SECTION I

Basics of Methodology in Human Sciences

CHAPTER 1

Introduction and Objectives

Goals of the Chapter

1. To obtain an overview of the importance of scientific inquiry and statistics in the human sciences.
2. To familiarize ourselves with the most common criteria for scientific knowledge.
3. To orientate ourselves and examine the research material in a critical manner.

This section handles the general basics of carrying out a research. To understand its contents, no special mathematical skills or prior knowledge about carrying out a research is required. It can also be referred to even if you do *not* have any existing knowledge of these topics. Readers who have basic research knowledge may find the topics addressed at the beginning of Volume 1 to be too trivial.

The structure of this section follows the common structure of a scientific article or enquiry; the different phases of research project are dealt with. The first phase is to choose a topic and formulate a clear research problem. This section helps learn how to carry out a critical analysis of the existing literature on the selected topic and how to define the key concepts of the enquiry. Then, the research design, sampling, and measurement instruments are addressed. Finally, one will be familiarized with the concepts of validity and reliability, both of which relate to the trustworthiness of the instruments used and thus, to the overall trustworthiness of the research. Chapters 8 and 9 describe the general features of *results* and *discussion*. The aim of this section is to prepare one for conducting an independent research and critical evaluation of other reports. The other sections of this material focus primarily on data analysis.

Structure:
- Topic
- Literature
  - Concepts
  - Models/Theories
- Problem
- Methods
  - Sample
  - Design
  - Instruments
  - Trustworthiness
- Results
- Discussion

There are a vast number of good methodological guidebooks and web sites in the field of human sciences (see, e.g., 40 books discussed in the *Preface*). These books are just examples of what is the required reading if one wants to investigate the field of the behavioral sciences and its research methods in depth. These books are usually quite lengthy and difficult to comprehend by those whose native language is not English. In many cases, naturally, one finds methodological books in the reader's own native language as well.

1.1 Basic Characteristics of Scientific Knowledge

Science corrects itself

The fundamental feature of generating scientific knowledge is that science corrects itself. This means that reliable evidences can be verified by new findings, but fallacious ones cannot be verified or will be proven incorrect. An essential feature of scientific knowledge is that new knowledge builds on the foundations that have been laid by old knowledge that is, what is already known serves as the basis for new knowledge. It is rare for a researcher to find something so exceptionally original that no research has ever been conducted on that topic.

New knowledge builds on the foundations that have been laid by existing knowledge

Another essential characteristic of scientific knowledge generation is that it must be possible to obtain the same result repeatedly. In other words, the results must lend them to replication. It means that if you have obtained a result, other researchers must be able to get the same result in their own research by using the same methods. If no one can obtain the same results as you did, it is likely that you have made an error or your results were exceptionally rare incidents.

Scientific knowledge must be capable of being replicated

Furthermore, yet another characteristic of scientific knowledge is that it is acquired using identified methods. This makes the new knowledge comparable with the existing knowledge. Scientific knowledge must be acquired using an objective approach and researchers must usually avoid employing subjective stance that can be considered as mere speculation on the topic. On the other hand, a scientific inquiry can never be free from values; a researcher's ideas of the nature of reliable knowledge and reality have a strong influence on the findings. In this sense, the perfect objectivity is an ideal: a goal to strive for, not a requirement. Scientific knowledge is not *the truth*. It is usually something much more and cannot be captured through research. Scientific inquiry can be said to consist of a multitude of methods that have been found to produce reliable results with the minimum likelihood of errors.

Methods are known
- objectivity
- avoiding subjectivity

Scientific knowledge is not free from values

Scientific knowledge is not "the truth"

Science is not infallible—all people make mistakes, even though some blindly believe in science. On the other hand, sometimes consumers are manipulated into buying things that are *scientifically* *proven to be the best* or *scientifically tested*, or *proven to be the best in their class*. In the history of science, there have been cases where the researchers deliberately manipulated the data to achieve the desired results. Sometimes, the sponsors of a study would not permit publishing the results that could harm their companies or products. Despite the proverb "Don't bite the hand that feeds you," reporting distorted results is not acceptable.

Science can be misused

Because these aberrant ways of using research exists, it is essential for researchers at every academic or nonacademic level to be able to analyze the literature, the methods used, and the results obtained in research in a critical manner. It is also important to form one's own individual opinion about the trustworthiness of an enquiry. The primary requirement is the critical attitude toward one's own research, but this attitude should also extend to reports by other researchers.

A critical attitude is a virtue

1.2 Special Characteristics of the Human Sciences

In the human sciences, the focus of inquiry is humanity from a wide perspective. Educational research focuses on the essence of education and in the nursing sciences the focus is on the essence of nursing in its broader sense. This material, however, does not concern the essence of either education or nursing, as these topics have been discussed in other books. Roughly speaking, one can divide research in the human sciences into five fields: (a) research concerning the subject/object of an action (e.g., a student or client/patient), (b) research concerning the primary actor (e.g., a teacher or a nurse), (c) research concerning an action (e.g., educating or nursing), (d) research concerning the results of the action (e.g., learning or caring outcomes), and (e) research concerning the broad nature of those things which modify or restrict the action, or which explain the results.

