid becoming a matter of flights of the literary fantasy of the person analysing the narratives: Achieving this is the task of *coding* of narratives. It is possible to distinguish between “substantive” coding, which has the purpose of identifying specific direct references to the object of study, and “theoretical” or “higher order” coding, which involves finding content units that refer to abstract and general ideas that form discipline-based concepts such as, in the case of psychology, “motivation,” “coping strategies,” “self-image,” and the like.

Clarifying and summarizing: In narratives, what the respondent wants to say may be relatively obscure or hidden. The first step in qualitative coding thus involves identifying fragments of information (such as statements in a narrative) that contain an understandable idea. At its simplest, this is necessary because “raw” narratives often contain a great deal of material that cannot be used in a particular study, especially when the text is an interview protocol. Material may be (a) meaningless or unintelligible (an obvious example is expressions in an interview such as “Ah,” “You know,” or “Mmmm?”), (b) repetitious (repetition may indicate the strength of a participant’s views or the importance of an issue for that person, so that it may well be important, but it does not introduce new themes), or (c) irrelevant (“irrelevant” in the present context means that a piece of information does not refer to the specific object of study; irrelevant content need not be trivial – it may be very important to the respondent and fascinating to the interviewer, but simply not informative in the specific context).

Mayring (1983) outlined a number of specific steps in this first stage of qualitative coding, of which two are of particular interest

at this point: (a) *summarizing* and (b) *clarifying*. Summarizing involves identifying *relevant* material in a narrative by picking it out from the general mass of information in the narrative; as I said above, this general mass may be interesting, but for research purposes we only need the information that is directly relevant to the object of study. Clarifying requires making the respondent's intended meaning unambiguously clear. For example, respondents often use slang or special terms that have a particular meaning to them personally or to specialists such as other workers in the same area. They may also use metaphors or other indirect forms of expression. Clarifying often involves turning slang into standard language. It sometimes requires the use of a dictionary, or consultation with a person who is thoroughly familiar with specialized terminology. When such aids do not resolve the respondent's meaning, careful reading of the text and placing of the difficult content unit in context will often succeed. Petersen (1989) gave a helpful example in her *Diplomarbeit* (equivalent to a master's thesis) on hobby authors. In an interview, one respondent referred to the role in her creativity of a "fantasy cage." This term was unknown to the researcher and is not found in dictionaries. However, reading of the relevant statement in its special context made it plain that what the participant was referring to was her fear of expressing her own ideas to other people – she felt as though her fantasies were locked away in a cage. This idea is understandable and of obvious relevance to a discussion of factors affecting the productivity of hobby authors. Summarizing and clarifying are practical steps that are frequently necessary before the actual analysis can even commence. After these steps, a "clean" text containing intelligible, relevant statements is available.

Developing a codebook: Once the fragments of a narrative (i.e., the content units) have been rendered comprehensible (i.e., clarified and summarized) and those containing information relevant to the issue being studied identified, the next step in qualitative coding is to establish what these content units *mean* in a *more general* sense. Apparently different content units may refer

to the same underlying idea such as, let us say, anxiety, a need to become famous, or a sense of entitlement, even though the specific words used to express these ideas may differ from content unit to content unit. Identifying the more general meaning of specific content units makes it possible to group those units that contain a similar idea, even if they are located at different places in a single narrative or in different narratives. This is possible despite the fact that the specific wording of statements may be quite different from person to person. In the example of the “fantasy cage” just mentioned, only one respondent used this particular expression, but several referred to the general idea that their creative thoughts were locked in or blocked in some way. Thus, despite large differences in the exact words used, several content units contained the same idea, in much the same way that “water,” “Adam’s ale,” and “aqua pura” all refer to the overarching idea of water, although they look quite different. Identifying and recording such overarching meanings of content units is the task of the second aspect of coding.

Whereas the first notion of “coding” (see preceding paragraphs) involves identifying and tidying up individual content units, the second involves expressing the *meaning* of these units in more general, but stable, terms: In the fantasy cage example, for instance, the more general idea underlying the specific words chosen by the particular respondent was that her novel ideas were somehow blocked, so that in the codebook this content unit would be labelled something like “blocking of ideas” (“blocking of ideas” is a label or *code* that transcends the exact wording of the original remark and instead emphasizes its general meaning). This aspect of the coding of qualitative data is much more subjective and much more threatened by individual foibles of the researcher than the first aspect (eliminating errors, identifying content that is relevant, etc) which is in effect a mere technical procedure. It is here that the need to avoid flights of fancy becomes acute. What is needed is a research process that is “logical, traceable, and clearly documented, particularly on the methods chosen and the decisions made by the

researchers” (Hannes et al., 2011, p. 4). In simple terms, what is needed is what Hannes and colleagues labelled “dependability.” Consequently, establishing a stable set of rules for identifying the underlying general meaning of different statements (both within and between protocols) is an important aspect of qualitative research. These rules are recorded in the *codebook* by the researcher.

In a sense, the codebook of a particular study resembles the scoring key for a psychological test, except that in the case of standardized tests responses are usually delivered in a pre-defined form, for example by checking one of a number of pre-specified alternative answers (no clarifying and summarizing are needed), the test manual remains the same for every setting in which the test is applied, and assumes that there are standardized answers that transcend the individuality of the people being tested, and answers are often objectively right or wrong (e.g. any respondent who ticks, let us say, box (b) is right, anyone who ticks any other box is wrong – no discussion needed). In a qualitative study, the codebook refers to a specific set of responses from a specific group of participants under particular conditions, and also reflects the influence of a particular researcher (who him- or herself also possesses a specific, partly privately, partly socially constructed reality).⁴⁸

The codebook contains a list of more precise, discipline-related labels such as “thinking processes” and provides – for each overarching term – a list of specific, concrete statements taken from the narratives being analysed that refer to the overarching term in question (these specific concrete statements are what I call “content units”). For example, the codebook might specify that the content unit “I spend my time exploring options” and other closely related specific, concrete statements (content units) found in various narratives (such as “I like to try different approaches,” or “Thinking

48 An example of the pitfalls connected with clarifying the intended meaning of idiosyncratic expressions is seen in the coding of interview data in a recent research report. The respondent said, “I’ll bet you’d look good in that.” The researcher erroneously coded this content unit as expressing the participant’s desire to gamble!

things over is my greatest strength) are all to be coded as referring to the more general psychological area of “thinking processes.” Thus, the codebook functions as something like a dictionary: once it has been constructed you can look up a particular content unit in a new narrative and code it as set out in the codebook. You can also return to a narrative that had already been analysed earlier and code or recode it according to the codebook, even if the codebook did not yet exist at the time you did the original coding of that particular narrative.

A major issue that was touched upon above is that the assignment of specific content units to particular codings is relatively subjective: One researcher might code “fantasy cage” differently from another. For this reason, it is important that (a) the rules for coding content units should be formally recorded (i.e., that there is a written codebook) and (b) that the rules should not be simply products of the special experience and private opinions, or even the passing moods, of a particular researcher, but should reflect the commonly understood meaning of language or other symbol systems that constitute the narrative. One way of avoiding excessive influence of the particular researcher on the contents of the codebook is for it to be constructed by several coders, who work together to establish the overarching meaning of content units. One approach to what might be called “team coding” is discussed in a different context in the next chapter where the process of *objective hermeneutics* is presented (see p. 228). Team development of a codebook can be seen as a specific application of objective hermeneutics.

In the curriculum case study outlined in the following box, specific content units such as “making unusual associations,” “linking ideas,” “suggesting different options,” and “generating surprise” were identified by a coding team as all referring to “thinking processes.” Other overarching codes the team identified for different blocks of content units were “definition of creativity,” “personal properties,” “classroom activities,” “teacher support

measures,” and “teacher training.” These codes were then used as more general labels for grouping content units that had a single idea in common. Coding is normally carried out one narrative at a time, starting – unsurprisingly – from the very first one, so that the codebook grows as additional narratives are analysed and additional codes identified. However, practical experience suggests that after perhaps 12 narratives *convergence* usually occurs: additional intelligible and relevant content units may be identified, but they refer to already existing codes and the principle of *parsimony* indicates that the codebook is complete.

In an analysis of written documents (Patston, Kaufman, Cropley & Marrone, 2021), the discussion of creativity in several national school curricula was analysed in terms of three pre-defined questions (i.e., according to existing discussions of creativity – not tabula rasa):

1. How is creativity defined?
2. Where is creativity placed in the curriculum?
3. Is there concrete advice for teachers about nurturing creativity in their students?

Coding started by identifying statements (content units) in each document which referred unmistakably to one of the themes just listed. A set of rules was developed for deciding how to code statements (i.e., a codebook). For example, “exploring options and alternatives, implementing ideas and taking action” (a statement [content unit] in one curriculum document) and “seeing existing situations in a new way, identifying alternative explanations, and seeing or making new links” (a statement [content unit] in a different curriculum document) were both coded as referring to “thinking processes.” Other statements were coded as referring to different overriding categories such as “playfulness,” “curiosity,” “openness,” “self-confidence,” and “autonomy. A number of themes emerged

(see the discussion of “emergent” themes on p. 195) that went beyond the initial three research questions listed above, and so the documents were coded a second time, searching for content units that referred to these new themes. Thus, the analysis was an iterative and reflective process. It involved an initial theory-based definition of the dimensions of the object of study (definition of creativity, place of creativity in the curriculum, concrete advice for teachers), identification of additional themes that emerged from the narratives examined for the purposes of the study, addition of the new themes to the codebook, and a re-analysis of all content units.

Iterative coding: Not infrequently, codes are identified in later narratives that were not detected when earlier narratives were analysed, or it may become apparent that in some earlier narratives content units were misunderstood. In the aircraft freight loader case study, for example, the code “dissatisfaction with senior management” was not recognised in the analysis of the first few narratives but emerged during construction of the codebook. This means that earlier narratives may need to be re-visited using the now complete codebook. Coding is thus an iterative process. In quantitative studies there is usually no iteration, since the quantitative equivalents of “codes” are already specified in the test manual and scoring procedures are pre-defined. In qualitative research, by contrast, codes “emerge” from the data and may not be identified (“teased out,” as I put it above) until several narratives have been analysed.⁴⁹

49 See the discussion of emergent issues on p. 195 above.

Chapter 8: The process of analysing qualitative data

Qualitative analysis can be regarded as a process of moving from concrete, everyday-language statements (narratives) that describe the experiences of individual respondents (i.e., single “cases”) to increasingly abstract and general, “scientific” or “discipline-based” statements that apply to many cases. This movement from specific statements “content units” (or “codes”) towards a smaller number of increasingly abstract and broad underlying principles involves three steps: (1) identifying content units, (2) grouping blocks of content units with related content into “categories” or “sub-themes” and (3) combining categories or sub-themes into overarching “concepts” or “themes.” In the case of steps (2) and (3) this grouping is carried out in accordance with rules set out in a codebook constructed by the researcher for each specific

research project. Both categories and concepts (sub-themes and themes) may be revised or added to as the analysis of individual narratives proceeds, with the result that narratives analysed in accordance with an earlier version of the codebook may be re-assessed (and indeed the codebook itself revised). The process of coding may be carried out *deductively* or *inductively*. There are two fundamental ways of doing this: the holistic-intuitive approach (e.g., analytic induction) and the atomistic-logical approach (e.g., grounded theory). There are different variants of both approaches.

Chapter Glossary

<i>analytic induction:</i>	reading the first narrative as a single whole and forming an overall impression of its general meaning, then refining this understanding by reading further narratives one after the other (contrast with grounded theory – see below)
<i>atomistic-logical analysis:</i>	breaking individual narratives into fragments (content units), with each fragment containing a specific, relevant idea
<i>deductive reasoning:</i>	working out specific details from a general premise (such as a theory) which already exists (contrast with inductive reasoning)
<i>grounded theory:</i>	deconstructing narratives into highly specific content units (see atomistic-logical analysis) then recombining

	these units to form more general ideas (see deconstruction and reconstruction)
<i>holistic-intuitive analysis:</i>	forming an overall impression of the meaning of narratives by reading each one as a single overall whole (contrast with atomistic-logical analysis; see above)
<i>inductive reasoning:</i>	combining separate bits of specific information to form a new general premise (contrast with deductive reasoning)
<i>sensitizing concepts:</i>	concepts which emerge during an analysis and subsequently direct raters' attention to specific ideas in narratives
<i>sub-themes:</i>	more general, but still relatively narrow, ideas underlying a group of statements found in a single narrative or present in only a few narratives
<i>thematic analysis:</i>	an approach to ascertaining the general, abstract meaning of groups of related content units. This often has two phases: identifying sub-themes followed by identifying themes
<i>themes:</i>	general ideas underlying a high proportion of narratives, often identified by combining sub-themes (see above)
<i>theoretical sampling:</i>	after a provisional understanding of the object of study has been established deliberately selecting

new narratives that are likely to support this understanding, in order to enrich it

Working out the Meaning of a Text

The previous chapter identified and explained the nature and purpose of the various elements of a qualitative study such as “content units,” “sub-themes,” and the like. The present chapter focuses on the *processes* through which such elements are obtained. For example: How do you actually go about developing content units?

Inductive vs. deductive analysis: In the broadest terms, there are two ways of approaching a qualitative content analysis: the inductive approach and the deductive approach.

1. A “pure” *inductive analysis* begins, at least in theory, *without any pre-existing expectations whatsoever*, i.e. without any theory. In this oversimplified vision of analysis, narratives⁵⁰ are regarded as a diffuse mass of separate bits of information. The researcher’s task is to dig out broad, overarching ideas that are embedded in the mass of information. The question that has to be answered is, “What broad concepts or themes are contained in this narrative?” As a rule, it is possible to identify such originally-hidden topics; people do not usually talk at random when discussing things they know a lot about (i.e., matters about which they possess

⁵⁰ Remember that although the word “narrative” most often refers to spoken statements and that these statements are very frequently transcribed into written texts by the researcher, “narratives” sometimes consist of existing documents such as press reports, speeches, school curricula, laws, political proclamations, poems, films, etc, or even non-verbal forms of communication such as paintings. Although the remarks that follow refer explicitly to transcriptions of spoken communications such as interviews, they can readily be adapted and applied to existing texts or other forms of communication such as those just mentioned (documents, films, works of art, etc).

narrative competence). Such topics are said to “emerge.”

After reading two or three narratives (for instance interview protocols), it is usually possible for researchers to identify a relatively small number of themes mentioned in the narratives that are “interesting” for the particular object of study: Concepts/themes are interesting when they cast light on some matter that a researcher wishes to pursue because the matter is relevant, useful, informative, and so on (qualities connected with external validity). Typically, the inductive approach makes it possible to identify a small number of more general themes in the larger number of diverse statements interview partners make in the course of discussing their construction of reality (such as the theme of trust in senior management, or lack thereof, that emerged in the freight-loader case study). Once a concept/theme has been identified, the researcher can look either backwards and check whether and how this theme emerged in *earlier protocols*, or forwards by focusing on the issue just identified when reading additional, as yet unread protocols.

2. The *deductive approach*, by contrast, *begins with focused expectations*, usually derived from existing theory: The general ideas (i.e., concepts or themes) already exist. The researcher’s task is to discover possible differences in the details of how different interview partners understand these principles. For example, in the national curricula case study outlined in Chapter 7 (see p. 205), the broad dimensions through which the curriculum developers constructed creativity in their minds had already been worked out in a review of the literature. These were: (a) the definition of creativity, (b) the place of creativity in the curriculum, and (c) ways of nurturing creativity. The research task was to identify specific references to these issues in different narratives – in this study, a narrative was a curriculum document⁵¹ and the object of study was (a) the way

51 See previous footnote: Referring to curriculum documents as “narratives” may seem odd. I do it here to emphasize that texts being analysed (i.e., narratives) need not be interview protocols, even though they usually are.

creativity is defined in the various different curriculum documents, (b) what recommendations are made concerning how to incorporate creativity into the structure of children's studies – for instance as a separate discipline or by means of a once-a-week creativity orgy, and (c) what specific classroom activities are recommended in order to foster creativity. In analyzing a narrative (a curriculum document) researchers could ask themselves questions like, "Is there a definition of creativity at all in this particular narrative?" "What different specific form(s) does the definition take in the narratives?" "What role is assigned to creativity in the narrative?" "Where does the narrative see creativity as coming from?" or "How is creativity seen as interacting with other factors?"⁵²

The next step – a controversially discussed aspect of qualitative coding – involves the strategies for identifying specific examples of an already existing general understanding of the object of study (deduction in the sense just explained) or for identifying the more general meaning of narratives in the absence of a pre-existing understanding (induction, as explained above). Suppose that a narrative does contain a content unit that has been identified as relating to the role of creativity (i.e., that has been clarified and summarized): What is the more general meaning of this content unit? There are two ways of working out what the content unit means, which are referred to here as the "holistic-intuitive" approach, on the one hand, and the "atomistic-logical" approach, on the other. Both of these can be applied inductively or deductively, although they are more commonly associated with inductive analyses.

The holistic-intuitive approach: When using this approach the researcher reads the entire first narrative,⁵³ possibly

⁵² Remember, because this is an example of a *deductive* analysis, these questions were already defined prior to the commencement of the analysis. In the case of an *inductive* analysis the questions would *emerge* as the first step in the analysis.

⁵³ Although the following discussion applies to all qualitative analyses, for ease of understanding I will assume in the following sections that the analysis is being carried out with interview protocols which are available in written form (i.e., in the form of a text).

several times, until he or she is ready and able to propose a speculative interpretation of its *contents as a whole*. A second text is then read and the original interpretation revised as necessary in order to develop a corrected interpretation that holds true for both texts. The procedure is repeated on the basis of additional narratives and the interpretation revised each time the new narrative deviates from the previously existing interpretation, so that the interpretation applies to all the narratives that have been analysed. After a number of narratives have been analysed and a first understanding formulated of what the texts as a group are saying about the subject of interest this understanding can be broadened and deepened by adding new themes or the existing themes can be deepened, differentiated or clarified. The procedure is repeated with further protocols, one after the other, and the interpretation “corrected” step by step, until the addition of further new protocols leads either to no further change in the interpretation or to only small, insignificant changes. At this point what I call “convergence” has been reached: Additional information brings improvements so minor that, applying the law of *parsimony*, no new cases are needed.⁵⁴

The atomistic-logical approach: This approach does not focus on the text as a whole and does not depend primarily on the researcher’s intuition. After clarification and cleaning of the text (see for example p. 234), the narrative it contains is deconstructed into small units consisting of brief statements (groups of related words, incomplete sentences, etc). These statements are referred to in this book as *content units* – each content unit contains a *single* idea. These content units are the building blocks of which the original narrative consisted. As is explained in more detail in the next section, content units can be identified relatively objectively according to logical rules specified by the researcher and stated in writing in the codebook. They are not the result of intuitions of the researcher but are physically and literally present in the

54 My practical experience is that this often occurs after about 12 narratives (see the discussion of sample size).

narrative; they should not be “interpreted into” the text. As long as the codebook rules for identifying content units are clear, concrete, specific and easily understandable, it can be expected that a second researcher would deconstruct a given text into the same content units as the first. In other words inter-rater *reliability* is enhanced.

Once the content units have been “revealed” or “discovered,” they are combined into groups of content units with related content (i.e., “categories” or “sub-themes”), and subsequently into higher-order groups of groups (“concepts” or “themes” – see below). In this way, a new structure is built out of the content of the original text. This approach involves the systematic *deconstruction* (i.e., identification of specific content units) and subsequent *reconstruction* (i.e., combination of content units into categories/sub-themes and concepts/themes) of the original narrative. When the point is reached where the deconstruction of a new protocol yields no new content units or only content units that add nothing to the existing reconstruction, convergence has been achieved and parsimony requires that no further protocols are needed. The atomistic approach is much more systematic and objective than an approach based on intuitions: Summing up, the holistic-intuitive approach is based on “felt” interpretations, whereas the atomistic-logical approach is based on objectively worked out “facts.”

Finding content units: When identifying content units, the following should be remembered:

1. Content units need not be complete sentences. They are often grammatically incorrect fragments couched in everyday language.
2. They are usually brief, possibly only a single word. Aim at units that are as brief as possible (up to 3 or 4 words), and avoid fragments larger than 7 or 8 words, except in unusual circumstances (e.g., a respondent says something longer that captures a complex idea with exceptional eloquence and clarity).
3. Focus on statements that are specific and concrete,

rather than vague or general.

4. At least in the first run through a narrative, eliminate from consideration meaningless, repetitious or irrelevant material fairly ruthlessly. Concentrate on contents that are understandable and relevant to your object of study. Your task is to investigate an object of study, not to summarise everything the respondent thinks about everything!

5. Be conscious of the reasons for excluding particular statements from the analysis (e.g., the statement is too general or does not refer to the object of study) and be able to state these reasons if you are asked to do so. Give examples.

6. Do not “over-interpret” respondents’ statements. This happens, for instance, when the researcher reads into a statement a meaning that may well have been in the respondent’s mind, but is not actually in the statement in question. In the example below of an interview on a social worker’s feelings about her job, the respondent commented that residents of refugee hostels “don’t seem to make much progress.” It would be possible to argue that the statement referred to a sense of disappointment or frustration. However, as it stands the statement is not about the respondent’s feelings, in the way that, “*I am frustrated* by their lack of progress,” would be.

Theoretical sampling: Glaser and Strauss also proposed a further phase, which is not absolutely necessary for grounded theory but is sometimes employed – the phase of *theoretical sampling*. Theoretical sampling occurs when, after an initial understanding of the broad meaning of a set of narratives has been developed and substantiated by several consecutive cases, further cases are deliberately selected to cast light on this understanding. There is no pretence of randomness in selecting additional cases. On the contrary, there is a deliberate focus on cases that have special potential for expanding, clarifying or refining the newly-

acquired understanding. This means that the cases studied in the phase of theoretical sampling are very **unlikely** to contradict the explanations already formulated: Indeed, they are selected because it is anticipated that they will support these explanations! Their function is not to correct the analysis that has been carried out, but to elaborate and clarify it. Theoretical sampling is more or less the opposite of random sampling, a crucial procedure in experimental designs and therefore of great importance in typical quantitative studies.

Sensitizing concepts: Both analytic induction and grounded theory can be seen as procedures for using empirical evidence to develop theory. They give this process a certain rigour. Fundamentally, the end product of the research is theory. However, Blumer (1969) argued that theory has no purpose except to make the observed world easier to understand, i.e. theory is valuable, but only as an instrument for achieving better understanding of concrete, empirical events. This means that what is needed from an analysis is broad, general statements (i.e., theory) that show researchers how participants understand the world around them. Emerging or partial understanding (theory) developed early in an analysis “sensitizes” the researcher to certain kinds of material. This idea is at the core of the concept of “sensitizing concepts.” This discussion might seem to belong to earlier sections on the role of theory in beginning an analysis. However, sensitizing concepts do not exist prior to the commencement of the analysis and are developed during the analysis (i.e., it is a preliminary result). After a theory of this kind has *emerged* in this phase it can “guide” the balance of the analysis: the researcher is sensitized for data that are compatible with the newly-formulated theory and looks specifically for such data – his or her interest is directed by the sensitizing concept(s). Once such theory exists it should, according to Blumer, be checked out in new situations involving different objects of study in order to establish its generality and refine and test it. One qualitative way of doing this would be to observe in which real life

situations it applies, what special forms it takes in different settings, and so on.

Analytic Induction and Grounded Theory

Analytic induction: Analytic induction is the classical example of the holistic-intuitive approach outlined above. The initial foundation of analytic induction is usually attributed to Znaniecki (1935). It is a process of formulating an understanding of the general meaning of narratives and generalizing this understanding to other cases or situations. The practical steps involved in this approach are as follows:

1. The problem on which research is to be conducted is identified at the beginning of the project, even if initially in a fairly rough, approximate way (e.g., problems of families with children suffering emotional disturbances, failure of drivers to use seatbelts). This is what I call the “object of study.” The object of study can be derived from existing knowledge (e.g., previous research or a widely accepted theory), or subjective (e.g., personal experience or intuitions).
2. Data on the problem such as interview protocols or written narratives are collected from a number of appropriate people (e.g., parents of emotionally disturbed children, drivers fined for failure to fasten their seat belts) – people who possess narrative competence).
3. The researcher reads the narratives carefully, several times if necessary, in order to become thoroughly familiar with their contents. This is the process of “immersion” in the protocol.

4. On the basis of this intense reading a working hypothesis about the phenomenon under study is formed (e.g., “Parents experience feelings of guilt about their children’s disabilities,” or “Drivers refuse to wear seatbelts because they regard themselves as fighting for freedom”).
5. Further narratives by new respondents are read.
6. The contents of the new narratives are examined in order to see how well they fit the working hypothesis.
7. Wherever a discrepancy occurs between the working hypothesis and the new data, the hypothesis is revised until it fits all cases studied to date (e.g., “Parents *who receive no psychological support* experience feelings of guilt,” or “*Authoritarian* drivers refuse to wear seat belts because they regard themselves as fighting for freedom”).
8. Steps 5-7 are repeated (i.e., further new narratives are read and used to test the working hypothesis, which is revised if necessary).
9. When new data reveal no need for any further revision of the working hypothesis the procedure comes to a halt; the latest version of the working hypothesis becomes the finding of the study.

It is important to note that the analytic induction procedure just outlined

- involves an idealized approach that usually cannot be realized in practice. To reach a point where all additional cases fitted perfectly with the existing hypothesis would usually require either analysing a very large number of cases or formulating the hypothesis in an extremely complicated or an extremely general form, or both of these. In practice, it would be more common to stop after a relatively small number of iterations of steps 5-7. A practical reason for stopping would be time constraints – graduate students and researchers typically want to complete their thesis in the foreseeable future, and cannot

allow unlimited time for data analysis! The extent to which the latest version of the working hypothesis still fails to fit new cases then forms part of the discussion in the research report;

- does not separate data collection and analysis – in the way that quantitative studies usually do – into two discrete phases, the first (data collection) being completed before the second (data analysis) commences. On the contrary, the two aspects alternate. This means that in writing up research in a qualitative report it may not be possible to describe “Instruments,” “Procedure,” “Data Analysis,” and “Findings” in neatly separated chapters. It is also common for steps 2 and 3 to consist initially of a kind of pilot study in which the object of study is refined, the ability of respondents to make sense of the situation or their willingness to reveal themselves confirmed (or not), questions or cues revised or clarified, and so on. Thus, there may be a cycle with steps 2 and 3 being repeated at least once, and possibly several times (i.e., iteration);
- does not start with a strict hypothesis. On the contrary, the hypothesis emerges during the process of data collection and analysis (i.e., the process is inductive). Furthermore, the final result of the investigation is itself more in the nature of a hypothesis or a working framework for describing the phenomenon under investigation. Miles and Huberman (1994) suggested a compromise which pre-structures the process of data reduction from the beginning, but does not involve using the analysis to test a completely worked out theory. The researcher may state a general *frame of orientation* before commencing content analysis and then look for content units referring to this orientation.

Grounded theory: It is apparent from what has just been written that analytic induction has a major weakness in

comparison with quantitative methods: It does not yield clear data that are capable of “proving” a hypothesis such as “Parents of children with multiple handicaps who receive counselling experience significantly less guilt than uncounselled parents,” or “Drivers with an authoritarian personality structure reject the use of seatbelts significantly more often than other drivers.” As a result, its theoretical statements lack the apparent objectivity that is characteristic of quantitative methods. A procedure that goes some way to dealing with this problem is *grounded theory*, first described by Glaser and Strauss (1967). Grounded theory is the prototypical atomistic-logical procedure. It involves the following steps:

1. The problem on which research is to be conducted is identified at the beginning of the project, even if initially in a fairly rough, approximate way (e.g., problems of families with children suffering emotional disturbances or failure of drivers to use seatbelts). This is what I call the “object of study.” The sources from which ideas for defining the research focus (the object of study) can be obtained may be *formal*, such as previous research, or *subjective*, such as personal experience or intuitions. This topic was discussed in more detail on p. 197.
2. Data on the problem such as interview protocols or written narratives are collected with a number of appropriate people (e.g., parents of emotionally disturbed children, drivers fined for failure to fasten their seat belts).
3. The researcher breaks down the text into its building blocks by identifying content units.
4. The researcher forms groups of content units with similar content into categories or sub-themes (e.g., “guilt feelings,” or “aggression” (these categories really were identified in the study of parents of children with emotional disturbances).
5. On the basis of either a re-examination of the existing data or of the analysis of new data (or both), the researcher

clarifies the content of these categories and makes their definition more concrete in order to understand their content and their importance for the participants better. Glaser and Strauss call this process “saturation.”

6. After saturation of the categories the researcher builds more general overarching blocks of categories/sub-themes by combining related, specific categories/sub-themes (for instance, “extra-punitive vs. intra-punitive reaction” as a combination and generalization of the two categories/sub-themes “guilt feelings” and “aggression” given above in point 4). I refer to these overarching blocks of categories/sub-themes as “concepts” or “themes.” An example can be taken from a study of hobby authors which is described in greater detail later (see Chapter 9). In this study the categories/sub-themes “problems getting started,” “goals,” and “need for recognition” were combined to form the overarching concept/theme “Motivation.”⁵⁵

7. The researcher looks for connections among concepts/themes and formulates theories about such connections. For instance, in the hobby authors study just mentioned it was noted that there appeared to be a connection between the concepts/themes “Motivation” and “Interpersonal Relations,” with this link involving the facilitating or inhibiting effect of other people on productivity of authors.

8. The hypothesis formulated in step 7 can then be linked up with existing theories to yield a new or modified theory. In the hobby authors study, for instance, the researcher established links with Maslow’s hierarchical model of motivation, which distinguishes among other

⁵⁵ The terms “category” and “concept” are appropriate when the grouped content units (categories) or grouped groups of content units (concepts) refer to discipline-related phenomena such as “confirmation bias” or “group cohesiveness.” “Sub-theme” and “theme” are more appropriate when the groups of content units refer to practical, everyday life-related issues such as “living with a disability” or “reacting to mandatory vaccination.”

things between survival needs and needs for status and recognition, and concluded that the motivation to write satisfies needs at more than one level. To turn to other earlier examples, data on parents of emotionally disturbed children might be related to theory on locus of control, while data on migration could be related to theories of personality development.

Where step 8 is achieved, it is possible to speak of “grounded theory.” The theory formulated in steps 6 and 7 may then be tested by separate empirical studies, including quantitative studies.

Both analytic induction and grounded theory make use of the *method of constant comparisons*. Cases are analysed one at a time (starting of course with the first one), and as each new case becomes available it is analysed and its results compared with the existing analysis, possibly causing changes in this analysis, or even causing changes in the data collection itself. This process of case-by-case comparison is kept up until the addition of a new case leads to no changes. In quantitative research progressing in this way would be regarded as close to cheating: quantitative researchers would usually not commence the data analysis until all data had been collected, on the grounds that knowledge about how the analysis was proceeding could easily influence the new data. Suppose, for instance, that a researcher knew that it was important for the fate of the null hypothesis that the next case (“subject”) failed a particular test. Either consciously or unconsciously the researcher might influence the outcome of the test in order to obtain the desired result. However, in the constant comparative method phase of content analysis a case-by-case analysis is necessary.

Analytic induction and grounded theory – a comparison: The most striking difference between analytic induction and grounded theory is that analytic induction involves forming a theory on the basis of a general impression of the narrative being studied and seems to depend largely on intuition in the initial stages. It does not seem to be as purposeful, systematic

and organized as the grounded theory approach, mainly because it does not break down the narrative into specified, separate elements, as does grounded theory (i.e., analytic induction is more holistic). It may be difficult to explain in more or less objective terms what aspects of a narrative led to the explanatory hypothesis developed in analytic induction, or how clearly the hypothesis fits new data. In fact, the initial explanatory hypothesis may stem from the researcher's own expectations or subjective theories. By contrast, through the identification of content units grounded theory has a more objective basis than the more holistic analytic induction. The content units (building blocks) are almost like numerical data: they are discreet, clearly defined elements which are recognizable to any observer and are at least quasi-objective, with the result that people such as the readers of the final report usually find it easier to understand and accept the process of building categories and concepts. In other words, grounded theory leads to well-founded concepts that have a strong claim to reliability and validity. External observers usually have little difficulty in seeing how categories were formed or in seeing the logic of the way they are linked to obtain super-ordinate concepts. In other words, grounded theory leads to *grounded* concepts that seem to be more obviously *reliable* and *valid*, whereas those in analytic induction are more subjective.

In the later stages, analytic induction relies on negative cases for the refinement of the theory that has already been formed – the hypothesis already formed is only altered if a new case does not fit. This procedure follows the traditional approach of quantitative research, according to which a hypothesis can never be proved, only disproved. It is not the supporting case that aids the development of theory, but the negative one. This is a strength of analytic induction. If the cases studied – that have all been shown to fit the theory (otherwise it would have been altered) – are large in number or cover the range of cases that exists in real life (i.e., in quantitative terms if the sample is “representative”), a certain degree of “proof” of the generalizability of the theory exists. However, this point in some

ways contradicts the core of qualitative methods, and also raises issues of sampling that are more typical of quantitative methods. By contrast, in grounded theory the emphasis on systematic – even hierarchically organized – steps permits a more rigorous approach to analysing qualitative data. The process begins with specific, narrow concrete statements (content units) that are recognizable to any observer, and proceeds in a goal-directed way to the level of abstract and general formulations. It involves a systematic progression from the narrow and specific to the broad and general, and not just a vague “immersion” in the data in the hope of being inspired by them.

Bryman and Burgess (1994) discussed the extent to which grounded theory is actually applied in everyday research practice. Their comments are revealing and helpful for novice researchers, even bringing a sense of relief. To put it briefly, they express their doubt that grounded theory – and by implication analytic induction – are ever really applied in the strict way described above. Grounded theory’s importance is probably more because it reminds researchers of the need for an interplay between description and theory than because it is applied literally. It also involves a systematic even hierarchical set of steps commencing at the level of specific, narrow and concrete statements (referred to in this book as “categories/sub-themes”) and proceeding purposefully to the level of abstract and general formulations (“concepts/themes”), even if the steps are not followed exactly. In other words, it encourages a rigorous approach to the process of induction – movement from the small and particular to the broad and general – rather than a vague “soaking” in the data in the expectation of receiving inspiration.

Other Ways of Analysing Qualitative Data

Although they dominate analysis of qualitative data, analytic induction and grounded theory are not the only approaches that exist: There are other ways of working out the meaning of qualitative data. Eight of these are outlined in the following section.

The hermeneutic approach: The hermeneutic approach is most frequently used for the analysis of biographical narratives. Kromrey (2009, p. 300) compared it with the literary analysis of a text, an exercise that may well be familiar to readers in their memories of literature lessons at school. The purpose of a hermeneutic analysis is to grasp what the author of a text wanted to say. A hermeneutic analysis has four central characteristics. According to Madison (1988) it must display:

- *coherence*: the interpretation must be unified and free of contradictions or inconsistencies;
- *comprehensiveness*: the meaning of the text *as a whole* must be considered;
- *penetration*: “hidden” ideas that may have been expressed indirectly must also be considered;
- *contextuality* (the analysis should take account of the sociocultural context in which the narrative was constructed, and look for links to the situation beyond the text itself).

The first crucial point is that the totality of the text must be considered. Like the links of a chain, the parts of a text acquire their meaning through their place in the whole. A second principle that is particularly emphasized in the hermeneutic approach is the necessity for the researcher/interpreter to be aware of his or her own “biases.” These biases are essential in that they provide “fields of meaning” and allow an initial or preliminary understanding of

any new situation (Bleicher, 1980). (This initial understanding derives from researchers' constructions of their own experience, and is therefore, among other things, socially constructed.) However, it is the researchers' *awareness* of their own biases that enables them to distinguish between initial impressions and new meanings that arise from the text. This point is reminiscent of the view that interviewers should not try to eliminate social construction but should take account of it.