Five types of research:

There is no methodology that is only typical for the human sciences. Knowledge in the fields of education and nursing is acquired using the exact same methods as those employed in other branches of science. Our goal is to find general models or rules to understand the phenomena of education or nursing. A delicious flavor is added to the human sciences when one thinks of human beings as thinking, intentional, motivational, attitudinal, image-making, and goal-oriented actors; all of these things affect our results. This is why the human sciences are not only both interesting and challenging but also vulnerable and sensitive. Is it even possible to find any stable or consistent model or theory which could explain the phenomena in the field of, for example, education or nursing?

Humans are an interesting target for study

Is it possible to create a model of a human?

SECTION II

Basics of Test Theory and Test Construction

CHAPTER 10

Introduction and a Brief History of Test Theory

Goals of the Chapter

1. To orientate to learn the terms of test theory.
2. To recognize some remarkable individuals of test theory.
3. To get a general picture of the history of test theory and statistics.

In the first few pages, there is no need for an extensive knowledge of mathematics to understand the text. Further on, the amount of mathematical expressions increases. This may cause some difficulty for a person who does not have the basic knowledge of statistics or who cannot take some time to learn the new terms and notations. The learning process may be the same kind as in learning different languages. It is almost impossible to teach the principles of test theory without some mathematics. Thus, a reader without a mathematical background should not be scared of the exotic notation. For the novice reader, it is important to make a difference between the *population parameter* which is denoted by Greek letters (e.g., μ for the population mean), and an *estimate* of this parameter (e.g., $\bar{X}$ for the sample mean). For example, the variance of the sum variable X is denoted by the symbol σ_X^2, when speaking of the theoretical population variance, and by the symbol s_X^2 when estimating the value based on the sample. Also, such symbols as the correlation (ρ_{gX} and r_{gX}) and the true value (τ_{ag} and T_{ag}) will become more familiar. If necessary, the reader can turn to the last pages and the notation list.

When preparing the text, I have tried to think of a reader without a statistical degree. This chapter introduces some of the leading contributors in the history of test theory. In Chapter 11, different kinds of measurement types are discussed. In addition, the basic principles are discussed, with which one can construct good items for a test. Chapter 12 concentrates on the starting phase of the test construction. Here, some preliminary ideas of test construction are applied to a real situation.[1] Chapters 13 and 14 handle the validity and reliability of the test and the measurement from the classical test theory viewpoint. Chapter 15 handles item analysis from the parametric[2] viewpoint, that is, from the IRT modeling viewpoint, however, quite superficially. Chapter 16 concentrates on the principles of item banking, Chapter 17 on a practical example of the development of a large test battery, and Chapter 18 on an example of the development of a new item bank.

There are several good books concerning test theory. A classic book is Gulliksen (1987[1950]), where the most advanced results of its own time were collected into one book. Another good—but heavy—book, especially for parametric models, is one by Lord and Novick (1968). More recent papers can be seen in the list of references.

10.1 Turning Years in the History of Test Theory

Beginnings of Test Theory: 1900-

The beginnings of test theory can be traced as far back as 1890, when J. McKeen Cattel was trying to create new ways of predicting the achievement of college students in the USA. According to Gulliksen (1987[1950], 1), researchers were very disappointed when they found out that the correlations between test scores and grades were almost zero—maximum in the class of 0.19. At the beginning of the century (1904–1913), Charles Spearman developed—simultaneously with Karl Pearson—some of the basic formulae of quantitative research (one might remember the Spearman rank-order correlation ρ or Spearman-Brown prophesy formula and Pearson product-moment correlation coefficient r_{xy}).

Early development in 1920–1950

During 1920–1950, the British statistician Ronald A. Fisher gave a strong impetus to the improvement of statistics; according to Cronbach, Gleser, Harinder, and Rajaratnam (1972, 1), Fisher radically renewed statistical thinking. His influence is today seen in the p-values, and χ^2- and F-distributions (Fisher's daughter has said that it took years for her father's and the inventor of Student's t-distributions time to compute the values of distributions—this may be the reason why we are still using p-values at three different levels.). In the 1930s, Kuder and Richardson created several well-known indices for assessing reliability; coefficients like KR20 and KR21 are from that time. Lee J. Cronbach himself developed some

[1] Preliminary ideas are introduced also in Chapter 7 in Section I.
[2] Of the concept and using parametric distributions, see closer Section VI.

features of the well-known *Alpha* coefficient at the beginning of the 1950s and concentrated later on the generalizability theory. From the 1970s onward, the great popular concept has been SEM that has also contributed to the construct validity of the measurement instrument; the Swedish statistician Karl G. Jöreskog developed a maximum likelihood estimation that has been heavily used, especially in confirmatory factor analysis.