The hermeneutic emphasis on the analysis of the "parts" and the "whole" means that the interpreter of a text should pay particular attention to the narrative sequence as it develops during the course of narration. In other words, it is important to consider how what is described during the initial part of the narrative is related to what is described in later sections. The specific details of how this narrative sequence is considered varies from study to study, according to the needs of the particular research question. One example of the way the parts acquire additional meaning through their relationship to the whole is to be seen when discrepancies or contradictions occur from one part of the narrative to another.

As has just been emphasized, the hermeneutic approach requires considering the individual parts in order to understand the whole. However, the parts can be looked at in many different ways. In fact, psychologists who engage in hermeneutically oriented qualitative research often note that the process of such research does not follow a prescribed formula, is based upon the specifics of each research project, and is dependent upon the creative, imaginative and empathic resources of the researcher. Bruner (1990), for example, in presenting an analysis of texts from interviews with six members of the same family indicated that the process was based on principles of literary and discourse analysis (see also the reference to Kromrey above). Freeman (1993) similarly noted that the process of hermeneutically oriented qualitative analysis of autobiographical narratives (including those elicited during the course of an interview), involves various aspects of literary analysis.

Freeman explained his approach in the following way: “When I sit down with the text of a life history before me and try to make sense of what it is that’s being said – in line with the specific questions I bring to it – I had better be willing to exercise my own imagination” (Freeman, 1993, p. 225). The results of a hermeneutic analysis are thus, almost of necessity, strongly linked to the particular researcher, which raises the problem not only of reliability but also of both internal and external validity. This problem of qualitative data analysis will be discussed more fully in later sections.

Critical incidents analysis: Critical incidents analysis can be used with any of the kinds of data outlined in earlier sections, although biographical narratives are especially appropriate. The critical incidents approach identifies content units related to specific concrete incidents that interview partners regard as of particular significance in their lives. These must be incidents that the interview partners personally experienced in their lives, not for instance vague opinions or feelings. A critical incident can be a turning point in the respondent’s life in connection with family, health, loss, sexuality, intellectual or spiritual life, money, work, success or failure, and so on. As with all qualitative analyses, it is possible to start by defining in advance the kind of incident that is to be investigated, on the basis of existing theory or because of a particular interest in a pre-defined area. This could be done by searching protocols for incidents related to, let us say, school. However, it is also possible to allow the object of study to emerge in the course of the investigation.

The critical incidents approach can also be used as a method for *collecting* data. This can be done in an “open” way – in this case interview partners are simply asked to describe critical experiences in their lives in a general way. The interview could begin with: “Please tell me about good or bad experiences in your life.” It is also possible to conduct a structured interview, for instance by asking about a certain kind of critical incident. The Novalis study discussed on p. 138 is a good example of an analysis focused on a

specific kind of critical incident (death of a loved one), although the data were not interview protocols. The theme of death had already been specified prior to the commencement of the study, because it was already known that this topic played a special role in Novalis's poetry. The way Novalis dealt with the theme of death in his poems was shown to be linked to real deaths in his life. It is also conceivable that at the beginning of the project a critical incidents approach was not foreseen at all, and that the researcher only recognized the usefulness of this approach for understanding the object of study in the course of the investigation.

Comparative analysis: In a comparative analysis cases are directly compared with each other in order to identify differences between people in the way that a general way of constructing reality manifests itself. Contrasts between different cases may make special characteristics of the two cases more readily apparent. This can greatly facilitate the clarification of narratives and assist in the building of concepts or recognition of themes. Systematic differences between two (or more) interview partners in the way they construct reality may also indicate the existence of "polarities" which differ from participant to participant and reveal contrasting "orientation patterns." These can be linked to individual people (orientation figures): For instance two divorced women may construct the process of divorce in highly contrasting ways, the one as a happy release, the other as a personal failure. This contrast may help to understand the general process of building and dissolving personal relationships better (remember that the purpose of qualitative research is not to "diagnose" specific participants and recommend treatment, but to understand some more general phenomenon better – in this case divorce). Orientation patterns may also be linked to areas of life (for instance, an orientation to work as a source of principles for constructing reality vs. an orientation to home life).

Objective hermeneutics: Hermeneutics is based on working out the overall meaning of a narrative (see above). This is done by carrying out a "reading" of a narrative, the term "reading"

referring not to the technical process of perceiving written symbols on a page and turning them into words, as you are doing at this moment, but of assigning meaning in much the same way as we “read” another person’s behaviour. It is thus an intuitive-holistic procedure. In the case of objective hermeneutics, the analysis involves *a team of readers*, not a single researcher as is usually the case. Each team member initially develops his or her own “reading” of the first narrative. The team members then discuss the different readings and work out rules for how to understand the narrative, what to focus on, and so on. Members then read the next narrative with the benefit of the rules that have just been worked out, if necessary (as is usually the case) modifying the earlier readings, as well as modifying or adding to the rules of interpretation. A third narrative is then read and its implications for the initial alternative “readings” discussed in the group, possibly eliminating some, which are seen to be obviously false, and modifying others, while the rules of interpretation are also modified on the basis of the additional experience gained with the third case. This procedure (sequential analysis of narratives) is continued until consensus is reached that no more modifications of the readings are needed – there is, in effect, now only a single reading, on which the team agrees. Although it is essentially holistic-intuitive in nature, this approach has the advantage that the rules of analysis are worked out in a transparent manner and stated in a relatively objective way. It is clear that they are not just fantasies of the researcher.

Framework analysis: Framework analysis is an atomistic-logical procedure based on the identification of content units and the development of categories and concepts as already explained above. However, the broadest and most general categories (the “concepts”) are not psychological dimension such as “Motivation” or “Defence mechanisms,” but are themes or topics from a person’s life. In psychological research such themes might be “Joblessness,” “Emigration,” or “Raising a handicapped child.” Less general categories (subtopics) are then identified and assigned to

a specific theme: in the example of the theme of “Joblessness” one relevant subtopic might be “Failure,” another “Promotion policy.” Results of a framework analysis are typically presented in the form of a matrix for each theme (hence the label “framework”), in which the rows are individual participants and the columns subtopics. Each cell then contains content units defining the particular subtopic for a particular participant. It is then possible to examine each column in order to see how the subtopic in question emerges for different respondents and to identify patterns, contradictions, relationships, and so on.

Types and patterns analysis: An approach that has some overlaps with Framework Analysis involves identifying types and patterns. It is often possible to identify consistencies in a group of protocols which can be regarded as patterns or types. There are at least four forms of these:

- *Types of people* within the group of respondents,
- *Types of reaction* to the object of study,
- *Differing ways of dealing* with specific reactions,
- *Types of situation* which provoke a specific reaction.

In the airfreight case study already described, types could clearly be identified. For example, two types of loading personnel could be seen: those who were relatively satisfied with the level of competence of their managers and those who were clearly dissatisfied. The first type were almost all newcomers to the job, young, and less well trained, while the second type consisted of workers who were older, more experienced on the job, better qualified and strongly established with the company; on the one hand, beginners versus old hands. After identifying these two types, the analysis was able to focus on a new question: What were managers doing which drove away well-established employees but not beginners?

It was also possible to identify and focus on not types of person but types of negative reaction (e.g., aggression versus apathy). The aggressive type of reaction not infrequently involved

leaving the job out of the desire to “show them a thing or two.” The contrasting type of negative reaction involved a feeling of isolation and powerlessness, and resigning was an act of despair or burnout. Such contrasting reactions could also be classified into more general, “psychological” types such as *extra-punitive* (the managers are to blame) versus *intra-punitive* (I am to blame).

A final variant of types and patterns analysis involves identifying typical situations which are of special interest for understanding particular aspects of the situation being studied. For example, this might involve a typical situation which had had a particularly strong effect on a particular type of respondent or had evoked a specific kind of reaction. This approach made it possible, for example, to show that younger loading personnel were inclined to react in an intra-punitive way when they tried to avoid the early morning shift in the middle of winter. For instance, they felt guilt about reporting in sick on particularly unpleasant days even though there was nothing wrong with them, although this did not stop them from continuing the practice! Such insights provide practical advice on training of supervisors and counselling new employees.

Interpretative phenomenological analysis:

Phenomenology is concerned with how different events or states are experienced subjectively. It focuses on what experiences “feel like” to the individuals experiencing them. Interpretative Phenomenological Analysis (IPA) was developed specifically within psychology in order to use qualitative research methods to get at such feelings. It is now being used widely by researchers in health, clinical and social psychology: What does it feel like to be sick or depressed or lonely, for instance? This is, of course, a variant of the basic question in qualitative research: “How do people who are ill, depressed or lonely construct their situation?” What is special about IPA is that it *explores the individual’s personal experience of a state or an event*, rather than describing the state or experience itself.

IPA can be used as a clinical diagnostic tool focused on

a single case for the purpose of making recommendations for treatment, but for our purposes here it is better understood as a research tool for analysing a psychological phenomenon (see the examples above). In this case, as is customary, several cases are analysed one after the other and the phenomenological description is progressively refined, until the analysis of a new case produces no further refinement or only minor refinement, in which case the parsimony rule indicates that no further cases are needed. Typically, IPA specifies the object of study in advance (i.e., it is less emergent than some analytic induction or grounded theory approaches), and uses semi-structured interviews, in order to be sure that the respondent focuses on the particular phenomenon under investigation, while still allowing the researcher to probe interesting areas which emerge. In theory, other qualitative data collection methods can also be used, e.g. diaries, letters or other narratives.

Thematic analysis: Like all procedures for analysing narratives, thematic analysis emphasizes providing an understandable – but still systematic and rigorous – overview of people’s understanding of the phenomenon about which they are providing a narrative (Braun, & Clark, 2006). As with all qualitative analysis, its focus is on the grouping of individual content units into blocks of units with related content. This grouping process can be seen as involving two broad steps in moving from a specific person’s individual remarks about his or her understanding of the object of study to a more general understanding: (a) combining *individual content units* with closely related meaning to form blocks of related units followed by (b), combining blocks formed in (a) that turn out to have closely related meaning in order to form broader super-blocks. In earlier discussions I have used the term “category” as a label for the blocks formed in step (a) by applying processes of analytic induction, grounded theory or other meaning-identifying procedures (see Chapter 8), and the term “concept” as a label for the higher-order blocks of related “categories.” However, especially

in a discussion that has the goal of making “the *practical* aspects of qualitative research understandable and usable” (p. 6), labels that are closer to everyday ways of giving meaning to life are more practical, understandable, and useable. For this reason, I find the terms “sub-theme” and “theme” used by Braun and Clark and other writers very helpful as synonyms for “category” and “concept.”

Braun and Clark’s (2006) description of the phases of thematic analysis closely resembles the steps in building “categories” and “concepts” outlined above. However, the terms “sub-theme” and “theme” focus attention on understanding *the story* being told in a narrative, whereas the terms “category” and “concept” focus more on the specialist terminology of the researcher’s discipline. For a psychologist this might involve thinking propcesses used by respondents to fashion a coherent narrative (cognition), the aspects of their personality that modify these processes (attitudes, values, feelings, etc), or the forces driving them (motivation, emotions, etc). Thus, whereas “category” and “concept” are usually closely related to the way *researchers* reflect upon respondents’ narratives, sub-theme” and “theme” are more expressive of the way respondents sort out their understandings when they reflect upon their experiences with the world. For this reason, the two pairs of terms are more or less interchangeable, depending on the purpose of the analysis. In the example of parents trying to make sense of their child’s disability (see pp. 72-73), for example, thematic analysis with a focus on *themes* like “What is the nature of the disability?” “Why did it occur?” “Who or what caused it?” “What consequences will it have for the child?” “Is it treatable?” “Who will help us and how will they do it?” provides a concrete and down-to-earth understanding of how the parents gave meaning to their situation in their own minds. By contrast, in the psychological study of young women who had lost their fathers (p. 168), *concepts* like “extra-punitive reaction” versus “intra-punitive reaction” are more helpful for a more psychologically oriented analysis.⁵⁶

56 At a very down-to-earth level it could be said that “categories” and “concepts”

A Concrete Example

In order to make the whole procedure more concrete, I present in the following section the summarised and clarified text of an interview with a respondent who took part in a study of social workers' reasons for choosing to work with refugees and their psychological reactions to the job (Cropley, 2002). Although the example is based on an actual study it is presented here in a heavily modified form to make it suitable for the purposes of the present book. The first passage contains the social worker's remarks before summarization and clarification.

[The interviewer asks about reasons for taking up social work with refugees.]

Well, of course, I ah, I mean everybody is the same, well, I wanted a job, and this one they ah offered it to me because it was you know it was available. It seemed OK, ah when I thought a bit I thought more or less it was just right when you think about it for me. I ah suppose I have to say that I, you know since I was young, have always been fascinated by the like exotic or what you call unusual. I wanted to see more, like experience what they call new horizons, broaden them you could say. That's me, I guess that's the sort of person I am, or any way I think that that's what I am. I'm social, you know I like to be with others, ah a social like creature too. I don't have any ah problems you could say that I get along well with people, most people I think ah even you know the kind that most people don't like that kind of people, so that I wasn't

summarize the way that a particular scholarly discipline understands the object of study, whereas the terms "sub-theme" and "theme" summarize the way the people providing the narratives understand it. Returning to the example on p. 72, a theme might be "Whose fault is my child's disability?" whereas a concept might be "intra-punitive versus extra-punitive reaction."

worried about, I didn't have any anxiety about ah actually I thought it was ah good to work with strangers. It wasn't just that, though. I mean it's good when some-one has problems to be there, I ah feel OK when I help some-one. Especially when, mostly when, they have you know problems. But I am a person who likes, well I don't want some-one telling me "Do this, Do that!" I like to ah decide myself. A bit OK, but not telling me like all the time.

As can be seen at once, the passage contains many *meaningless* elements that had to be removed in the process of *summarizing* and *clarifying* (see above). After summarizing and clarifying, the passage was much easier to understand, as can be seen from the passage that follows. The significance of the sections of the text that are now written in bold face will be explained shortly:

[The interviewer asks about reasons for taking up social work with refugees.]

Well, of course, I wanted a job, and this one was available. It seemed just right for me. **I have always been fascinated by the exotic** or unusual. **I wanted to broaden my horizons.** That's the sort of person I am. I'm a social creature too. **I get along well with people**, even difficult people, so that **working with strangers interested me.** It wasn't just that, though. **I like to help people in trouble.** However, I'm an independent type. **I want to be able to work without someone telling me what to do** all the time.

The summarised and clarified version of the text can be compared with the original form given above. This comparison makes clear the value of these steps before commencing the actual process of developing content units. I turn now to this process. The passage above and the two boxes that follow are all part of the data collected

in the social worker project. The researcher's object of study was social workers' reasons for entering their particular branch of work and the psychological problems they encounter there, and the identification of content units therefore focused on these areas. The groups of words written in boldface, both in the passage above and in those that follow, are the content units identified as relevant to the object of study.

[The interviewer asks: "Are you satisfied with your job?"]

What worries me though is that **I want to do my job better**. I can see that I need to learn more about people. You learn a lot on the job and that pleases me, but **I need more academic training**. Otherwise, I feel that **I am not really good enough** as a social worker. Conditions in the hostel aren't good. We don't get enough financing and the politicians don't care. **The clients are often very difficult** too. They quarrel among themselves and some try to get more than the others. **I don't know if I am really helping them**. They don't seem to make much progress. **Why can't they avoid trouble themselves?** Like when they don't pay their fare on the train. Then they try to get me to do what they want. **I'm sick of them trying to control me**. When I talk like this **I feel guilty**, because **I feel disloyal to my clients**.

Statements such as, "Well, of course, I wanted to earn money, and this job was available," were disregarded because they were not specific (wanting to earn money refers to any job, and not specifically to social work with refugees). In a similar way, the statement "We don't get enough financing and the politicians don't care," refers to a difficulty of the job, to be sure, but it is not a statement about the respondent's psychological reaction to the job (i.e. it is irrelevant to the object of study). The same is true of "Conditions in the hostel aren't good," and "They don't seem to make much progress." The

content units identified as relevant (i.e., those in boldface) were subsequently related to psychological categories. The results of the analysis of this case (the first one analysed in the study in question) are shown in Table 8.1 (see below). Tentative concepts defined by the “categories” were also identified, and these are also shown in Table 8.1. The contents of this table can be regarded as representing the completion of step 5 in the description of grounded theory given on pp. 219ff.

[The interviewer asks how the respondent feels about the interview.] Well, it has helped me to do something for myself for a change. It gave me a chance to overcome my isolation and get new ideas about my job. I need such conversations in order to understand the point of my work. Otherwise it just seems to be doing the same thing over and over again. I feel so helpless. I often ask myself, “Do you really need this job?” I am too involved and it wears me down.