1970 on: SEM in assessing the construct validity

One of the fathers of modern psychometrics and test theory is L. L. Thurstone, who affected in the 1930s and the 1940s. His student, assistant, and later colleague H. Gulliksen wrote the first textbook on test theory in 1950. To some extent, this section is based on that classic book. Nevertheless, more recent publications are also referred to—in the area of classical test theory, the facts have remained the same. At the end of this section with parametric models of IRT, the essential developers were the Danish mathematician G. Rasch (one-parameter logistic model [OPLM] 1960) and American statisticians A. Birnbaum, F. M. Lord, and M. R. Novick (two- and three-parameter logistic and the normal ogive model). Lord especially published a remarkable amount of scientific papers concerning IRT in the 1950s and the 1960s. In 1968, Lord and Novick wrote the heavily cited classic book on test theory. Later literature represents the developments of partial credit models, where the classical IRT is also enlarged to handle other than dichotomous (i.e., which can have only two values—like correct and incorrect) items (e.g., OPLM software created by CITO, the national testing organization in the Netherlands, see Verhelst, Glas, Verstralen, 1995). A good article is also the work of Hambleton (1993), even though the text does not cover the very latest advances of the IRT. One can read some of the later advances from Stout (2002).

Father of modern psychometrics and test theory

Gulliksen's book is a classic

Rise of IRT-modeling

10.2 New Winds in the Test Theory

The last 50 years have been triumphant ones for the IRT modeling or, specifically, Rasch modeling, in relation to the Classical Test Theory (CTT) or the Classical Item Analysis (CIA; Gulliksen, 1987[1950]). Contributions to the development of the IRT modeling have been made by laboratories and universities all over the world since the days of Lord (1952), Rasch (1960), and Lord and Novick (1968). Some of the most active contributors appear to be the Educational Measurement Lab in the University of Illinois at Urbana–Champaign (see, among others, Stout, 2002), MESA (Measurement, Evaluation, Statistics, and Assessment) Laboratory in the University of Chicago (see, among others, the works of Linacre, Smith, and Wright,[3] CITO

IRT modeling has, in practice, superseded the classical item analysis in 50 years

[3] See http://www.rasch.org/publish.htm (accessed on April 27, 2016).

SECTION III

Basics of Qualitative Research

CHAPTER 18

Introduction to the Qualitative Research

Goals of the chapter

1. To orientate to learn the terminology of qualitative research.
2. To sketch the structure of the section and to orientate in the contents.
3. To understand that the qualitative and the quantitative research are not exclusive.

Researcher: -Around the world, an attempt is being made to examine what kind of man really is.

Prime Minister: -Where does such survey truly aim?

Researcher: -Man is tried to get behaving according to the research results.

—Carsten Graabaek, *The Prime Minister*

Question 1. What is the Purpose of the Research?

Advanced reader of the section consisting the *basics* of the qualitative research methodology soon detects that the handling is wide but superficial and many issues are not handled at all. To support the reading, it is crucial that the reader should also have an opportunity to attend practical exercises, which would be an opportunity to deepen the concepts that are presented in the section more practically.

Reading is not enough. One needs practice, too

The outline of this section is based on the *Handbook of Qualitative Research* (Denzin & Lincoln, 2000; 2011 ; Metsämuuronen, 2006d), from where some articles are also used. In Chapter 19, the ideas and territory of qualitative research are compared to the elder sibling: the tradition of the quantitative research. In Chapter 20, some strategies of the qualitative research are presented: case study, phenomenology, ethnography, phenomenography, the grounded theory method, action research, and discourse analysis. In Chapters 21 and 22, some methods for acquiring knowledge and for the analysis of qualitative material are presented: interview, observation, the analysis of written material, and computer-aided analysis. Chapter 23 addresses reporting the qualitative results. Finally, Chapter 24 presents different possibilities to combine qualitative and quantitative research. At the same time, the concept of *triangulation* is considered.

On top of Denzin and Lincoln (2011), there is a great deal of literature of the area available in English—surely, native language materials are also available. Some materials are collected in the appendix of this section (provided at the end of the section), where 40 living textbooks of qualitative research methodology are compared on the basis of their contents. Some of the general books with wide contents are Carson, Gilmore, Perry, and Gronhaug (2005), Creswell (2006), Flick, Kvale, and Angrosino (2007), Hesse-Biber and Leavy (2010), Lichtman (2009), Lindlof and Taylor (2011), Mariampolski (2001), Marshall and Rossman (2010), Padgett (1998), Patton (2001), Seale, Gobo, Gubrium, and Silverman (2004), and Silverman (2010). On top of these general books, there are also more specific publications covering, maybe, only one strategy (such as case study, ethnography, or phenomenology), the data acquiring method (such as an interview or participation method), or the analysis methods (such as content analysis). The latter possibilities are referred to in the chapters to come.

Our goal is to find 'truth' from the phenomenon studied

When conducting research, qualitative as well as quantitative, it is essential to ask as Pilate asked 2000 years ago: *What is truth?* When searching for truth, it is not essential whether we draw the nearest with the qualitative or quantitative methods; what is essential is to reach the truth as close as possible. For this reason, a chapter is included at the end of this section in which the combination of two research traditions is discussed.

A good qualitative survey requires as much the work as a good quantitative survey–maybe even more

The purpose of this main section is not to frighten a beginner researcher, but rather to inspire to look for and to find new knowledge by using qualitative methods. Conducting a good qualitative survey is just as difficult as conducting a good quantitative or statistical survey. At the same time, I would say that it is easier to conduct a *low*-quality qualitative study than a low-quality statistical survey. Why? In the methodology of statistical survey, diverse methods have been developed to avoid wrong inferences. Instead, the qualitative process of research is intensively based on the researcher's own intuition, his/her own interpretation, the sense of reasoning, convergent skills, and classification abilities; there are a great number of ways to make conclusions—even inconsistent—for the same material. In this section, it is attempted to present generally acceptable ways to conduct a qualitative study so that another researcher would make the same conclusions from the same data.

SECTION IV

Basics of Futures Studies

CHAPTER 25

Introduction to Futures Studies

Goals of the Chapter

1. To get a general view of the basic assumptions of futures studies.
2. To preliminarily become familiar with the nature of changes occurring in time.
3. To orientate with studying methods, catching the futures.

Futurology is a scientific domain which is interested in the probable, possible, and preferred futures. The course description of the University of Turku[1] explains the futures studies as follows:

> *As an academic field of research, Futures Studies generates knowledge of the ways and processes of how individuals and organisations deal with the uncertain future.*

Does Futures Exist?

The central studying object in the futures studies is *present*. All of our observations are based on the present (or in the discipline of history, the *past* observed from the *present*). On the other hand, we must notice that although future does not really exist—it always slips away from our hands—it *does* exist, because we can imagine it. Therefore, the future affects us the same way as a thought of a (imaginary) lottery win may make us experience both psychological (joy, happiness, excitement, and exhilaration) and physiological effects (cold shivers, raising hair, filling the lottery coupon, etc.).

Alternative Futures

Futures are characterized by alternatives. They are a set of different *possible* futures. Hence, we usually use the plural form *futures*, when discussing about futures studies. According to some writers (e.g., Mannermaa, 1991), periods of transition and uncertainty make

[1] See http://www.utu.fi/en/units/ffrc/publications/Pages/home.aspx (accessed on April 28, 2016).

possible circumstances for the so-called bifurcation or trifurcation, which can be visualized with Figure 25.1:

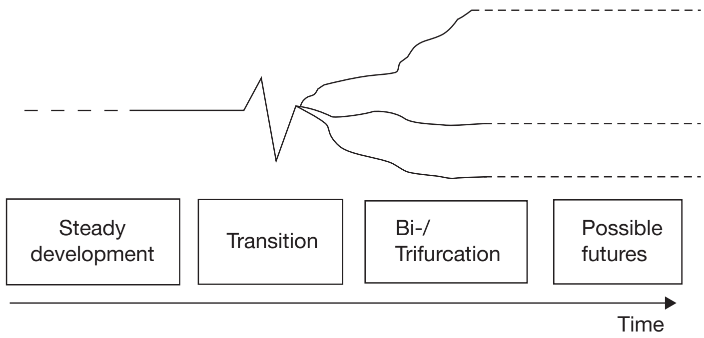


Figure 25.1 Trifurcation and Futures

One of the paradigms of futures studies is evolutionary futures studies

According to this thinking, development would lead from steady time periods to the transition phase and from there to the different potential states of futures (in this case, three different possibilities). The closest analogy given to the development visualized in Figure 25.1 is evolution. This idea is central to the so-called evolutionary paradigm on futures (Mannermaa, 1991). Though the paradigm is quite clear, it created some discussion in the futures studies society when published because as it may embed some weaknesses: one of them is that although transitions make it possible for potential futures of a different kind, only one is really happening. Therefore, one may ask, does real bifurcation or trifurcation exist? In any case, nowadays, the very popular—and popularized—chaos theory gives mathematical possibilities to study nonlinear trends, transit, or *chaos*. Another central idea of evolutionary futures studies is the idea of development. In reality, one cannot speak exactly of *development* in this context, but probably more of *change*.

If futures do not exist, is it possible to acquire the accurate information of it? Obviously, it is not. On the other hand, futures can be caught with methods giving more or less certain and reliable information about the futures. To use these methods requires some level of a preliminary view of the scientific basics of futures studies.

Finnish researchers have been quite active in the field of futures studies.[2] All in all, Finns seem to be very active in the

[2] See the lists of publications of, for example, Finland Futures Research Centre with 65 researchers at http://ffrc.utu.fi/en/publications/ (accessed on April 28, 2016).
Mika Mannermaa at http://www.kirjaseuranta.fi/kirjasto/M/Mannermaa_Mika/ (accessed on April 28, 2016).
Tuomo Kuosa at http://tuomo.tel/ (accessed on April 28, 2016).

futures thinking, as Metsämuuronen, Kuosa, and Laukkanen (2013) noted. Several institutions for futures thinking have been established in Finland, such as a permanent Committee for the Future in the Finnish Parliament,[3] the Finnish National Fund for Research and Development (SITRA),[4] and the Finnish Funding Agency for Innovations, TEKES (FinnSight, 2015). Also, the main actors in the governmental field—for example, the Ministry of Education and the Finnish National Board of Education—have their own futures thinking activities (see Metsämuuronen, 2006c).[5] During recent years, a group of good basic books has appeared in Finnish for the support of futures studies. The older classic *How Do We Research Futures* (Vapaavuori, 1993) has published its second renewed edition (Vapaavuori & von Bruun, 2003). Also, the massive *Futures Studies, Basics and Applications* edited by Kamppinen, Kuusi, and Söderlund (2003) has got a second renewed edition. Additionally, *How Do We Research Futures* from the Future Studies Network Academy (Tulevaisuudentutkimuksen VerkostoAkatemia, 2004) is representing newer literature. In Finland, the scientific *Journal of Futures Studies* (*FUTURA*) has been also been published for tens of years. During the years 1998–2000, it turned toward a more scientific format under the editorial of the writer of this material.