Table 8.1: Content units, categories (sub-themes) and concepts (themes) identified in the social-work interview

Content Units (specific, concrete statements made by participants and couched in everyday language)	Categories/Sub-Themes (more general, more abstract summaries made by the researcher and couched in fairly scientific language)	Concepts/Themes (abstract, scientific labels related to theory)

to work without supervision to help people in trouble	desire for independence helper syndrome	MOTIVATION
to broaden my horizons working with strangers I get along with difficult people	openness to novelty desire for novelty tolerance for uncertainty feeling of special personal qualification	PERSONALITY
I want to know more about people I need more academic training I feel guilty I feel disloyal to my clients	self-doubt need for greater knowledge feeling of personal deficiency negative affect about herself	SELF-EFFICACY
I want to do my job well I don't know if I am really helping them The clients are often very difficult Why can't they avoid trouble themselves I'm sick of them trying to control me	success orientation uncertainty negative feelings towards clients feeling of being manipulated	JOB-RELATED STRESS
to understand the point of the work to overcome my isolation (the job) wears me down Do I really need this job? I feel helpless	sense of meaninglessness isolation sense of personal destruction sense of meaninglessness loss of control	BURNOUT

Completing a Qualitative Analysis

The typical result of a qualitative content analysis conducted as outlined above is a set of concepts or themes (such as “Motivation,” “Personality characteristics,” “Self-doubt,” “Job-related stress,” and “Burnout syndrome”). These do not stand alone, however. They *summarize* in psychological terms the way the participants in the interviews conceptualized the phenomenon under discussion. In the example in Table 8.1 this involved social workers in the specialized area of refugees. Also available to the researcher are the original interview texts (i.e., the full interview context) as well as informal information such as emphasis, tone of voice, body language and observable nonverbal behaviour such as gestures or even smiles, crying and the like. The researcher’s task is now to

- describe the way in which the object of study is understood in the particular group of participants with whom the interviews were conducted, *using the concepts worked out in the content analysis as a framework*
- interpret the relationship of the participants to the object of study, *using the concepts worked out in the content analysis as a framework*. Occasionally, the purpose of the study is to examine the way the concepts apply to particular individuals rather than to particular situations. In this case, a special phase is necessary.
- generalize the interpretation just mentioned (i.e., develop an understanding of the object of study that goes beyond the particular individuals studied). This is the first level of *development of theory*
- formalize the generalization not just to other people, but to relevant psychological and educational theory
- apply the results of these steps to concrete, practical problems (i.e., make the analysis *useful*).

In the actual study on which the table is based, analysis of further interviews yielded many more content units that fitted well into the existing framework. The analysis then moved on to use the concepts just identified in order to develop a general description of the social workers in the study, i.e., a *substantive analysis* (see p. 258) was carried out. This analysis also made use of supplementary information such as repetition of statements, raising of the voice, or various forms of body language. The analysis finally went beyond description of this particular group of respondents in order to develop a general, psychological description of problems and difficulties of social workers working with refugees, i.e., a “formal” analysis was conducted. How to do this in a truthful, confirmable and useful way is the subject of Chapter 9.

Chapter 9: Requirements for a rigorous qualitative report

The task of a research report is to *communicate* the results of a scholarly investigation to readers (i.e., the report is at the heart of the phases of communication and validation – see Table 4.1). This requires more than a literary or journalistic report, even though the study being presented may have involved descriptions of everyday experience couched in non-scientific language. A research report must (a) deal with clearly stated *scientific* issues and state them in a formal, scientific way, (b) describe in detail the way in which information was obtained and analyzed, and (c) leave readers knowing more than they did before they read it. Thus, simply recording and reporting narratives is not enough, even if the report is organized into an interesting, structured account. It must contain some “surprise.” However, the kind of surprise that is needed does

not result simply from departing from the usual or expected in an undisciplined way: A research report must be a *scholarly* document displaying *scientific rigour*: what I call scientific “artisanship.” In this chapter I deal with the question of how to report qualitative research in a way that expands existing knowledge, while at the same time meeting the required standard of scholarship and rigour necessary for a scientific report (i.e., I will look at how to deal with problems such as the apparent *arbitrariness*, lack of *proof*, and *triviality* of qualitative research). Although there are no set procedures for achieving scholarliness, general principles can be laid down, and this will be done here.

Chapter Glossary

<i>artisanship</i> :	specialized skill in carrying out a demanding practical task that leads to an effective and pleasing product
<i>causal conclusions</i> :	conclusions that use the information obtained in an analysis to infer cause-and-effect relationships (typical of quantitative research)
<i>effective surprise</i> :	novel (i.e., surprising) ideas that are plausible, relevant, and useful (i.e., effective)
<i>flexibility</i> :	a quality of research that allows the object of study, sampling, data collection, and the like to be changed in the course of an investigation. Until recently this was referred to as “openness,” but this term now

	has a different meaning (see “open research” below)
<i>functional conclusions:</i>	conclusions that involve generating a new theory about the underlying factors that lead to different content units relating to each other
<i>genetic conclusions:</i>	conclusions that involve transforming existing theories
<i>higher order theory:</i>	theory that enlarges understanding of a field of study (also referred to as “formal” theory)
<i>open research:</i>	research that is conducted in accordance with strict and exact rules of procedure that are specified in detail in advance and are thus “open” in the new sense of the word
<i>open research reporting:</i>	reporting of research that reveals everything that was done in a study in concrete terms (sometimes referred to as “confessional” reporting)
<i>rigour:</i>	a combination of adherence to rules of good practice, thoroughness, precision, and accuracy
<i>propelling a field:</i>	changing the way people conceptualize a phenomenon or area of study
<i>pseudo-creativity:</i>	ideas that are unexpected but not novel
<i>quasi-creativity:</i>	ideas that are novel but have no link to the real world
<i>substantive theory:</i>	theory that enlarges understanding of a specific object of study and/or group of participants

<i>science communication:</i>	the study of factors that facilitate/ hinder popular understanding of research findings
<i>transparency:</i>	a property of a research report that makes what was done, how, and why clearly understandable
<i>usefulness:</i>	propelling a field in a way that can be applied in practice

Criteria of a Rigorous Report

There is much talk nowadays of “open science” (e.g., Open Science Collaboration, 2015). The related discussion emphasizes above all the admirable goal of making research findings replicable (reliable). This is not merely an abstract academic issue: As the German Psychological Society put it (DGfP, 2017), critical discourse based on *informed* opinion is the indispensable foundation of democracy. According to the DGfP, a failure of research to provide readily understandable information that stands up to scrutiny threatens “uncertainty and mistrust” of science. Edwards and Roy (2017) summarized several studies which identified a “crisis of credibility” and warned of the “devastating consequences to humanity” if public trust in science were lost. Ibaviosa and Chin (2022) gave a specific concrete example of this endangerment in a domain far removed from research in education and psychology: They discussed the massive problems public distrust of research findings in forensic science would bring for the whole justice system. For example, loss of trust in research on DNA and fingerprints would threaten the very notion of what constitutes proof. Bromme and Kienhuis (2017, p. 167) even suggested a new scholarly discipline which they called “science communication.” Mixed research (see Chapter 6) raises special problems in this regard. It is based on what are far from

the “coherent and unified systems of thought” that underly “open” research. Ghiara (2020, n.p.n) pointed out that:

Members of the same community share some key ideas about how research should be conducted and recognize that they belong to the same community. However, there can be often deep divisions among members regarding the ontological, epistemological, and axiological assumptions guiding research.

In fact, much of this book, especially the previous chapter and this one, can be seen as involving a discussion of aspects of science communication such as transparency, credibility, and the like in qualitative research. At the level of research reports, the need for transparent presentation of all phases of investigations is emphasized. All aspects of the research process need to be clearly documented, and reports written in a way that makes plain what was done and why, what happened, and what it all means.

Different understandings of “openness”: Earlier in this book I repeatedly emphasized the “openness” of qualitative methods, using the term to refer to emergent properties such as changing the object of study part way through an investigation, changing the data-collection method, developing hypotheses after data have been collected and analysed (now known as “HARKing”), and similar actions. However, in very recent quantitative research “openness” has come to mean adhering closely to strict methodological rules. If all researchers always go about all research in precisely the same way, what was done in a particular study will always be clear, so that the research will be “open.” As of September 1, 2022 more than 300 journals require that the exact procedure of *planned* investigations must be registered *before* data collection is even started (for up-to-date figures, see the website: <https://cos.io/rr>), so that projects can be approved in advance as methodologically correct (or not). Some journals now guarantee to publish reports even if they yield no significant findings, provided that the methodology was pre-certified as correct. Thus, in the

recent “open science” discussion, the term “openness” means more or less the opposite of what I have used it to mean so far in this book. Because of the conflicting usage of the term, from now on I will refer to the *flexibility* of qualitative research. The issue with which this chapter is concerned then becomes the question of what a report should contain in order to be rigorous, despite the “flexibility” of the procedures it is reporting.

The next box demonstrates the flexibility of qualitative research. It summarizes a relatively “exotic” study in which (a) the researchers analyzed existing data, (b) participants were observed carrying out a specific action in which they were already well-versed (designing clothes), (c) participants were unaware that they were participants in a research project, (d) there was no direct contact between researchers and participants, and (e) raters carried out the data analysis after they had already been “sensitized” by an exploratory study that had already been carried out by completely different researchers.

Hogan, Murdock, Hamill, Lanzara and Winner (2018) investigated the object of interest “ways of thinking in fashion designers” by analyzing footage from the reality-TV show *Project Runway*. Their “sample” consisted of a complete season of the popular program. Their analysis of the behaviour they observed was pre-organized by an earlier study by completely different researchers which had identified eight “habits of mind” in design thinking in art classrooms (e.g., Hetland, Winner, Veenema, & Sheridan, 2013) and used these habits to structure their analysis of the videos of the episodes. A codebook was developed by a team of four expert raters who closely observed two episodes (see for example the discussion of development of a codebook on p. 201ff. and the discussion of team coding on p. 228). Subsequently, all verbal statements by all participants in all episodes were transcribed to form the customary written texts which were

analyzed by isolating content units and assigning these to what I call “categories” and “concepts” or “themes” and “sub-themes” according to the codebook (Hogan and colleagues referred to this as the “coding manual”). The process of coding involved a variant of objective hermeneutics (see p. 228).

With so much flexibility, what is it then that makes a scholarly report *rigorous*? Although they did not specifically use the term “rigour,” Altheide and Johnson (1998) suggested four criteria that I will apply here: *plausibility, relevance, credibility, importance*. To these Kvale (1995) added an important fifth criterion – “beauty.” By this he meant, in particular, internal logic, clarity and elegance of expression. These properties are summarized in Table 9.1. It is important to note that formal correctness is not mentioned in the list. This is because there is no simple, clear cut right or wrong way to do qualitative research (because it is *flexible*) or to write the corresponding report. In effect, *researchers can do what they like, providing that they describe it clearly and precisely* (i.e., “confess”). This does not mean, however, that readers are forced to accept whatever the researcher chooses to do. They can judge for themselves whether a study was rigorous or not, provided that the report contains the necessary information. Thus, researchers must make it clear that decisions made during the course of the investigation were based on specifiable systematic rules and were not arbitrary or *ad hoc*. As a result, the researcher’s personal *credibility* is of great importance, as Kvale (1995) emphasized. Unlike quantitative research, in the case of qualitative research credibility does not derive more or less from the formal correctness of the artisanship involved in applying fixed research paradigms, but from the convincingness of clearly and fully-stated scientific thinking, information gathering, and reporting, and establishing this credibility is one of the central tasks of a qualitative research

Table 9.1: Criteria of rigour
in a qualitative report

Property	Requirement
Plausibility	The report should be rationally related to existing scientific knowledge
Relevance	The report should link the topic to people’s experience of life
Credibility	The report should contain interpretations that are well founded, logical and believable
Importance	The research should generate ideas for further research and also for practice
Beauty	The report should be pleasing to read: clearly written, well-constructed, and well argued

Achieving rigour: I turn now to the question of how to fulfil these criteria. What properties of a report make its contents rigorous in the qualitative sense just outlined (i.e., plausible, credible, relevant, important and beautiful)? Building on suggestions of Hamel (1993), I divide a report into content areas: in order to be rigorous these must have the properties shown in Table 9.2.

⁵⁷ I have already discussed the importance of a codebook for this purpose in Chapter 7.

Table 9.2. The sections of a report and their key properties

Content Area	Property
Specifying the object of study	What the report is about must be clearly stated in both everyday and scientific language
Selecting the cases	The participants' demographic properties, how they were recruited, and how they are "representative" must be fully described
Describing the procedures	This section must describe what was done and how data were recorded
Analyzing the data	The steps must be described and examples given that make the steps understandable
Interpreting and generalizing	As much as possible, the report should expand existing knowledge (i.e. achieve surprise).

The "properties" just listed require more detailed explanation and this is provided in the present section:

- *The specification of the object of study*: A high quality report makes clear from the very beginning what question or issue is being examined. This can initially be stated in everyday language, but must be specified in scholarly terms in the course of the report.
- *The selection of the cases*: Well-chosen cases "represent" the object of study in a way that is obvious and compelling. The presentation of the case(s) in the report must (a) specify who or what they were (for example, demographic data), (b) explain why they were suitable for the particular

object of study, and (c) show how access to them was obtained.

- *The description of the procedure:* This must explain what technique was used to obtain data (e.g., participant observation, semi-structured interviews, examination of existing documents, etc) and under what circumstances, what preliminary steps occurred (e.g., a small number of pilot interviews), what aids were developed (e.g., rating schedule, interview plan, “prompts”), and how data were recorded (e.g., notes during an interview, video recording, etc). The description must contain sufficient concrete detail for other researchers to be able to repeat the study in a form closely resembling the original research. It must be clear, logical, well-organized and complete. It must distinguish clearly between description of the method used to obtain the data and details of the data themselves. Where the researcher made decisions such as concentrating on certain material and ignoring other material, the rules underlying these decisions must be explained so that it is apparent that they were objective and consistent. Examples are very helpful.

- *The description of the data analysis:* In particular, a rigorous qualitative analysis is systematic, logical, understandable, plausible, and convincing. Because no set formulae exist, the various steps of the analysis (see Chapter 7) must be outlined and made understandable with the help of concrete examples. It is important to distinguish clearly between data and interpretations of them.

- *The interpretation and generalization of the findings:* Conclusions can be of two kinds: “substantive” (relating to the particular case(s) in the present study) or “theoretical”/“higher order” (relating to the object of study in general). They should expand existing knowledge

(“create surprise” or “achieve originality” – see pp. 253ff.). They should be *plausible* (i.e., consistent with the evidence delivered by the respondents), *credible* (believable) and *consistent* (internally logical). Also important is a distinction between a scientific interpretation of data and personal opinions, political views, or even hopes or wishes about what the researcher thinks the world should be like. The author’s private opinions are perfectly acceptable in a scholarly report, but they must be identified as what they are and not presented as though they were strictly derived from the evidence obtained in the study.

Usefulness: To be externally valid a research report should be, above all, “useful” (Kvale, 1995): Findings should be relevant to real life and capable of advancing “sensible discussion” (Cronbach, 1980, p. 107). In order to be useful the report must be (a) *informative* (it must deliver an overview of the object of study), (b) *broad* (it must bring together a range of material), and (c) *practical* (it should offer ideas that can be applied to real life). However, the aim of useful research is to go further – it should advance thinking about the object of study. To do this it must be (d) *original* (explanations or interpretations should offer something new) and (e) *seminal* (explanations should show the way to or provoke further discussion and stimulate formulation of theory or offer new ideas on how to investigate the object of study more intensively). The criteria of usefulness are summarized in Table 9.3.

Table 9.3. Criteria of the usefulness (external validity) of a research report

Criterion	Explanation
Informative	The report must provide scholarly and scientific information about the object of study

Broad	It must offer a range of information and not be too narrow
Practical	It should offer generalizations that can be applied in real life
Original	It should go beyond common sense and offer new ideas
Seminal	It should suggest new hypotheses about the object of study, make suggestions for further research, and/or suggest novel ways of investigating the object of study

As with other properties and criteria such as those listed in Tables 9.1 and 9.2, these criteria form a package that is ultimately judged on its overall quality. This means that not all dimensions will necessarily be equally strong in a particular report: A report may have weaknesses in some areas, but still be of high overall quality (or, of course, not be), because particular strengths may compensate in part or in full for specific weaknesses. However, it must be borne in mind that glaring weaknesses in one area can also undo good work in another.

Achieving Originality

The last two criteria of usefulness listed in Table 9.3 (originality and seminality) differ from the first three in that they cannot be achieved simply by hard work and knowledge. These are necessary, it is true, but are not sufficient. Achieving originality and seminality requires more, since the goal is to add something new to knowledge. In the case of quantitative research, this is usually regarded as having occurred when a null hypothesis is rejected. From time to time some quantitative researchers argue that the *failure* to reject the null hypothesis can also expand knowledge, for instance in studies where there are overwhelming reasons for expecting it to be

rejected, but it is not: In terms of the following section this would be an example of a “surprising” result. However, very little quantitative research is published in which the null hypothesis was not rejected. Rejection of the null hypothesis permits generalisation of existing knowledge to new populations and/or new settings (deduction). In the case of qualitative research, by contrast, generalisation takes on a different form.

Generating surprise: To repeat Elliott’s (1999) comment mentioned earlier, qualitative researchers hope *to discover something new and unexpected*. Thus, a research report must seek to generate *surprise*. This means that it must go beyond common sense and develop an explanation of the data that may very well transcend explanations offered by individual respondents. Furthermore, it must do this according to the principles of a scholarly discipline (such as psychology or sociology), which are probably not obvious to the respondents, and in so doing expand and generalize their constructions. The scholarly conclusions may contradict individual respondents’ hunches, as these are based on a single case (themselves) and on pre-scientific ways of analysing and drawing conclusions.⁵⁸ Ideally, the interpretation of the data would *surprise the researcher too* by revealing an internal structure of the object of study that had previously been hidden. By doing so, a report acquires *surprise value*: Without this it simply repeats or confirms what we already knew or at the very least expected (hypothesized) to be true.

It is important, however, to remember that achieving surprise *in the scholarly sense* (as against for example in a movie or a circus performance) is always subject to the tenet of rigour (system, logic, understandability, plausibility, and convincingness). Naturally, it is easy to achieve surprise by wild speculation, but this is not what is meant here. Surprisingness without rigour involves what Cropley (2001, p. 14) called “pseudocreativity” (creating surprise simply

⁵⁸ See the discussion of “scientific” and “non-scientific” ways of understanding information on pp. 42-43.

by means of unexpectedness: for example belching loudly during a funeral service), or at best “quasi-creativity” (genuinely novel ideas that are completely unrelated to reality: for example, proposing to solve parking problems in large cities by opening up car parks on the moon). What is required is surprise, to be sure, but surprise that is “effective.”