Of course, there are a large number of good books of futures studies in English, too. Some of them have been cited further in the text. Some classical ones and some that are worth citing are, for example, the classics Armstrong (1985, 2001) Bell (1997, the latest printing 2011), Meadows, Meadows, Randers, and Behrens (1972), Meadows, Meadows, and Randers (1992), Naisbitt (1982), and Naisbitt and Aburdene (1990), and some more recent or otherwise powerfully shaking volumes of Aaltonen, Barth, Casti, Mitleton-Kelly, and Sanders (2005), Aburdene (2005), Masini (1993), and Masuda (1980)—to mention a few. An interested reader may also benefit by scanning journals such as the *FUTURA*, *Futures*, or *Technological Forecasting and Social Change*.

Jari Kaivo-oja at https://www.utu.fi/fi/yksikot/ffrc/yhteystiedot/henkilokunta/Sivut/jari-kaivo-oja.aspx just few to mention (accessed on April 28, 2016).

[3] See their publications at http://web.eduskunta.fi/Resource.phx/parliament/committees/future.htx

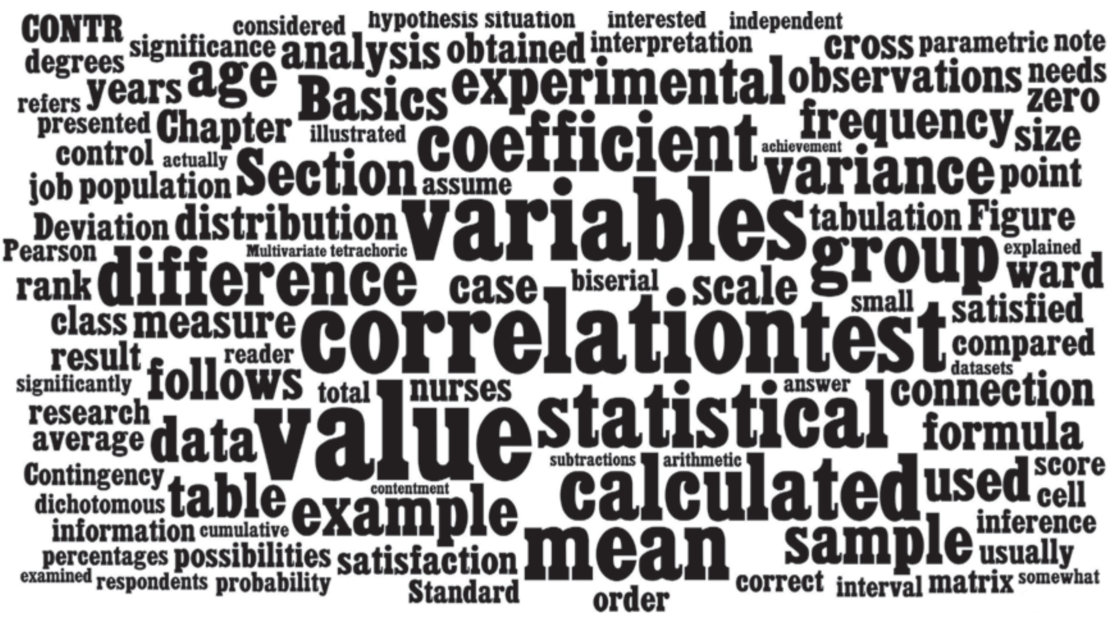
[4] Exhaustive list of publications is available at http://www.sitra.fi/en/archive (accessed on April 28, 2016).

[5] See, for example, http://www.minedu.fi/OPM/Tiede/julkaisulistaus?lang=en (accessed on April 28, 2016). and http://www.oph.fi/english/publications (accessed on April 28, 2016).

Some of the readers might wonder why this chapter was placed between qualitative research and quantitative research. I see it just the way I represent it: in futures studies, both quantitative and qualitative methodologies are used. Many of the methods are purely qualitative; hence, I have placed this section just between the two approaches.

SECTION V

Basics of Statistical Description



CHAPTER 29

Introduction to Statistical Analysis

Goals of the Chapter

1. To orientate to the statistical analysis of the collected data.
2. To orientate to the contents of the section.

Section I of this book addresses the basics of doing a research, by becoming familiar with things, such as defining concepts, choosing a theory, creating a measurement instrument for the analysis, sampling, and collecting the data. In Section II, the reader gains a deeper understanding on how to work out the construct of an accurate measurement instrument for human sciences. Now, it can be assumed that the measurement instrument has been tested and has been proven to be good: it is measuring the very things it is supposed (and, thus, the information is valid) to as accurately as possible (and, thus, the information is reliable). One has reached the point in the study in which quantitative data have been collected. Now, what to do with the collected data and what conclusions can be drawn from it? In this section, the simple basic things are discussed. In Section VI, the analysis of quantitative data will be deepened in the area of making statistical inferences on the basis of the sample. However, the topic is briefly handled already in this section as advanced material. Even if the reader has an understanding about the basics of doing research, she/he may learn something new from, for example, the section that addresses the coefficients of correlation.