Propelling the field: Effective surprise “propels a field” (Sternberg, Kaufman & Pretz, 2000). Although Sternberg and colleagues were discussing a whole area of knowledge (more like what I refer to as a “formal” object of study – see for example pp. 104ff.), for our purposes here a “field” is either formal (an entire scientific discipline or sub-discipline such as the psychology of interpersonal relationships) or substantive (a specific setting, process, product, etc, such as the way cargo loaders at a major world airport perceived their managers). According to Sternberg, Kaufman and Pretz, there are seven ways of doing this:

1. *conceptual replication* (transferring what already exists more or less unchanged to a new field);
2. *redefinition* (seeing the known in a new way);
3. *forward incrementation* (taking the known further in an existing direction);
4. *advance forward incrementation* (extending the known in an existing direction, as above, but going beyond what is currently accepted);
5. *redirection* (extending the known in a new direction);
6. *reconstruction and redirection* (returning to an approach previously abandoned and breathing new life into it);
7. *Re-initiation* (beginning at a radically different point from the current one and taking off in a new direction).

These ways of propelling a field can be thought of as forming a hierarchy. At the lowest level it is simply a matter of transferring what is already known to a new field. Successive levels involve greater production of novelty, but usually by adapting, extending,

refining, or redefining what already exists. Indeed, Miller (2000) concluded that *seeing the already known in a new light* is the core of creativity.⁵⁹ Only the highest level involves creating something entirely new. In other words, achieving effective surprise does not normally require the development of previously entirely unknown ideas in a burst of inspiration, but “merely” the clarification, refinement or extension of existing theory. (This is not to suggest that inspiration is to be ignored: The intention here is to guard against exaggerated expectations.) Sternberg, Kaufman and Pretz’s (2002) seven criteria can all be applied to the question of how to expand knowledge in qualitative studies. They can serve both as guidelines on how to construct original, surprising interpretations and also as a framework for self-assessment of degree of originality.

Generalizing: The implications of these ideas for the process of generalisation in qualitative research are summarized in Table 9.4. It must be admitted that the characterization of quantitative methods in this table is an oversimplification, but it is adopted here in order to create a clearly recognizable contrast. Put briefly, quantitative research applies existing general knowledge (in science, often summarized in a theory) to make sense of new information (deductive thinking), whereas qualitative research combines existing fragments of knowledge (content units) to formulate new, more general knowledge, such as a theory (i.e., inductive thinking). The “steps” listed in Table 9.4 as characterizing qualitative research are to be regarded as idealized goals – the task of the researcher is to go as far as possible towards fulfilling them. The way in which this is done in the expanding of theory is discussed in the following section. As has already been mentioned, the various aspects of creating surprise can receive different weight in a particular report, while strengths in one aspect (e.g., “development of new hypotheses” or “working out new research questions,” and so on) can compensate for weaknesses in another.

59 In scholarly writing, the simplest variant of this involves giving known results, effects, or phenomena a new name!

Table 9.4. Generalization in quantitative and qualitative research

Kind of research	Steps in generalizing
Quantitative	• Transfer of <i>existing</i> knowledge about the object of study to new settings or populations through testing of hypotheses formulated before the study commenced
Qualitative	• Creating surprise: Differentiation, extension or refinement of existing theory. • Development of new hypotheses. • Working out of new research questions. • Suggesting novel approaches to conducting research on the object of study

Judging the extent to which a report generates novelty: D. H. Cropley and Cropley (2016) developed an instrument that can be used to measure the extent to which a report drives a field and the way it does this. They call this instrument the *Creative Solutions Diagnosis Scale* (CSDS). Products are referred to as “solutions” in order to make it plain that any kind of product is a solution to some kind of problem. In the present case, a research report is a “solution” to the problem of how to present research in a way that is relevant, plausible, convincing, informative, and even “beautiful”). The scale assesses a solution on five dimensions:

- Relevance and Effectiveness
- Problematization
- Propulsion
- Elegance
- Genesis

These competencies are broken down into specific properties of the report (“indicators”) that make it possible to identify the presence or absence of the characteristics underlying the competencies.

The indicators (i.e., recognizable specific properties) include aspects of Relevance and Effectiveness (such as familiarity with the already known, accurate presentation of this knowledge, relevance of the presentation of the material for the object of investigation), Problematization (e.g. problem identification, prognosis), Propulsion (e.g. redefinition, redirection), Elegance (e.g. convincingness, pleasingness) and Genesis (e.g. seminality, germinality). The indicators can be regarded as characteristics of a research report that allow a “diagnosis” of its strengths and weaknesses in terms of advancing the field. For example, one research report could be strong in terms of problem detection but offer no new approaches, whereas another might reveal vast knowledge of an area but show little insight into the shortcomings of what already exists.⁶⁰ In a brief discussion of the CSDS, Kershaw, Clifford, Khatib, and El-Nasan (2022) mention studies confirming that the scale has high reliability and good factorial validity, although there is sufficient overlap between propulsion and genesis to suggest that the scale only measures four dimensions which I refer to as Relevance and Effectiveness, Problematization, Transfer, and Elegance.

Advancing Theory

A key element in achieving external validity in qualitative studies is the enlargement of knowledge (see Table 9.4 above). Very early in discussions of qualitative methodology, Znaniecki (see Jorgenson, 1989, p. 112) discussed the way in which qualitative studies can do this. According to him, it always involves drawing general conclusions about “components of a theoretical system,” i.e., it always involves theorizing. The theoretical conclusions can be

60 The CSDS can be used by students to analyze their own research reports. See <https://play.google.com/store/apps/details?id=com.csdsrubrik.id&hl=en>

“functional,” “genetic,” or “causal” in nature: *Functional conclusions* involve theory about the way elements relate to each other and are interdependent, *genetic conclusions* transform existing systems or even work out new systems, and *causal conclusions* use the information obtained in a content analysis to infer cause-and-effect relationships. The latter resemble most closely the typical result of quantitative studies. Despite this, all three kinds of explanation are possible in qualitative studies too, although it should be noted that the criteria of what is a proof are derived from the philosophy of truth, not inferential statistics (e.g., Kvale, 1996).

Two kinds of theory: Strauss and Corbin (1998, p. 161) made a differentiation that is very useful for the present discussion: They distinguished between “substantive” and “higher order” theory. The former focuses on theorizing about the object of study and the particular group of participants involved in the study in question. It can involve (a) interactions among people (cases) in a particular, specified setting, or (b) processes involved in a particular phenomenon (e.g., movement through phases, stages and the like, or relationships between changes in conditions in one area and changes in another). However, a scientific research report in psychology and education should go beyond substantive description of a particular object of study in a concrete situation and seek to advance knowledge about other situations too, i.e., to generalize. This is the phenomenon of “formal” or “higher order” theorizing. Higher order theory is more abstract and general. It not only advances knowledge about the particular case(s) studied in the specific research project, but it also enlarges the general body of educational and psychological knowledge. As Strauss and Corbin pointed out, this kind of theorizing is more common in quantitative studies than in qualitative ones, although this does not mean that it is impossible in qualitative research.

Criteria for evaluating advances in theory: Kvale (1995) emphasized that in qualitative research the worth of theorizing is measured by its “usefulness.” As Cronbach (1980,

p. 107) put it, this means that the study's conclusions "advance sensible discussion." Katz (1983) made a similar point in defining usefulness: It lies not in establishing absolute truths but in *advancing knowledge*. Judging how much a particular study has achieved is thus not a matter of what universal laws it "proved" but of how much it advanced knowledge. The methods of analysis outlined in Chapter 7 can then be seen as strategies for using data to advance knowledge. This can be done by

- altering ways of thinking about the object of study
- generating novel hypotheses about the object of study
- giving new impetus to research on the object of study, for instance by provoking novel lines of inquiry
- suggesting novel ways for looking at either existing lines of inquiry or new ones.

These are the basic criteria that should be applied to judging the usefulness of a qualitative study for advancing knowledge. They can be stated in the form of questions:

1. Does the study help us to see a psychological/educational issue in a new light?
2. Does it suggest new hypotheses about the issue?
3. Does it indicate what further research is needed on the issue, or open up novel lines of inquiry?
4. Does it suggest novel kinds of research?

Chapter 10: Rigorous, accessible, and transparent reporting

Qualitative research methods breach numerous rules of traditional procedure, such as forming the hypothesis only after the data have been collected, treating intuitions or hunches as knowledge, or allowing the researcher to specify what data mean (as against leaving this task to objective tools such as test manuals, instruments, and the like). Since procedures for collecting and analysing qualitative data are not highly standardized, in qualitative research reports these must be *clearly described* to readers with frankness and completeness. Thus, it is essential that qualitative reports make clear (a) what was done, (b) how it was done, and (c) why it was done in this particular way. Sometimes such reporting becomes almost a form of confession. The examples from research reports outlined in this chapter show how to report subjective aspects

of planning, implementation, analysis and reporting of a qualitative study in a flexible but transparent manner.

Chapter Glossary

<i>artisanship:</i>	working according to set procedures which are repeated in a highly standardized and highly skilled way
<i>accessible reporting:</i>	research reporting that it is easy for readers to understand
<i>confession in reporting:</i>	reporting that does not hide weaknesses, problems, compromises, and the like
<i>exploratory interviews:</i>	interviews conducted at the very beginning of the data collection in order to get a feel for what is likely to happen and to work out how to encourage respondents, for example through prompts
<i>explanatory interviews:</i>	interviews conducted late in the data collection in order to understand the practical implications of emerging information
<i>Milli-Vanilli effect:</i>	carrying out highly routinized behaviours in order to create the impression of behaving skilfully
<i>rigour:</i>	a quality of a report that reveals exactly what was done and enhances its methodological reliability and validity
<i>transparent reporting:</i>	reporting that presents all necessary

information in a convincing way

Writing a report

Quantitative research is carried out in accordance with a strict paradigm. In fact, the rise of “open science,” to which I have already referred in Chapter 9 and elsewhere, has increased this strictness, as a glance at recent research reports in highly respected journals shows: For example, in some journals nowadays even the abstracts have set sections marked off by headings, in order to standardize their thematic content and organization. I have already referred to the situation as involving “artisanship,” according to which highly refined set procedures are repeated in a skilful way. An example of how absurd blind adherence to the rules of good form can become can be seen in a recently published paper which included an introductory section of almost 500 words dealing with headings such as “Author contributions,” “Funding,” “University ethical statement,” “Data availability statement,” and “Conflicts of interest.” The height of absurdity was reached with a 4-line sentence explaining that the two authors had contributed equally to conceptualization, methodology, validation, formal analysis, investigation, original draft preparation, reviewing and editing, and acquisition of funding. Each heading (conceptualization, methodology, validation, etc) led to the same statement repeated over and over again (i.e., that the two authors were equally responsible for the area in question) and contained no other information.⁶¹

This situation contains elements more typical of show business and is reminiscent of the German rock band Milli Vanilli, which won a Grammy Award for the 1989 hit album “Girl You Know

⁶¹ This is just an example. I have not given the exact source because I do not wish to create the impression that my remarks here have the purpose of criticizing a specific publication.

It's True," despite the fact that they never sang a note, but simply lip-synched and pranced and postured. They looked like rockstars and behaved like rockstars, which was accepted as meaning that they *were* rockstars. Looking the part was enough to make them good. By contrast, qualitative research focuses not on faultless adherence to doing things in the approved way, but on actually doing the job (i.e., working out the meaning of narratives and presenting this in a clear and convincing manner). Nonetheless, the *Milli-Vanilli effect* must be avoided. Consequently, in addition to the "scientific" issues discussed in Chapter 9, a number of what might be called "stylistic" aspects of the preparation of a qualitative report need to be dealt with. Some of these will be discussed in the following paragraphs. The issue is that of making a report accessible and transparent, but also rigorous, not just putting on a good show.

"Confession" in reporting: The convention in reporting research is to smooth out irregularities, i.e., to present a cleaned and streamlined report as though everything went according to plan. However, hitches often occur, despite the fact that the tradition is to gloss them over. A simple example is that in reports on conditioning rats to run through mazes it is seldom mentioned that the animals sometimes refuse to leave the start cage. In qualitative research, which accepts a high degree of flexibility or indefiniteness in collecting data and of subjectivity in recording, analysing and interpreting the data, this problem is often much greater, and the question of how to handle this state of affairs when writing the final report is important. Referring to field research in sociology and anthropology, van Maanen (2011) described the "confessional" style with a considerable degree of irony, commenting that confession seldom goes so far as to involve admitting being wrong. Also writing as sociologists interested mainly in field research, Fontana and Frey (1998) made a similar point when they wrote with approval of "soul cleansing" in writing reports. For our purposes this means that it is appropriate to describe events in the development of a qualitative research project in the way they actually occurred, although without

exaggerated self-criticism.

Indeed, it may be necessary to do so. For instance, the object of study may have changed after a small number of interviews had been conducted, possibly because it became apparent that the original object was too complex or abstract for interviewer and/or respondents to work with, or because it proved too painful for them. The way in which interviews were conducted may have had to be altered, for a number of reasons (e.g., unwillingness of respondents to talk about the object of study using the originally planned format, need for greater structure than originally foreseen, etc). It may have proved impossible to obtain the cooperation of the group of participants originally planned. The recording of data may have had to be curtailed, with subsequent omission of certain kinds of information. As a result, it may prove to be impossible to maintain the fiction that the project started with a clearly formulated hypothesis that was systematically tested by a set of steps laid down at the beginning and followed strictly and without variation to their inevitable conclusion.

In a similar way, there may have been moments of uncertainty with the data analysis. Coding may have proved less clear cut than anticipated, for instance because the meaning of respondents' words was ambiguous, and there may have been a degree of uncertainty about some of the categories identified. Some decisions about the meaning of statements may have been taken with an obvious degree of subjectivity. Some interpretations may have a particularly strong subjective component and some conclusions may be speculations rather than strictly logical. (*The latter are allowable, provided that they are identified* for what they are. The main thing is to say clearly what happened and why – i.e., to “confess.”)

The language of the report: The data being reported in a qualitative study normally consist of “commonplace evidence” (Hamel, 1993, p. 33), usually couched in everyday language. However, the report – often a thesis or dissertation – is a scholarly document, not a piece of journalism or *belles-lettres*. Thus, it

must be couched in appropriate language, in the present case, the language of psychology and education or other social sciences. Students sometimes think that an insistence on using the specialized language of a scientific discipline involves pedantry, love of jargon, a wish to make relatively trivial or banal ideas seem more important than they really are by couching them in complicated language, an attempt to gain prestige by activating the positive associations connected with a discipline's special language, creating false insider status by using a kind of secret language shared only with specialists in a particular area, or even an attempt to shut out non-specialists who are not familiar with this language. However, the language of a discipline is not simply jargon, but is a specialized terminology that facilitates communication with other specialists, and by evoking associations connected with the discipline gives new material a particular "flavour," brings to mind related existing knowledge, reminds the informed reader of special issues such as conflicts of opinion among experts, and adds to the believability of the conclusions being drawn. *A report gains credibility from its ability to achieve a linkup with existing norms in an area of study, and this purpose is greatly facilitated by the use of appropriate specialized language.*

The status of the researcher in the report: It is important in achieving plausibility that the report reveals professionalism and self-confidence. Indeed, Kvale (1995) pointed out that the personal credibility of the researcher is a major factor determining whether readers will allow themselves to be influenced by a qualitative research report. Statements of opinion, interpretations, or even speculations are permissible, even desirable, but they should always be identified for what they are. As was pointed out above, they should not be confused with logical conclusions. They should be stated with an air of authority, even when they are identified as tentative, preliminary, or speculative, because to a considerable degree they define the extent to which a research report is interesting for readers. This does not mean claiming infallibility

or complete certainty: Speculations, hypotheses and assumptions can be identified for what they are without becoming absurd or the stuff of fairy tales. Much more important is to make clear what intellectual and scholarly basis statements have, and to show that they are rational and logically linked with existing opinion, even if they are open to informed disagreement. Students sometimes err by being overly modest and making themselves unconvincing, for instance by being too hesitant, excessively uncertain, or even presenting themselves as ignorant, inexperienced and incompetent. If there are weaknesses in execution in a qualitative study it is best to draw readers' attention to these in an apologetic way, but without cringing. One graduate student I supervised chose particularly inappropriate language in her thesis when she referred to herself as "this humble, insignificant person."

Transparent presentation: A case study

The following sections contain a concrete example of the writing up of qualitative research based on Petersen's (1989) study of the motivation of hobby authors, for which she obtained the German *Diplom* (comparable to an MA). The thesis was entitled "Motivation of amateur authors." The following sections consist partly of extracts from Petersen's thesis. These are presented in boxes. However, I must follow my own rules and confess that I have not only translated the original text from German into English, but have also edited the contents for the purposes of the present book. Of course, the passages given here are only brief extracts from Petersen's thesis, because extended quotation would have exceeded the space available. Both selection and adaptation of passages were conducted with the purpose of providing clear examples of what is meant by terms employed in this book. These passages are supplemented by brief summaries of other sections of the thesis, so

that readers can obtain a general understanding of its structure and contents. The statements in italics contain my comments on the Petersen material. In passages directly quoting Petersen's thesis, literature sources she cites are retained in the quoted text in order to show how she made use of the literature relevant to her object of study. However, the details of these sources are not included in the list of references in this book, since they do not refer to its object of study (qualitative research methodology). The passages reflect Petersen's original perspective: For instance, the expression "I" in the following sentences refers to Petersen herself, not to me.

Specifying the object of study:

"As a hobby author myself I have often participated in writers' circles. Discussions with other members of such groups suggest that they experience a strong urge to write down their ideas. Because of the apparent importance of a drive to write (motivation) suggested by these observations of hobby authors, I decided to investigate the question: **'What leads hobby authors to write?'**" *Here Petersen identifies her object of study with admirable brevity and clarity, although in everyday language.*

"Necka (1986) hypothesized that there may be a number of components of such motivation, including need for recognition, need to communicate, need to prove oneself, and so on, and that these interact. However, he provided no empirical evidence. I was thus particularly interested in whether there is a single general drive to write, or whether the motivation of hobby authors can be broken down into differentiated components and, if this should prove to be the case, what relationship these components have to each other. Of particular interest was the question whether there are different and special patterns of motivation resulting from difference combinations of motives." *Petersen defines the object of study in a more scientific,*

psychological way and also orients readers' attention to ideas that will later be central to her thesis. She writes as though she still does not know what will emerge from her research, although in reality these sentences were written after the research had been completed, so that she is really preparing readers for what she already knows is to come in her thesis. Petersen did not commence the analysis without any idea of what she was looking for – the analysis was not conducted "cold." None the less, she initially specified her object of study in general terms only.

Collecting the data: Petersen discussed various data collection procedures such as questionnaires or projective tests. However, she concluded that "the most appropriate way of collecting data would be interviews with the authors themselves (p. 5)." In making this decision, she decided to treat the authors as "experts."

"In order to avoid limiting the breadth of the interviews, it was decided to investigate the question of motivation via relatively open interviews. None the less, at least one opening question was needed, as well as a strategy that would make it possible to see to it that the participants did not wander too far from the main topic but did not limit too severely the contents they introduced into the interviews. It was also necessary to consider how the fundamental question, "Why do you write?" could be broken down into answerable sub-questions. ...Four questions were developed. These were as follows:

1. Can you remember the first time you wrote something of your own?
2. What happened after you wrote it?
3. Is there anything else you would like to mention?
4. Can you explain why you write, whereas others do not?"

Petersen lists the questions she used to give her initial interviews a certain degree of structure.

Two exploratory interviews were conducted in order to try out these questions. Petersen described these two interviews (where they were conducted and with whom, how long they took, how she recorded responses, and so on). The exploratory interviews were used to develop prompts.

“In the two exploratory interviews, which served as a pilot study, it was established that interview partners could understand these questions, found them interesting and were even eager to discuss them. Discussions in the pilot study showed that respondents spontaneously introduced eight more specific themes into their responses to the initial four questions. These themes were converted into questions that were then used as prompts in later interviews if the respondents in these interviews did not spontaneously introduce them into the conversation:

1. What was the role of external stimuli?
2. What are your thoughts about writing?
3. What are your feelings before, during and after writing something?
4. What is the role of communication in your writing?
5. What is your source of ideas for topics to write about?
6. What problems do you experience in writing?
7. How important are formal writing skills?
8. What is the role of social factors in your writing?”

Petersen carried out two preliminary interviews to see if she could persuade writers to participate and how difficult it would be to obtain their views. The data from these two interviews were incorporated into the final study and presented as though

they had been an integral part of it. Although the interviews in the main study were only loosely structured, because of the interviews in the pilot study Petersen knew what areas she wanted to cover and had the eight specific questions ready as prompts, in case respondents did not spontaneously discuss these areas. These questions were not imposed on the research by Petersen on the basis of a theory or of knowledge of the literature but emerged from the pilot study or trial run (i.e., they were “emergent” issues). They were thus the first results of the study. This demonstrates the dynamic nature of qualitative research.

Describing the participants: There were 10 participants (four men and six women), aged between 17 and 31. All participants had completed the German high-school graduation examination (*Abitur*). They were recruited via contacts with people who were hobby authors, and all of them were participants in writers’ groups. Thus, Petersen included in her sample people who obviously “represented” hobby authors by virtue of being active in this area. In the thesis she drew attention to this, described in some detail the way in which the interviews were conducted (usually in the home of the participant, although not always), as well as how long they took. She commented on the willingness of the respondents to participate and on the lack of difficulty they had in discussing the issues she was interested in and communicating their thoughts about these issues. In general, respondents dealt spontaneously with most of the issues listed above, which indicates that these were relevant to their own conceptualizations (constructions) of their situation and were not simply imposed by the researcher.

Coding the data: The interviews were tape recorded and then transcribed. Subsequently, they were subjected to the processes of summarizing and clarifying described in Chapter 7 (see p. 200ff.). This procedure was carried out one case at a time.

In the actual thesis the summarizing and clarifying steps were described in detail. The rules Petersen developed for doing this were explained, and examples of original text were given, followed by the summarized and clarified versions of this text, much as was done in Chapter 5 with the two versions of the passage on social workers' reasons for specializing in the area of refugees. In view of the fact that there were only 10 participants, each case was presented separately, with examples of how specific passages were summarized and clarified. This follows the principle of the quantum of effort introduced in Chapter 4. Even with only 10 interviews it was, of course, not possible to include the full transcripts in the thesis, because they amounted to more than 300 pages. In all, Petersen's presentation of the coding procedures in the thesis (i.e. her codebook) involved about 30 pages of text. In other words, she gave considerable attention to this aspect of her study. The purpose of this detailed presentation was to increase the credibility of her analysis by making its logic plain to readers.

Analyzing and Interpreting the Data

Content analysis of the summarized and clarified transcripts yielded many content units for each respondent. Following a procedure resembling analytic induction these were used to build categories (sub-themes) and subsequently concepts (themes), one case at a time. A brief summary of the analysis of each case was presented, one after the other, with additions, modifications and clarifications of categories and concepts being explained step by step. After analysis of the tenth interview, it became apparent that categories and concepts had converged (become stable), and the set of concepts that had emerged was presented in a single table. Table 10.1 gives examples of Petersen's coding of the interviews. Only a single content unit is given here for each category, but in the

actual study there were numerous additional content units defining each category, and in her thesis Petersen gave many examples. Naturally, Petersen also identified more than the three concepts that are presented here.

Table 10.1: Content units, categories and concepts in the interviews with hobby authors

Content Unit (Exact words from interviews)	Category/ Sub-theme (Psychological interpretations of sub-themes)	Concept/Theme (Broader themes combining sub-themes)
It helps when others like my ideas I don't like criticism It helps when I talk over my ideas People look up to me I need to tell others what I think	positive feedback negative feedback social support social status communication	SOCIAL RELATIONS
I often feel that I can't do it I feel very nervous Many ideas collide with each other I switch to something else	self-doubt anxiety internal conflicts avoidance behaviour	INTERNAL BLOCKAGES

Ideas just come from nowhere I just feel an urge to write I write when I feel excited	spontaneous ideation internal drive elevated mood	INTERNAL IMPULSE

Developing substantive theory: Returning to the individual transcripts, Petersen described the situation of the members of her group of hobby writers in some detail, structuring this description with the help of the concepts outlined in Table 10.1. Two examples of these descriptions follow:

Multicausal motivation: “The authors experienced ‘multicausal’ motivation that was the result of a developmental process going back into childhood. Their ideas sometimes derived from events in the world around them, but sometimes seemed to come from sources inside the individual, springing into existence in a way that the respondents experienced as spontaneous. The authors often experienced self-doubt or anxiety about the quality of their work, and this led to periods of reflection and doubt. They also experienced fear of being laughed at by other people. However, not only the internal satisfaction of completing a piece of work, but also external rewards, especially praise, helped them keep going. Interactions with other authors, for instance in discussion groups and clubs, provided supportive social contacts.”

This description of the situation of Petersen’s respondents goes beyond each individual person’s understanding of his or her own specific situation and constructs a more general picture of the motivation of hobby authors. However, it is still focused on the particular group of authors in the study. In this sense it involves only substantive theory.

Petersen continued with her analysis of multicausal

motivation and began to develop a more general psychological understanding of the processes through which such motivation develops.

Writing and the satisfaction of needs: “The structure of motives of each individual author was complex. All 10 indicated that writing satisfied a combination of needs in the sense of Maslow (e.g., need for belonging and acceptance; need for status; need for self-expression; need for knowledge and understanding; need for self-realization). However, the nature of the combination and the importance of particular needs differed from author to author. Differences between authors seemed to reflect not the existence of different needs in different people, but rather differences from person to person in the pattern of needs, especially the importance of particular needs for an individual person. For instance, feedback from other people was a very important factor, in general. However, how important it was depended on the particular *pattern of needs* of the individual. Those authors who displayed a pattern of motivation emphasizing need for acceptance, status, or mastery were more affected by feedback, while those in whom self-realization was emphasized were less affected by it.”

The concept of “pattern” or “profile” is very useful in qualitative studies. It makes it possible to construct more complex statements about kind or style. Petersen’s distinction here between two groups of participants involves what is, strictly speaking, a quantitative statement, since it is not based purely on kind of need (i.e., mastery need vs. self-realization), but also includes a statement about the relative strength of the needs (i.e., self-realization was emphasized more than mastery by some people). However, Petersen has avoided using language typical of quantitative studies such as, “The level of self-realization was significantly greater,” and has spoken instead of “patterns” (i.e., kind) and of emphasis within a pattern, thus

maintaining an essentially qualitative approach. Theorizing in this passage remains substantive, i.e., focused on describing the specific group of participants in this study.

Developing higher-order theory: Petersen then undertook an interesting and useful theoretical interpretation of her findings in which she drew conclusions not just about the people in her study but about the general nature of the development of (a) motivation and (b) creativity. In other words, she developed *higher order theory*. Parts of this section of her thesis follow.

Extrinsic vs. intrinsic motivation: “Amabile’s (1983) conclusion that intrinsic motivation is of special significance in creativity is important here. In her own research she showed that authors whose motivation is mainly intrinsic are more creative than those for whom it is mainly extrinsic. The present study neither confirmed nor disconfirmed this conclusion, because it did not look at the creativity of products. However, I have in several places drawn attention to the close links between extrinsic and intrinsic motivation. As far as these two areas are concerned, it is possible to see preferences or leanings in individual authors, but all participants displayed both extrinsic and also intrinsic motivation. This means either that Amabile’s assumption that extrinsic motivation blocks intrinsic is not correct, or else that all participants in the present study had developed effective strategies for limiting the effects of extrinsic motivation.”

Petersen begins to consider the implications of her findings for existing theory on motivation for creativity. In doing this, she goes beyond a description of her particular group of participants, i.e., she goes beyond substantive theory to higher order theory. She considers the implications of her

findings for Amabile's extremely influential theory, drawing attention to discrepancies between the contents of this theory and certain aspects of her findings. She suggests the need for a Redefinition (see p. 254) in this area.

Petersen continued her consideration of the implications of her findings for the theory of extrinsic/extrinsic motivation, focusing more on the development of motivation.

The dynamics of the “extrinsic constraint”: “It is also important to mention at this point that the ‘extrinsic constraint’ that can arise from extrinsic motivation and is regarded by Amabile as extremely inhibiting for creativity was hinted at in a few of the interviews, but clearly was not unfavourable to the degree described by Amabile. This was also true in the case of participants who displayed a clear pattern of extrinsic motivation. ... Amabile herself added that there is a not inconsiderable number of contradictory examples for the theory of inhibition of creativity by extrinsic motivation. She attributed these to the ability of the people in question to resist the pressure of extrinsic motivation as well as to the opportunities available to them to avoid situations in which the extrinsic constraint is activated or becomes particularly strong (Amabile, 1983, p. 96). The ability to avoid inhibition is one of the characteristics of creative people mentioned by Necka (1986). The present findings suggest that there is a continuum of extrinsic constraint and that its inhibiting effect depends upon an interaction between the strength of the constraint and the strategies for coping with the constraint that have been learned by the particular respondent.”

Petersen now turns to the theoretical model of a second influential creativity theorist. In doing so she achieves

forward incrementation or advance forward incrementation of Amabile's theory. In doing so she offers insights leading to a substantial increase in the degree of differentiation of existing theory.

Working out practical implications: One important aspect of the external validity of qualitative studies involves their ability to generate suggestions for practice. Petersen derived a number of interesting implications from her findings. Initially, the discussion was relatively general.

“Positive feedback was particularly important for the motivation to write, largely because this helped to develop self-confidence. Of great importance was that this positive feedback began in early childhood when the hobby authors first began to express creative ideas. This feedback can be regarded as positive reinforcement in the sense of Bloom (1985). It was particularly effective when the person providing the feedback was someone with high status in the child's eyes (e.g., parents, teachers and close friends), thus confirming the research of Bandura and Walters (1963) on social reinforcement of behaviour. The importance of self-confidence is scarcely surprising in view of the fact that this factor is consistently emphasized by creativity researchers (e.g., Cropley, 1982).

It is important to note that early positive feedback seemed to achieve its effect by making it possible for the authors to tolerate negative feedback or blockages later in life. This is consistent with Maslow's (1981) finding that people who experience satisfaction of a need early in life develop a higher level of tolerance for frustration of that need when they are older: a kind of immunizing or buffering effect. These findings support the view that the development of psychological properties that

are favourable for creative writing, and presumably other forms of creativity as well, is positively (or negatively) affected by interactions with other people and encourages the hope of educators that the disposition to creativity can be trained.”

Petersen goes beyond theory and begins to examine practical aspects of the development of creativity. Her remarks are useful, because they indicate that parents and teachers can influence the development of creativity. They are also made more credible by the fact that she supports them with references to the existing literature.

Petersen proceeds to more concrete suggestions for practice, focusing in particular on the classroom. Among these is a discussion of extrinsic motivation. Since a world in which there were no external rewards and punishments is almost unimaginable, it seems that extrinsic motivation will remain a fact of real life. As a result, it is highly appropriate to consider how extrinsic motivation and creativity can exist side by side.

“Certain factors seem to be favourable for creative activity, even in the presence of extrinsic rewards and punishments: Among these are high self-confidence, realistic expectations, and highly developed technical skills. On the other hand, there are also other factors that lead to negative effects of extrinsic motivation: for instance, low technical competence, low self-esteem, fear of failure. Taken together with other findings on the influence of childhood events on psychological development, this conclusion emphasizes the importance of appropriate experiences in the classroom. Among other things, these would include provision of opportunities to build up self-confidence through for instance the achieving of success in difficult situations such as examination or competitions, provision of differentiated feedback on the quality of work (including grades), and drill in basic skills.”

The positive comments on examinations, grades and basic skills is not only closely practical but it is also surprising. In discussions of creativity such activities are typically rejected as inimical to creativity. Thus, Petersen suggests a redirection in thinking on this topic.

A return to the list of ways in which qualitative research can propel a field given above (see p. 254) shows that Petersen clearly satisfied several of the criteria listed, e.g., redefinition, forward incrementation, advance forward incrementation, redirection. She also satisfied criteria listed in Table 9.3 by suggesting a revised way of thinking about motivation and creativity, and offered ideas capable of generating several hypotheses, even if she did not specifically identify the hypotheses in question. For instance, she suggested that there are differences in motivation between people who actually write and those who do not, and described the differences in a way capable of being tested in a quantitative study. She also suggested the existence of a connection between childhood experiences, the hierarchy of needs, and adult creativity. This centres on the extent to which security, status and mastery needs are met in childhood, and suggests a hypothesis that could be tested empirically by either quantitative or qualitative research. The suggestions just mentioned also open up new lines of inquiry such as the question of changes in motivation at different stages of the creative process. Indeed, research since Petersen's study (see Cropley, 2001) shows that a number of her suggestions were well founded.

Two Contrasting Examples

Petersen's study does not represent the only way of going about

qualitative research. A second example from a *Diplom* student (equivalent to a master's degree in the German system as it was organized at the time) at the University of Hamburg shows some alternative possibilities. Plate (1997) was interested in "the conditions that make it possible for workers to apply their creativity in job settings." *He states his object of study clearly and in relatively straightforward language.* The first two chapters of his thesis contained a review of the relevant literature on (a) creativity and (b) work settings. On the basis of this review and before talking with any participants at all he worked out concrete themes that he specifically addressed in the interviews (largely in the form of questions), unless the respondents dealt with the themes without being asked. These included (a) Describe three ideas you have had at work. (b) Where did your best idea come from? (c) What blocked you expressing your idea? (d) What would have happened if you had expressed (implemented) your idea? (e) What would have helped you express your idea? *Plate defined his object of study much more exactly than Petersen and structured the interviews according to fairly precise a priori expectations developed in his review of the relevant literature. It can be seen that he oriented the interviews towards discussion of blocks to creativity such as negative consequences of attempting to be creative or missing supportive factors. Implicit in his approach was the image of work as a negative environment for creativity. This reflected the position found in the literature he reviewed.*

Plate tape-recorded the interviews and also made notes. The tapes were transcribed and the resulting texts analysed in order to identify inhibiting or facilitating factors. The content analysis identified only content units relating to blocking factors in the work setting, since this was the object of study, negative consequences for the respondent of displaying or attempting to implement creativity, and factors that the participants believed would have been facilitating, had they existed (but they did not). *The analysis was highly focused and looked specifically for concepts that Plate's*

*review of existing research indicated would be interesting. There was only limited room for emergent issues (see p. 195), and these were of interest only when they helped to clarify issues that Plate had decided to concentrate on a priori.*⁶²

The results identified two basic kinds of blocking factors: “internal” and “external.” Internal factors were essentially psychological in nature: e.g., “image” (respondents’ concern about how others saw them), “identification (or not) with the idea,” “level of motivation,” “willingness (or not) to take risks,” and “willingness to risk conflict with colleagues.” Factors outside the individual included “goals of the organization,” “demand for creativity in the job itself,” “opportunities to work without supervision,” and “time pressure.” In the last chapter of his thesis Plate first used these dimensions to describe the situation of his respondents in a more general way (i.e., he constructed “substantive” theory). He then went on to relate his findings to the original literature he had reviewed in the opening chapters, especially indicating where his findings supported or contradicted existing research findings (higher order theory).

Plate commenced the thesis by developing an a priori theoretical framework, collected data with a specific orientation based on this framework (this almost involved formulating hypotheses), and coded his data in order to cast light on these expectations, finally returning in the last chapter to a discussion of the degree to which his findings supported existing theory. Thus, his approach had many of the characteristics of quantitative research, although it is clear that his design (data collection, analysis and interpretation) was strongly within the qualitative framework.