Much more can be acquired than this introductory section of the basics of statistical description. In Section II of Volume 2 as well as Volume 3, some advanced, or more complex, statistical methods will be introduced. In Volume 2, Section II focuses on methods where many variables are analyzed at the same time. In Volume 3, Section I focuses especially on the statistical analysis of small datasets. Section II focuses on true and quasi-experiments and, hence, on Analysis of Variance (ANOVA) and its many variations. Section III focuses on analyzing the datasets that are collected with clustered samples of different organizations—as is usually the case in student assessments collected from the schools. Section IV focuses on SEM analysis. Finally, Section V focuses on the analysis of datasets with dropouts.

The purpose of this section is to teach the basic analysis of data by analyzing the collected data through simple analysis. It is not intended to go into much detail. Instead, the goal is that the reader will gain increasing confidence in using basic statistics and carrying out basic analysis. The basic assumption is that the reader has no initiative readiness to do research. The necessary things are explained in a simple fashion and, thus, the reader barely needs any initiative mathematical skills. Chapter 30 discusses the basics of the data. Chapter 31 introduces the computing of basic statistics. Chapter 32 focuses on the constructing and analyzing of cross-tabulation. In Chapter 33, one becomes familiar with the concepts of *connection* and *correlation* and, especially, the *correlation coefficient*. It will be important to learn how to interpret the correlation coefficient. In the final chapter, one learns how to compare between two means and the tests related to the differences between two means.

Compared with the other sections of the volume, somewhat less references are given in this section. The history of basic statistics goes beyond thousands of years—it is not a convention to refer to those ancient masters who created formulae for means or variances, for example. However, when willing to enlarge the knowledge of statistical description, there are lots of good basic textbooks available. In the Preface section, some 40 textbooks were compared on the basis of their contents. Textbooks such as Ary et al. (2010), Babbie (2010), Bernard (2006), Malhotra (2009), Punch (2009), Wallen and Fraenkel (2001), and Wiersma and Jurs (2009) contain much of the same kinds of topics as this section. Punch (2009) especially, contains almost all of the first-level headers as this section. It is always profitable to compare different textbooks and the way they introduce their matters—all writers have found their own way to illustrate the topics. If the things are not opened straight with one textbook, the other viewpoint of another textbook may help to gain the whole picture of the problem.

CHAPTER 30

Data

Goals of the Chapter

1. To become familiar with the central concepts of processing data.
2. To be able to preliminary process the missing data units.
3. To receive abilities to tabulate own data.

For the beginning, one will become familiar with a hypothetical set of data, which concerns implementing a philosophy called primary nursing in an imaginary hospital. In the example dataset, there are two wards of the same hospital, where one of the wards began actively training nurses to a new philosophy, primary nursing, but in the other ward, no conscious changes were made. The idea was to study how the work contentment changes at the experimental ward.

A few variables depicting background information are collected, using the following form:

1. Ward ☐☐☐☐☐

2. Age ☐☐

3. Your vocational training

1. Specially trained nurse
2. Midwife
3. Nursemaid
4. Mental nurse
5. Practical nurse
6. Registered nurse
7. Public health nurse
8. Other, please, specify ____________

4. Your work experience in this ward

years ☐☐ months ☐☐

VAS = Visual Analogue Scale

The main questionnaire consisted of the set of questions of job satisfaction with the VAS[1]-type instrument as follows:

5. How satisfied are you with the following factors in your present job?
Please answer by marking a cross (X) on the line on the location that best describes your satisfaction with the following statement.

How satisfied are you with...

5.1. your opportunities to use all your knowledge and skills in your work?

Not at all satisfied ⊢――――――――⊣ Completely satisfied

Attitude testing with a Visual Analogue Scale

5.2. your opportunities to develop the knowledge and skills needed in your work?

Not at all satisfied ⊢――――――――⊣ Completely satisfied

5.3. your possibilities of seeing the results of your work?

Not at all satisfied ⊢――――――――⊣ Completely satisfied

5.4. your possibilities of making independent decisions in your work?

Not at all satisfied ⊢――――――――⊣ Completely satisfied

5.5. your possibilities of acting independently based on the decisions you have made?

Not at all satisfied ⊢――――――――⊣ Completely satisfied

5.6. your possibilities of influencing how things are made in your unit?

Not at all satisfied ⊢――――――――⊣ Completely satisfied

5.7. your possibilities of taking initiative improving your work?

Not at all satisfied ⊢――――――――⊣ Completely satisfied

5.8. your possibilities of creating undamaged wholes in your work?

Not at all satisfied ⊢――――――――⊣ Completely satisfied

5.9. in your work considered as a whole?

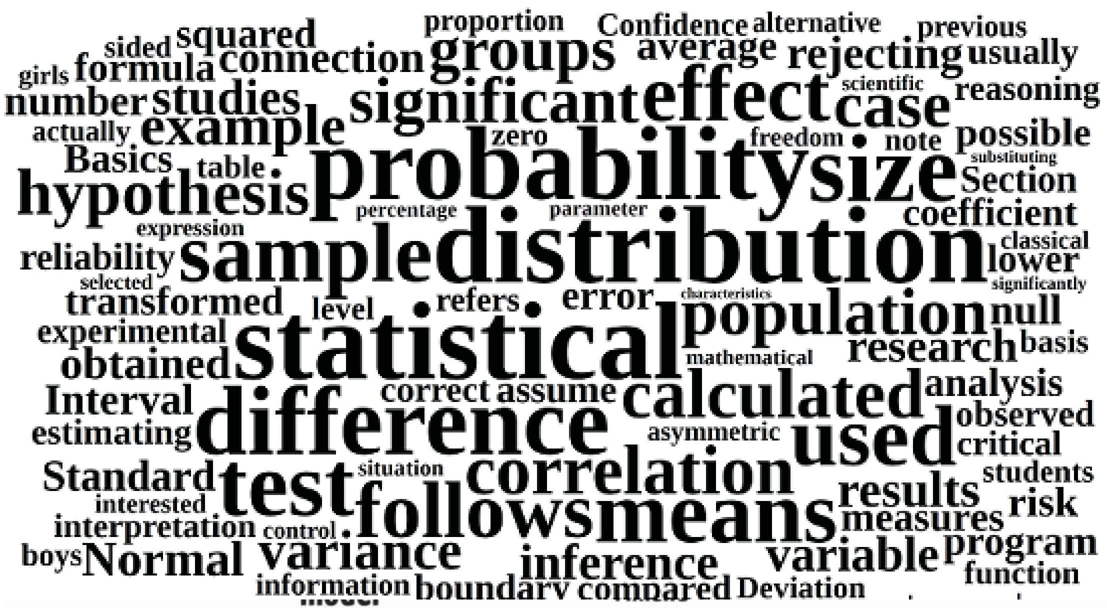
Not at all satisfied ⊢――――――――⊣ Completely satisfied

In the original form, the VAS line for responding to was exactly 10-cm long. Altogether 30 booklets—15 from the control group and 15 from the experimental group—were collected. In this section, the means and test statistics concerning the data are computed manually. In Section I of Volume 2, the statistics are computed with the SPSS program and some relevant issues—such as the estimation of reliability— are enlarged, too (see also Sections I and II).

[1] If the VAS type of measuring was not familiar, more introductory text can be found in Chapter 7 (7.1.3) in Section I.

SECTION VI

Basics of Statistical Inference



CHAPTER 35

Introduction to Statistical Inference

Goals of the Chapter

1. To orientate with studying the statistical inference.
2. To understand the concepts inductive, deductive, and abductive.
3. To know what the hypothetic-deductive method means.
4. To be able to assess the criteria for cause and effect.

The scientific research can be defined—as Kerlinger and Lee (2000, 14) do—as a systematic, controlled, empirical, and critical examination of the assumed connections of natural phenomena. When thinking about assumed connections, it is important to know how these phenomena are connected with each other. It is not sufficient just to say that there is a connection. In many cases, it is more important to know which is the direction of the connection, that is, which phenomena is the cause and which is the effect. For example, when knowing that there is a strong connection between smoking and appearance of yellow color on the fingers, one cannot say that the yellow fingers are the cause of smoking but the very opposite. Of course, the causal inferences of the phenomena are not always this easy. Do the hobbies of children originate from the parents' interests or vice versa? Is sportiveness caused by good physics or vice versa? This logic of cause and effect will be discussed deeper in section 35.2. In an introductory manner, three basic mechanisms of making scientific inference will be introduced in section 35.1: the classical *inductive* and *deductive* reasoning and more modern *abductive* reasoning, which connects two previous ones. The basic characteristics of probability are then handled in Chapter 36.

One definition of scientific research

It is not always easy to infer what is the cause and what is the effect

All statistics contain error

Another view to statistical inference comes from the fact that it is quite usual to infer something in the population on the basis of a sample drawn from that population. When, for example, one obtains the means of males and females in the sample, it is a convention to think that these values also reflect the values in the population. More, it is a convention to think that the differences between the males and females in the sample reflect a true difference in the population. This is not, however, evident, because all statistics contains a random error due to the sampling procedure. The difference which supersedes this random error is taken as a real effect. These matters are deepened in section 35.3 and Chapter 37 onward.

35.1 Induction, Deduction, and Abduction

Induction
= from single to general
Deduction
= from general to single

Generally speaking, in the inductive reasoning, one infers general laws on the basis of single cases and in deductive reasoning, on its behalf, on the basis of general laws, one infers special cases. The division is not always this simple (Niiniluoto, 1983, 29–30), but the separation may help to understand the differences of inference processes. Both the type of logic is needed and they have their own place in scientific reasoning. Haaparanta and Niiniluoto (1995), among others, say that inductive reasoning (i.e., to infer a general law on the basis of single cases) does not necessarily lead to the correct truth—however, it can enlarge our understanding of the matter. One can think, for example, the theories of motivation, discussed in Section I, or Piaget's models of intellectual growth in childhood, which may have been found on the basis of some single cases.