Gray and Kunkel’s (2001) study of ballet dancers and the way they experience themselves and their profession also demonstrates departures from Petersen’s design, as well as from Plate’s. Most striking of all, they did not conduct any interviews

62 I am grateful to Petersen and Plate for their outstanding work and the contributions they made to the development of my understanding of qualitative methods.

themselves at all, but carried out a content analysis of interviews found in a book that had already been published a number of years previously by a journalist (Gordon, 1983). This book described Gordon's observations of professional dancers and dance students in New York, Houston and Boston. She observed them in classes, in rehearsals and in social interactions, and recorded what they said and did, i.e., she was a participant observer. She wrote down a large number of verbatim statements by the dancers themselves about many aspects of their lives. Gray and Kunkel treated the contents of Gordon's book as though they were in effect a series of interview transcripts. *I have already mentioned the possibility of using various kinds of existing records as a source of data. This would be particularly useful in studies with a historical orientation, since it is in principle possible to examine events of the past in this way. Although it is obvious that the cases studied by Gordon did not constitute a random or stratified sample, it is also obvious what object of study they "represented" – ballet dancers' view of their work – and it is hard to deny that they would be likely to know about this subject.*

Gray and Kunkel subjected Gordon's text to a content analysis conducted as strictly as possible according to the principles of grounded theory. They focused on finding content units in which the dancers described their own understanding of their professional experiences and of themselves as dancers (as against their views on other aspects of their lives, or statements revealing the perspective of other people such as parents or teachers). They tried to include all content units that satisfied this criterion. Thus, they did not select material on the basis of an *a priori* interest in particular aspects of the dancers' perspective, but they did limit themselves to material that referred to dancing. They identified a total of 167 content units that reflected dancers' constructions of this aspect of their lives. These were used to define categories such as "dancers are an expression of other people's fantasies," "the dancers pay for the viewers' pleasure," or "dancers enter a

fantasy realm.” The categories defined five concepts: “Fantasy creatures,” “Mechanization,” “Commodification,” “Infantilization,” and “Sacrificial transcendence.” *Of course, the texts analysed by Gray and Kunkel had already been selected and edited by Gordon. However, this can to some extent be regarded as corresponding to the steps of summarizing and clarifying. It is not known to what extent Gordon focused on areas of special interest to her, but this does not mean that the concepts identified by Gray and Kunkel contain no insights into how dancers construct their own lives.*

Gray and Kunkel described dancers in terms of the concepts just listed (“substantive theory”), and then attempted to relate these categories to psychoanalytic theory, feminism, social psychological models of power relationships, and the like, giving considerable emphasis to issues like object relations, autonomy, individuation and identity. *In doing this they developed higher-order theory that went beyond description of the situation of one particular group of dancers and related their data to general concepts in psychology. In addition, they helped to clarify the meaning of terms like “autonomy” and the nature of processes like “individuation,” by showing how these manifest themselves in a real-life situation. Finally, the authors discussed the possible role of art as a buffer against the depersonalizing effects in society of advancing technology, and suggested that training of dancers should give more emphasis to “artistic integrity” than is presently the case. At the end of their report Gray and Kunkel turned to the practical implications of their findings, in particular for the education of dancers, thus attempting to satisfy the criterion of usefulness.*

References

- Alasuutari, P. (1995). *Researching culture: Qualitative method and cultural studies*. London: Sage.
- Allport, G. W. (1965). *Letters from Jenny*. New York: Harcourt, Brace and World.
- Altheide, D. L. & Johnson, J. M. (1998). Criteria for assessing interpretive validity in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (pp. 283-312). Thousand Oaks, CA: Sage.
- Andersone, M. (2002). *Datorspēles un jauniešu dzīvestils* [Computer games and youth lifestyle]. Riga: Unpublished MA thesis, Department of Sociology, University of Latvia.
- Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, 193, 31-35.
- Bandura, A. & Walters, R. H. (1963). *Social learning and personality development*. New York: Holt, Rinehart and Winston.
- Barthes, R. (1973). *Mythologies*. New York, NY: Hill and Wang.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bleicher, J. (1980). *Contemporary hermeneutics: Hermeneutics as method, philosophy and critique*. London, UK: Routledge.
- Blumer, H. (1969). *Symbolic interactionism*. Englewood Cliffs, NJ: Prentice-Hall.
- Braun, V. & Clark, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Breuer, F., Muckel, P., & Dieris, B. (2018, 3rd ed.). *Reflexive Grounded Theory. Eine Einführung für die Forschungspraxis* [An introduction to research practice]. Wiesbaden: Springer.
- Briggs, C. (1986). *Learning how to ask: A sociolinguistic appraisal of the role of the interviewer in social science research*. Cambridge: Cambridge University Press.

- Bromme, R. & Kienhuis, D. (2017). Gewissheit und Skepsis: Wissenschaftskommunikation als Forschungsthema der Psychologie [Certainty and scepticism: Science communication as the subject of psychological research]. *Psychologische Rundschau*, 68, 167-240.
- Bruner, J. S. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Bryman, A. & Burgess, R. G. (Eds.). (1994). *Analyzing qualitative data*. London: Routledge.
- Buckley, D. L. (1975). *Gymnasium im Wandel*. Weinheim: Beltz.
- Casey, E. (1987). *Imagining: A phenomenological study*. Bloomington, IN: Indiana University Press.
- Cattell, R. B. (1950). *Personality: a systematic theoretical and factual structure*. New York: McGraw Hill.
- Charmaz, K. (2014, 2nd ed.). *Constructing grounded theory. Introducing qualitative methods*. Thousand Oaks, CA: Sage.
- Cicourel, A. V. (1974). *Theory and method in a study of Argentine fertility*. New York: Wiley.
- Coghlan, D. (2011). *Action research: Exploring perspectives on a philosophy of practical knowing*. *Academy of Management Annals*, 5, 53-87.
- Cook, T. D. & Campbell, D. T. (1979). *Quasiexperimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Corbin, J. L. & Strauss, A. M. (2015, 4th ed.). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.
- Couch, A. & Keniston, K. (1960). Yeasayers and naysayers: Agreeing response set as a personality variable. *Journal of Abnormal and Social Psychology*, 60, 151-174.
- Crang, M. & Cook, I. (2007). *Doing ethnographies*. Thousand Oaks, CA: Sage.
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.
- Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Clegg Smith, K.

- (2011). *Best practices for mixed-methods research in the health sciences*. Washington, DC: National Institutes of Health.
- Cronbach, L. J. (1980). Validity on parole: How can we go straight? *New directions for Testing and Measurement*, 5, 99-108.
- Cropley, A. J. (2001). *Creativity in education and learning*. London: Kogan Page.
- Cropley, A. J. (2002). Problems and concerns of welfare paraprofessionals working with refugees. *British Journal of Social Work*, 32, 233-238.
- Cropley, A. J. & Cropley, D. H. (2008). Resolving the paradoxes of creativity: An extended phase model. *Cambridge Journal of Education*, 38, 355-373.
- Cropley, A. J. & Cardey, R. M. (1975). Contact with the dominant culture and cognitive competence in Canadian Indians and Whites. *Canadian Journal of Behavioural Science*, 7, 328-338.
- Cropley, A. J. & Lüthke, F. (1994). Psychologische Aspekte der Adaptation von Zuwanderern. In: A.J. Cropley, H. Ruddat, D. Dehn, & S. Lucassen (Eds.), *Probleme der Zuwanderung*, Band 1, *Aussiedler und Flüchtlinge in Deutschland* [Problems of migration, vol 1, Returnees and refugees in Germany] (p. 19-32). Göttingen, Germany: Verlag für Angewandte Psychologie.
- Cropley, A. J. & Lüthke, F. (1995). Strategien für die psychologische Beratung von Zuwanderern [Strategies for counselling immigrants]. In: A.J. Cropley, H. Ruddat, D. Dehn & S. Lucassen (Eds.), *Probleme der Zuwanderung*, Band 2, *Theorien, Modelle und Befunde der Weiterbildung* [Problems of migration, vol. 2, Theories, models and findings of inservice training] (p.146-160). Göttingen: Verlag für Angewandte Psychologie.
- Cropley, D. H. & Cropley, A. J. (2016). Promoting creativity through assessment: A formative CAA tool for teachers. *Educational Technology*, 56(6), 17-24.
- Cropley, D. H. & Cropley, A. J. (2021). *Core capabilities for Industry 4.0: Foundations of the cyber-psychology of 21st century engineering education*. Bielefeld: Bertelsmann Verlag.

- Cropley, D. H. & Marrone, R. L. (2021). Automated scoring of figural creativity using a convolutional neural network. *PsyArXiv Preprints* (Preprint DOI: 10.311234/osf.io/8qe7y).
- Crowne, D. P. & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence*, New York: Wiley
- Denzin, N. K. & Lincoln, Y. S. (1998). *Collecting and interpreting qualitative materials*. Thousand Oaks, CA: Sage.
- Denzin, N. K. & Lincoln, Y. S. (Eds.) (2017, 5th ed.). *The Sage handbook of qualitative research*. Thousand Oaks, CA: Sage.
- DGP (2017). *DGP unterstützt den March for Science!* [German Psychological Society supports the March for Science!]. [Downloaded from http://www.dgps.de/index.php?id=143&tx_ttnews%5D=1790@cHash=0044153a64e82bb4113feebdd5065d7c on July 3, 2017.
- Douglas, J. D. (1985). *Creative interviewing*. Newbury Park, CA: Sage.
- Ebbinghaus, H. (1983). *Über das Gedächtnis* On memory [1885]. Passau, Germany: Passavia-Universitäts-Verlag.
- Edwards, M. A. & Roy, S. (2017). Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science*, 34, 51-61. (downloaded from <https://doi.org/10.1089/ees.2016.0223>)
- Eidson, J. I. (2021). *Mentorship, identity, and self-efficacy in mothers returning to the workforce after an extended leave*. Doctoral dissertation, Grand Canyon University, Phoenix, Az.
- Elliott, R. (1999). Editor's introduction to special issue on qualitative psychotherapy research: Definitions, themes and discoveries. *Psychotherapy Research*, 9, 251-257.
- Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, 208, 1181-1182.
- Erikson, E. (1968). *Identity, youth and crisis*. New York, NY: Norton.
- Fechner, G. (1860). *Elemente der Psychophysik* [Elements of psychophysics]. Leipzig, Germany: Breitkopf und Härtel.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA:

- Stanford University Press.
- Flanders, N. A. (1970). *Analysing teacher behavior*. Reading, MA: Addison-Wesley.
- Fontana, A. & Frey, J. H. (1998). Interviewing: The art of science. In N. K. Denzin and Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (pp. 47-78). Thousand Oaks, CA: Sage.
- Freeman, M. (1993). *Rewriting the self: History, memory, narrative*. London: Routledge.
- Friedman, H. S. & Schustack, M. W. (2012), *Personality: Classic theories and modern research*, Boston, MA: Pearson Allyn & Bacon.
- Freud, S. (1964 [1910]). *Leonardo da Vinci, A memory of his childhood*. Oxford, UK: W. W. Norton.
- Gergen, K. J. (1985). The social constructionist movement in modern psychology. *American Psychologist*, 40(3), 266-275.
- Geertz, C. (2000). *The interpretation of cultures*. New York, NY: Basic Books.
- Ghiara, V. (2020) Disambiguating the role of paradigms in mixed methods research. *Journal of Mixed methods Research*, 14(1). [Downloaded from <https://doi.org/10.1177/1558689818819928> on June 30, 2021]
- Glaser, B. G. & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine.
- Gordon, S. (1983). *Off balance*. New York: Pantheon Books.
- Gray, K. M. & Kunkel, M. A. (2001). The experience of female ballet dancers: a grounded theory. *High Ability Studies*, 12, 7-26.
- Gunersel, A. B. (2009). A qualitative study of the impact of environmental and personal factors on prominent Turkish writers. *Psychology of Aesthetics, Creativity and the Arts*, 3, 222-231.
- Gusdorf, G. (1980). Conditions and limits of autobiography. In J. Olney (Ed.) *Autobiography: Essays theoretical and critical* (pp. 28-48). Princeton, CA: Princeton University Press.

- Hamel, J. (1993). *Case study methods*. Thousand Oaks, CA: Sage.
- Hammersley, M. (1992). *What's wrong with ethnography?* London, UK: Routledge.
- Hannes, K. (2011). Critical appraisal of qualitative research. In J. Noyes, A. Booth, K. Hannes, A. Harden, J. Harris, S. Lewin, & C. Lockwood (Eds.), *Supplementary Guidance for Inclusion of Qualitative Research in Cochrane Systematic Reviews of Interventions*. Cochrane Collaboration Qualitative Methods Group.
- Harris, M. (2006). *Cultural anthropology*. Boston, MA: Allyn and Bacon.
- Hawking, S. & Mlodinow, L. (2010). What is reality? In *The grand design* (pp. 39-59). New York, NY: Bantam.
- Herborth, B. (2011). Methodenstreit – Methodenzwang – Methodenfetisch [Methodological dispute – methodological rigidity – methodological fetishism]. *Zeitschrift für Internationale Beziehungen* [Journal of International Relations], 18(2), 137–152.
- Heron, J. (1992). *Feeling and personhood. Psychology in another key*. London: Sage.
- Hetland, L., Winner, E., Veenema, S., & Sheridan, K. (2013). *Studio thinking 2: The real benefits of visual arts education*. New York, NY: Teachers College Press.
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010, 3rd ed.), *Cultures and organizations: Software of the mind*. New York, NY: McGraw-Hill.
- Hogan, J. Murdock, K., Hamill, M., Lanzara, A., & Winner, E. (2018). Looking at the process: Examining creative and artistic thinking in fashion designers on a reality television show. *Frontiers in Psychology*, 9, doi: 10.3389/fpsyg.2018.02008
- Hollis, M. (1994). *The philosophy of social science*. Cambridge, UK: Cambridge University Press.
- Holstein, J. A. & Gubrium, J. F. (2015). *The active interview*. Thousand Oaks, CA: Sage.
- Huber, G. L. & Gürtler, L. (2012). *Manual zur Software AQUAD 7*

- [originally published in 2003]. Tübingen: Ingeborg Huber Verlag.
- Ibaviosa, C. M. & Chin, J. M. (2022). Beyond CSI: Calibrating public beliefs about the reliability of forensic science through openness and transparency. *Science & Justice*.
<https://doi.org/10.31222/osf.io/tvcm6>
- Jorgensen, D. L. (1989). *Participant observation*. Newbury Park, CA: Sage.
- Josselson, R. & Lieblich, A. (1993). A narrative introduction. In R. Josselson & A. Lieblich (Eds.), *The narrative study of lives*, Vol. 1 (pp. ix-xv). Newbury Park, CA: Sage.
- Kampylis, P. & Valtanen, J. (2010). Redefining creativity – Analyzing definitions, collocations, and consequences. *Journal of Creative Behavior*, 44, 191-214.
- Katz, J. (1983). A theory of qualitative methodology: The social science system of analytic fieldwork. In R. M. Emerson (Ed.), *Contemporary field research* (pp. 127-148). Boston, MA: Little, Brown.
- Kelle, U. & Kluge, S. (1999). *Vom Einzelfall zum Typus* [From the single case to the type]. Opladen, Germany: Leske und Budrich.
- Kershaw, T. C., Clifford, R. D., Khatib, F., & El-Nasan, A. (2022). An initial examination of computer programs as creative works. *Psychology of Aesthetics, Creativity, and the Arts*.
 [Retrieved from <https://doi.org/10.1037/aca0000457>]
- Kivunja, C. & Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education*, 6(5), 31-41.
- Knapper, C. K. & Cropley, A. J. (1980). Interpersonal factors in driving. *International Review of Applied Psychology*, 29, 415-438.
- Knapper, C. K. & Cropley, A. J. (1982). Public attitudes to property and property insurance. *International Review of Applied Psychology*, 31, 425-441.
- Kremer, J.-F. (2010). Das Führen leitfadengestützter Interviews. In N. Schöneck-Voß & N. Voß (Eds.). *Methodenintegrative Forschung*

- (7-14). Bochum, Germany: Ruhr-Universität Bochum.
- Kromrey, H. (2009). *Empirische Sozialforschung* (12. Aufl.). Stuttgart, Germany: UTB.
- Kuckartz, U. (2007). *Einführung in die computergestützte Analyse qualitativer Daten* [Introduction to computer-based analysis of qualitative data]. Wiesbaden, Germany: VS Verlag.
- Kuckartz, A. M. & Kuckartz, U. (2002). Qualitative text analysis with MAXQDA. [Retrieved from https://www.researchgate.net/publication/239926542_Qualitative_Text_Analysis_with_MAXQDA on June 20, 2016.]
- Kvale, S. (1995). The social construction of validity. *Qualitative Inquiry*, 1(1), 19-40.
- Kvale, S. (1996). *InterViews*. Thousand Oaks, CA: Sage.
- Kyne, D. (2021). What is mixed-method research? A definition and why it's becoming so popular. [Retrieved from <https://medium.com/the-full-stack-researcher/what-is-mixed-methods-research-a-definition-and-why-its-becoming-so-popular-b435629e1acd> on July 3, 2021.]
- Lamnek, S. (1988). *Qualitative Sozialforschung. Band 1: Methodologie* [Qualitative social research, vol. 1, methodology]. Weinheim, Germany: Psychologie Verlags Union.
- Lamnek, S. (1989). *Qualitative Sozialforschung. Band 2: Methoden und Techniken* [Qualitative social research, vol. 2, methods and procedures]. Weinheim, Germany: Psychologie Verlags Union.
- Leiris, M. (1985). *Das Auge des Ethnographen* [The ethnographer's eye]. Frankfurt a.M, Germany: Suhrkamp.
- Levinson, D. J. (1996). *The seasons of a woman's life*. New York, NY: Knopf Publishers.
- Lewin, K. (1997). Field theory and experiment in social psychology. In D. Cartwright (Ed.), *Field theory in social science: Selected theoretical papers by Kurt Lewin* (pp. 130-154). New York, NY: Harper & Brothers Publishers.
- Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

- Lokot, M. (2021). Whose voices? Whose knowledge? A feminist analysis of the value of key informant interviews. *International Journal of Qualitative Methods*, 20, 1-8.
- Madison, G. B. (1988). *The hermeneutics of postmodernity*. Bloomington, IN: Indiana University Press.
- Mahto, M., Hogan, S. K., Hatfield, S., Coppola, M., & Kulkarni, A. (2021). *Looping in your new sidekick. The role of machine learning in qualitative data analysis*. Retrieved from [<https://www2.deloitte.com/us/en/insights/focus/technology-and-the-future-of-work/machine-learning-qualitative-data.html> on November 20, 2021.]
- Malinowski, B. (1929). *The sexual life of savages in North-Western Melanesia*. New York, NY: Eugenics Pub. Co.
- Maslow, A. H. (1962). *Towards a psychology of being*. Princeton, NJ: Van Nostrand.
- Mason, J. (1994). Linking qualitative and quantitative data analysis. In A. Bryman & R. G. Burgess (Eds.), *Analyzing qualitative data* (pp. 89-110). London, UK: Routledge.
- Mayring, P. (2008). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* [Qualitative content analysis] (10. Auflage) [10th ed.]. Weinheim, Germany: Beltz.
- Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 34, 103-115.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Miles, M. B. (1979). Qualitative data as an attractive nuisance. *Administrative Science Quarterly*, 24, 590-601.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: A sourcebook of new methods*. Beverly Hills, CA: Sage.
- Miller, A. I. (2000). *Insights of genius*. Cambridge, MA: MIT Press.
- National Centre for Vocational Education Research (2013). *Are we there yet?* Report presented at NCVER National Forum on research findings from the Longitudinal Surveys of Australian Youth (LSAY). Sydney, Australia, Thursday 11 April 2013.

- Necka, E. (1986). On the nature of creative talent. In A. J. Cropley, K. K. Urban, H. Wagner, & W. H. Wiczerkowski (Eds.), *Giftedness: A continuing worldwide challenge* (pp. 131-140). New York: Trillium.
- Niederberger, M. & Dreieck, S. (2020). Die qualitative Inhaltsanalyse in den Gesundheitswissenschaften. Ergebnisse eines systematischen Reviews einschlägiger Fachzeitschriften [Qualitative content analysis in health sciences. Findings of a systematic review of relevant journals]. *Forum Qualitative Sozialforschung*, 21, 1-48. [Downloaded from file:///C:/Users/Arthur/Downloads/3423-Article%20Text- 14359-2-10-20191208.pdf on May 24, 2021]
- Olney, J. (Ed.). (1980). *Autobiography: Essays theoretical and critical* (pp. 28-48). Princeton, CA: Princeton University Press.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251).
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Thousand Oaks, CA: Sage.
- Petersen, S. (1989). *Motivation von Laienautoren*. [Motivation of hobby authors]. Unveröffentlichte Diplomarbeit [unpublished master's thesis], Universität Hamburg.
- Pike, K. L. (1967). *Language in relation to a unified theory of the structure of human behavior*. The Hague, NL: Mouton.
- Plate, M. (1997). *Bedingungen für kreatives Handeln im Beruf* [Promoting creative behaviour in the workplace]. Unveröffentlichte Diplomarbeit [unpublished master's thesis], Universität Hamburg.
- Probst, P. (1997). Behinderungsbezogene Gesundheitsschemata bei Eltern autistischer Kinder [Handicap-related health schemata in parents of autistic children]. *Zeitschrift für klinische Psychologie, Psychiatrie und Psychotherapie [Journal of Clinical Psychology, Psychiatry, and Psychotherapy]* 41(1), 1-15.
- Riessman, C. K. (1993). *Narrative analysis*. Newbury Park, CA: Sage.

- Ravenstein, E. G. (1885). The laws of migration. *Journal of the Statistical Society*, 48(2), 167-235.
- Ricoeur, P. (1991). Narrative identity. *Philosophy Today*, 35(1), 73-81.
- Riddoch, M. J. & Humphreys, G. W. (1987). A case of integrative visual agnosia. *Brain*, 110(6), 1431-1462.
- Ritchie, J. & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman & R. G. Burgess (Eds.), *Analyzing qualitative data* (pp. 173-194). London: Routledge.
- Rogers, C. R. (1980). *A way of being*. Boston, MA: Houghton Mifflin.
- Rosenthal, R. (1976). *Experimenter effects in behavioral research*. New York, NY: Irvington.
- Sampson, R. J. & Wilson, W. J. (1995). Toward a theory of race, crime, and urban inequality. In J. Hagen & R. G. Peterson (Eds.), *Crime and inequality* (pp. 37-56). Stanford, CA: Stanford University Press.
- Sarbin, T. R. (Ed.). (1986). *Narrative psychology: The storied nature of human conduct*. New York, NY: Praeger.
- Schaefer, R. (1992). *Retelling a life: Narration and dialogue in psychoanalysis*. New York, NY: Basic Books.
- Scheppele, K. L. (2004). Constitutional ethnography: An introduction. *Law & Society Review*, 38, 389-406.
- Schmid, T. (2016). *Creativity and mental well-being in older Australians: A mixed-method study*. (Unpublished doctoral thesis) Charles Sturt University, Thurgoona, NSW, Australia.
- Sebre, S. (1992). *Autobiographical childhood narratives: Processes of remembering and reconstructing*. Unpublished doctoral dissertation, City University of New York.
- Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2014). *Research methods in psychology* (10th ed.). New York, NY: McGraw Hill.
- Silvia, P. J., Kaufman, J. C., Reiter-Palmon, R., & Wigert, B. (2011). Cantankerous creativity: Honesty-humility, agreeableness, and the HEXACO structure of creative achievement. *Personality and Individual Differences*, 51, 687-689.

- Smith, J. K. & Heshusius, L (1986). Closing down the conversation: The end of the qualitative-quantitative debate among educational researchers. *Educational Researcher*, 15, 4-12.
- Spradley, J. (1979). *The ethnographic interview*. New York, NY: Holt, Rinehart and Winston.
- Stake, R. E. (1998). Case studies. In N. K. Denzin and Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (pp. 86-109). Thousand Oaks, CA: Sage.
- Sternberg, R. J., Kaufman, J. C., & Pretz, J. E. (2000). *The creativity conundrum: A propulsion model of types of creative contributions*. New York, NY: Psychology Press.
- Strauss, A. & Corbin, J. (1998). *Grounded theory: Grundlagen qualitativer Sozialforschung*. Weinheim, Germany: Beltz.
- Strauss, A. & Corbin, J. (1998). Grounded theory methodology. In N. K. Denzin & Y. S. Lincoln, (Eds.). *Strategies of qualitative inquiry* (pp. 158-183). Thousand Oaks, CA: Sage.
- Tashakkori, A. & Teddlie, C. (2010). *Sage handbook of mixed methods in social and behavioural research*. Thousand Oaks, CA: Sage.
- Taylor, S. J., Bogdan, R., & DeVault, M. (2015). *Introduction to qualitative research methods*. New York, NY: Wiley.
- Thomas W. I. & Thomas, D. S. (1928). *The child in America*. New York, NY: Knopf.
- Thomas, W. I. & Znaniecki, F. (1927). *The Polish peasant in Europe and America* (2nd ed.). New York, NY: Knopf.
- Titchener, E. B. (1912). The schema of introspection. *American Journal of Psychology*, 23, 485-508.
- van Maanen, J. (2011). *Tales of the field: On writing ethnography* (2nd ed.). Chicago, IL: University of Chicago Press.
- Verplanck, W. S. (1956). The operant conditioning of human motor behavior. *Psychological Bulletin* 53, 70-83.
- von Saldern, M. (2001). Klassengröße [Class size]. In: D. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (2. Aufl.) [*Handbook of educational psychology*] (2nd. Ed.) (pp. 326-331). Weinheim: Beltz.

- Weckowicz, T. H., Younge, K. A., Cropley, A. J., & Muir, W. (1971). Objective therapy predictors in depression. *Journal of Clinical Psychology Monograph Supplement*, 1-27 (Whole No. 31).
- Weitzman, E. A. (2000). Software and qualitative research. In N. K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 803-820). Newbury Park, CA: Sage.
- Widdershoven, G. A. M. (1993). The story of life: Hermeneutic perspectives on the relationship between narrative and life history. In R. Josselson & A. Lieblich (Eds.), *The narrative study of lives*, Vol. 1 (pp. 1-20). Newbury Park, CA: Sage.
- Wilson, V. (1996, September). "The last blue mountains barred with snow? An exploration of managing qualitative research." Paper presented at the conference of the Scottish Council for Research in Education, Dundee, September 24-26, 1996.
- Wisdom, J. & Creswell, J. W. (2013). Mixed methods: Integrating quantitative and qualitative data collection and analysis while studying patient-centered medical home models. North Bethesda, Md: Agency for Healthcare Research and Quality.
- Woodman, R. W. (2014). The science of organizational change and the art of changing organizations. *Journal of Applied Behavioral Science*, 50, 463 –477.
- Znaniecki, F. (1935). *The method of sociology*. New York: Holt, Rinehart and Winston.

Appendix A: Checking the quality of a thesis

The criteria that have been outlined earlier in this chapter can be used to check whether a thesis or dissertation possesses at least the *formal* properties needed for its acceptance as a scientific report (i.e., structure, organization, completeness, etc). Of course, satisfying formal or technical criteria does not necessarily mean that the *contents* of a thesis are consistent with scholarly opinion or that the novelty it produces is useful. Elegance, surprisingness, usefulness, and similar properties can only be judged adequately by people who possess the necessary expertise based on knowledge and experience. None the less, if the formal requirements are not satisfied, the document is not ready and should not be submitted. If it has been submitted, it should not be accepted.

The following are questions that students should ask themselves about their own work before submitting it. (The questions can also be adapted for use by supervisors and examiners considering accepting a thesis or grading an assignment.) If the answer to any question is “No,” there is a corresponding weakness in the report and serious consideration should be given to revising it.

1. *Specification of the object of study:* Has the object of study been clearly specified in easily understandable language in the first page or two? Have you explained the source of your interest in the object? Have you presented the object (at least briefly) as a psychological/educational phenomenon? Have you indicated whether you started the study with any specific expectations? If you did, have you described these, as well as indicating their origin (e.g., previous research, your own intuitions)?
2. *Description of the participants:* Have you described

the group of participants demographically and indicated how you obtained their cooperation. Have you described the sample in such a way that it is obvious what they “represent”?

3. *Description of the data collection*: Have you explained what procedure was employed to collect the data (e.g., participant observation, unstructured interview, in-depth interview, etc)? Have you given relevant details such as the question used to start interviews, prompts that were used to structure the interviews, etc? Have you described the circumstances under which the data were collected (e.g., in the participants’ homes, at work, with or without other people present, how long the contact with the participants lasted, etc)? Have you explained how the narratives were recorded (e.g., tape recording, notes, minutes)?

4. *Description of the initial data analysis*: Have you explained how the narratives were converted to a written text (e.g., transcription of tape recordings, etc)? Have you explained how you went about the analysis (e.g., according to *a priori* categories or “cold”)? If there were *a priori* categories where did they come from (e.g., from existing theory, from a personal interest, from an initial reading through of one or more transcripts)? Was there any summarizing and clarifying? If there was, what rules were used to guide this procedure? Have you given sufficient examples of passages from actual narratives in order for readers to understand what you did and reproduce it (at least approximately), if they wished to?

5. *Description of the content analysis*: Did you follow a procedure resembling analytic induction or grounded theory? In what way did you decide which content units to identify (e.g., all content units referring to the object of study, an initial theory, a theory formed during the analysis)? Have you given a number of examples

of content units from several cases? Have you used content units to form *categories*? Have you named the *categories* you identified and used psychological or educational terminology as much as possible? Have you explained the rationale of your categories? Have you given several examples of content units defining each category (preferably in a table)? Have you combined categories to form *concepts*? Have you given the concepts psychological/educational names? Have you explained the rationale of the concepts?

6. *Formulating theory*: Have you used the concepts identified to develop a general description of the way your object of study manifests itself in your participants (i.e., substantive theory)? Have you linked this specifically to actual content units and categories? Is this substantive theory couched in scientific language? Does it go beyond what the respondents said? Does it go beyond simple everyday knowledge? Does it *create surprise* by propelling the field (your object of study)? Have you at least attempted to indicate the significance of your substantive theory for more general psychological or educational knowledge or theory (i.e., higher-order theory)?

7. *Demonstrating usefulness*: Have you discussed the implications of your findings for psychological or educational practice (e.g., childrearing, counselling, therapy, teaching, training, learning, personnel selection, work conditions, etc)?

If you have answered “Yes,” to every question and the “Yes,” is really justified, you should have few problems with your thesis. Remember that this list describes the perfect document: Some deviations from perfection can be tolerated. It should also be borne in mind that weaknesses in one area can, to some extent, be compensated for by strengths in another, especially where a thesis is very strong in some areas. The checklist that follows specifies

these criteria more explicitly. I recommend that you use it to check whether you have met the criteria of a rigorous thesis or dissertation.

Checklist		
Specification of the object of study	Yes/No	Page
The object of study has been stated clearly in the first 2 or 3 pages.		
I have explained the source of my interest in the object of study.		
I have stated the object of study in a scientific way.		
I have indicated how I structured the investigation.		
Description of the “sample”		
I have explained in what sense the participants are representative.		
I have described the group of participants demographically.		
I have explained how the participants were recruited.		
Data collection		
I have described the data collection procedure clearly and fully.		
I have explained what setting the data were collected in.		
If there was one, I have described the pilot study.		
If there was a pilot study, I have explained how it is linked to the main study.		
I have given a full and detailed description of the collection of data.		

I have explained how the data were recorded.		
Data analysis	Yes/No	Page
I have explained how the data were turned into a text for analysis.		
I have explained the role of theory in the initial stage of the analysis.		
I have explained where the theoretical structure was derived from.		
I have explained the rules according to which any preliminary “cleaning” of the data occurred.		
There are enough concrete examples to make the analysis understandable.		
I have explained the rules for building content units.		
There are enough understandable examples of the coding of data.		
My categories and concepts have scientific names.		
I have explained the logic of the concepts in an understandable way.		
Expansion of knowledge		
I have described the participants and/ or the setting in terms of the concepts worked out in the analysis.		
This description goes beyond the statements of the participants and is more than simple common sense.		
I have developed “substantive” theory.		
I have demonstrated this theory’s connection to the content units.		

I have formulated this theory in scientific language.		
I have developed “superordinate” or “higher-order” theory.		
I have demonstrated the links between my theory and existing theory.		
Usefulness and Seminality		
I have given an overview of knowledge about my object of study.		
I have covered a wide spectrum of relevant material.		
I have established links to everyday life.		
I have worked out the implications of my findings for practice.		
I have made suggestions for further research on my object of study.		
Beauty		
The report is well organized.		
The report is well written.		
I have supported my arguments and descriptions with enough examples.		
My arguments are logical and convincing.		
Creativity		
I have created effective surprise.		
I have propelled the field.		
The following aspects of my report have created effective surprise and propelled the field in the following way:		

Appendix B: Helpful introductory texts with a practical orientation

- Babbie, E. (2015). *The practice of social research* (14th ed.). Belmont, CA: Wadsworth. (Good discussion of how to conduct structured interviews.)
- Berg, B. L. & Lune, H. (2018). *Qualitative research methods for the social sciences* (9th ed.). Chennai, India: Pearson India. (Good overview of all the issues touched upon in the present book: very useful as supplementary reading and as a sourcebook for students wishing to go into various issues more deeply.)
- Bryman, A. & Burgess, R. G. (Eds.). (1994). *Analyzing qualitative data*. London: Routledge. (Overview of methods of data analysis.)
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage. (The title is self-explanatory.)
- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, 8, 597-607. [Available online at <http://www.nova.edu/ssss/QR/QR8-4/golafshani.pdf>]
(Overview of the early discussion contrasting the understanding of reliability and validity in qualitative research with the way it is understood in quantitative research.)
- Hedlund-deWitt, N. (2013). CODING: An overview and guide to qualitative data analysis for integral researchers. *IRC Resource Paper No.1*. [Available online at https://www.academia.edu/9864164/Coding_An_Overview_and_Guide_to_Qualitative_Data_Analysis_for_Integral_Researchers]

(Helpful introduction to coding data in qualitative research.)

- Lofland, J., Snow, D. A., Anderson, L., & Lofland, L. H. (2005). *Analyzing social settings: A guide to qualitative observation and analysis* (4th ed.). Belmont, CA: Wadsworth. (Good discussion of how to conduct *unstructured* interviews.)
- Moustakis, C. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage. (Very detailed description of hermeneutic methods.)
- Palmer, C. & Bolderston, A. (2006). A brief introduction to qualitative research. *Canadian Journal of Medical Radiation Technology*, 37(1), 16-19. (Very clearly stated overview of the basic ideas in qualitative research.)
- Pauwels, L. & Hardyns, W. (2009). *Measuring community (dis)organizational processes through key informant analysis*. *European Journal of Criminology*, 6, 401-417. (Clear and compact discussion of ecological reliability – including a statistical test – and ecological validity.)
- Silverman, D. (2013). *Doing qualitative research – a practical handbook* (4th ed.). Thousand Oaks, CA: Sage. (An introduction containing separate chapters on special topics such as using computers in data analysis.)
- Smith, J. A., Harré, R., & van Langenhove, L. (1995). (Eds.). *Rethinking methods in psychology*. Thousand Oaks, CA: Sage. (An introduction with a psychological orientation that contains separate chapters on, for instance, *semi-structured* interviews.)
- Stewart, D. & Shamdasani, P. (2014). *Focus groups: Theory and practice* (3rd ed.). Newbury Park, CA: Sage.
- Yin, R. K. (2017). *Case study research and applications: Design and methods* (6th ed.). Thousand Oaks, CA: Sage.