There are several types of induction, of which *direct reasoning* may be the most interesting when it comes to statistical inference. When one can assume a technique—widely speaking—to eliminate fault alternative hypotheses, it is called eliminative induction. One such technique can be to use a general statistical fact. To measure the correctness of the inference, one uses information that is obtained from the sample. The following example describes the situation:

- The average age of nurses in hospitals is 41.5 years (a statistical fact).
- A is a randomly selected sample of all nurses working in hospitals.
- In sample A, the average age of nurses is 41.5 years.

What if the average age in the sample is *not* 41.5 years? Logically, the average age should have been 41.5, but it was not. The premise, that is, the output assumed beforehand, could be expressed in a slightly different way (the numbers are not factual but imaginary):

Premise = basic assumption

- With 95 percent probability, the average age of nurses in the hospitals ranges 40.5—42.5 years.
- A is a randomly selected sample of all nurses working in hospitals.
- In sample A, the average age of nurses ranges from 40.5 to 42.5 years.

Again, one can ask the same question: what if the average age in the sample is *not* between 40.5 and 42.5 years? It seems evident that in the example, the sample does not represent the population of nurses working in the hospital.

The deductive reasoning, that is, inferring some single consequences on the basis of a found law (of nature), holds *the truth*. In what follows, the testing of hypotheses is based on deductive logic: the null hypothesis carries *the truth* and it is handled as the law of nature. In the classical deductive reasoning process, the logic can be as follows:

Deductive reasoning holds the truth

- All humans die. [A single case, premise I]
- The teacher is a human.[A single case, premise II]
- The teacher is mortal. [Inference]

Classical deductive reasoning

In both Niiniluoto (1983) and Haaparanta and Niiniluoto (1986), the *hypothetic-deductive* method has been handled considerably widely. Kerlinger and Lee (2000, 15–16) use the term *reasoning-deduction* of the same phenomenon. The idea is that the researcher seeks and finds a hypothesis as the ground for his/her study—such as *Commitment to the studies makes the learning outcomes better*. After defining the concepts, the researcher reasons as follows: if the commitment to the studies really makes the learning outcomes better, it has to be so that within the more committed students, one should find a higher amount of those with higher achievement than within the group of less committed students. It is good to note that the hypotheses do not come *from air* but they are based on previous results or general *truths*.[1]

Setting hypotheses is based on deductive reasoning

In the hypothetic-deductive method, the hypothesis is tested by the data. If the data speak against the hypothesis, the hypothesis

[1] Another viewpoint to these kinds of hypotheses and theories comes from the Grounded theory method(ology), and some other qualitative approaches: justifiably one can ask where the *theory* emerges? One of the strengths of the qualitative approach is that they may produce empirically created theories to be tested and validated. These methods are handled closer in Section III.

Hypothetic-deductive reasoning: Hypothesis will be rejected, if it does not get support from the data.

has to be rejected. On the other hand, if the data give support to the hypothesis, the hypothesis *may* be correct.

There are several criteria for the hypotheses. Haaparanta and Niiniluoto (1986) list the following five:

Criteria for hypotheses:

1. Hypothesis should explain the facts and regularities underlying the research.
2. Hypothesis should be logically coherent, explicitly formulated, and should fit the previously accepted theories.
3. Hypothesis should be testable.
4. Hypothesis should be informative.
5. Hypothesis should be as simple as possible.

Previously, the topic of setting the hypotheses has been handled in a less technical manner and the list of criteria, or the latter discussion, does not change what has already been said.[2] This section concentrates on testing hypotheses and, thus, several examples of the topic are used to illustrate the matter in Chapter 39.

Abductive reasoning

Abductive reasoning is placed here between inductive and deductive reasoning. When in the deduction, one generalizes accepted assumptions or laws (of nature) to infer something of single cases; in the induction, one draws general conclusions from single cases; in the abduction, the researcher—on the basis of collected data—finds or orders such combinations of the phenomena of which there were no argued explanation or rule in the previous knowledge base (Reichertz, 2007, 219). The term *abduction* can be traced as far back as 1597, when Julius Pacius used the term in translation for the Aristotelian concept *apagoge*. The meaning of abduction as it is used today originates from Pragmatist Charles S. Peirce. Peirce's *three-stage discovery procedure* starts with the abductive phase, where the tentative hypothesis for a scientific discovery or theory is found. In the second phase, on the basis of this hypothesis, the researcher makes the predictions of the details to be found during the research. In the third, inductive phase, one seeks the facts to support the preliminary hypothesis. If no (or enough) facts can be found, the process starts again from the beginning (Reichertz, 2007, 222). Practically thinking, the abductive reasoning means that, in the first phase of the scientific inquiry, the researcher tries to use all of the available information to form the tentative hypothesis.

Researcher tries to utilize all information available

Though abductive reasoning is meaningful in all scientific inquiries, it is the most valuable in the phase when the phenomenon is

[2] The topic is handled in Chapter 5 in Section I.