Part I

INTRODUCTION

This textbook requires a multi-layered view of 'statistics in psychology'. Within the Bachelor's curriculum it is only possible to demonstrate the correct use of the most important techniques. For the Master's curriculum, however, a certain understanding of these methods is necessary: for the Master's thesis, where usually a scientific question has to be worked on single-handed but under supervision, the student has to refer to statistical analyses in literature concerning the topic, and if necessary to improve the choice of the method used for analysis. For doctoral studies, understanding alone is not enough; a willingness to reflect critically on the statistical methods must be developed. The statistical methods used in the doctoral thesis, which means the entrance to a scientific career, have to be oriented on state-of-the-art methodological developments; the ability to follow these developments requires profound knowledge as well as the aptitude to evaluate new statistical methods regarding their shortcomings.

Since even for doctoral students a repetition of the basics of statistics on an elementary level is often useful, with this book they can be picked up individually where their powers of recollection end – if necessary at the beginning of the Bachelor's education. And in contrast, Bachelor's students are often interested in the contents of a Master's curriculum or where the textbook leads. They can get a taste of that now.

Finally, even lecturers will find something new in this textbook; according to our experience 'statistics for psychologists' is not taught by professional statisticians but by psychologists, mostly by those at the beginning of their academic careers; anecdotes may at least help them didactically. These casual reflections can of course also be academically amusing for students.

Accordingly, the three to four mentioned target groups are guided through the book using distinctive design elements.

The running text, without special accentuation, is directed at all target groups. It is information essential for the further study of the textbook and its practical use – as is this introduction before Chapter 1. Also the terminology used in the book has to be conveyed in a standardized way. Finally, some contents, which should be familiar to doctoral students, are nevertheless aimed at all target groups because we think that repetition is useful.

Moreover, special symbols and labels on the outer edge of some pages signal the target group that the information is aimed at. Target groups other than the ones indicated with the symbol can skip these passages without being in danger of missing the respective educational aim.

The symbol Bachelor indicates that the material in these passages is aimed particularly at Bachelor's students since it deals only with the *Ability to Use*. The symbol Master on the outer edge indicates that here the reader finds an explanation of the underlying methods, without using a mathematical derivation that is too detailed; this is about *Understanding*. The symbol Doctor on the outer edge of the page announces that the shortcomings of the method will be discussed and that common misuses will be indicated; this is about *Critical Reflection*. Finally, the note **For Lecturers** signals didactically useful observations, entailing understanding of the respective topic in a very demonstrative way.

In order to bring all target groups together again, occasionally a *Summary* is given. At the beginning of every chapter a short description of its contents is given.

1

Concept of the book

In this chapter, the structure of the book and accordingly the didactic concept are presented to the reader. Moreover, we outline an example that will be used in several chapters in order to demonstrate the analytical methods described there.

In six sections this book conveys the methods of the scientific discipline of 'statistics' that are relevant for studies in psychology:

I. Introduction (Chapters 1 to 4)

II. Descriptive statistics (Chapter 5)

III. Inferential statistics for a single character (Chapters 6 to 10)

IV. Descriptive and inferential statistics for two characters (Chapter 11)

V. Inferential statistics for more than two characters (Chapters 12 and 13)

VI. Theory building statistical procedures (Chapters 14 and 15).

Chapter 1 explains the concept underlying our presentation of the methods. Furthermore an empirical example that will be used as an illustration in various parts of the book is provided.

Chapter 2 will demonstrate that quantifying and measuring in psychology is not only possible but also very useful. In addition we would like to give the reader an understanding of the strategy of gaining knowledge in psychology as a science; the approach however is similar to other scientific fields, which is why this book can be used in other fields too.

In Chapter 3 we will address the issue that empirical research is performed in several steps. For all scientific questions that are supposed to be answered by the study (as diverse as they might be regarding contents), exact planning, careful collecting of data, and adequate analysis are always needed.

Statistics in Psychology Using R and SPSS, First Edition. Dieter Rasch, Klaus D. Kubinger and Takuya Yanagida.

Within this context we wish the reader to realize that a study does not always have to include all the people that the research question is directed at. Out of practical reasons, most of the time only part of the group of interest can be examined; this part is usually called sample, whereas the group of interest is called the population. Chance plays an important role here. It will be shown that we have to make probability statements for the results of the statistical analysis; the probability calculus used for this is only valid for events for whose occurrence (or non-occurrence) chance is responsible. For example, a certain event might be that a specific person is part of the study in question. We will treat this topic in Chapter 4, as well as in Chapter 7. Since 'chance' often has a different meaning in everyday use as opposed to its general meaning in statistics and therefore in this book, we will point out at this early stage that a random event is not necessarily a rare or unanticipated event.

Finally, if data concerning one or more person(s) or character(s) that are of interest have been gathered within the framework of the study, they have to be processed statistically. The data in their totality are too unmanageable to be able to draw conclusions from them that are relevant for answering the scientific question. Therefore, special methods of data compression are necessary. We will deal with this issue in Chapter 5. The decision of which one of these methods is applicable or most appropriate is substantially based on the type of data: for example, whether they have been derived from physical measurements or whether they can only express *greater*/*less than* and *equal to* relations. In the latter case it is important to use methods that have been specially developed for this type of data.

Mathematical-statistical concepts are needed, especially for the generalization of study results; these will be introduced in Chapter 6. For readers who are unpracticed in the use of formulas, this chapter is surely difficult, although we try to formulate as simply as possible.

If the generalization of the study results is the aim, then a prerequisite for the use of appropriate methods is that the collected samples are random samples; information on this topic can be found in Chapter 7.

In Chapter 8 an introduction to statistical inference, in particular the principle of hypothesis testing, will be given. Because of the fact that random samples are used, it is necessary to take random deviations of the sample data from the population into account. Through hypotheses that have been formulated before data collection we try to find out as to what extent these deviations are systematic or can/must be traced back to chance. The aim is to either accept or reject a hypothesis based on the empirical data.

Chapter 9 pursues a similar objective, but this time the focus is on two populations that are compared with each other.

The implied separation between planning, data collection, and analysis is true for the classic procedure for empirical studies. In this book, however, we also want to promote a sequential approach. Thereby the gradual collection of data is constantly interrupted by an analysis. This leads to a process that looks like this: observe–analyze–observe–analyze . . .; this goes on until a predetermined level of precision is reached. This procedure is also described in Chapters 8 and 9.

Special methods are needed in studies that examine a certain character of the research unit (which in psychology often is a person or a group of persons) not only under constant conditions but also under varying conditions or when the study includes more than two populations. In Chapter 10 we cover situations where there are three or more different conditions or two or three treatment factors, with at least two values of each (treatment or factor levels).

In psychological research hardly ever is only one character used. If more than one character per person is observed, then a certain connection between them may exist; we refer here to statistical relationships. If these relationships are of interest, then the statistical methods described primarily in Chapter 11 are needed.

If there really are relationships between several characters – or if there is reason to think so – then one needs very special methods for comparing several populations. Chapters 12 and 13 describe these.

Finally Chapters 14 and 15 give an introduction into theory-building techniques that establish or test models regarding content.

The appendix of the book is split into three parts: Part A lists the data of Example 1.1 which will be illustrated below, and in part B one can find tables, helpful for some analyses; often it is faster and more convenient to look up a value than to calculate it with the help of some software. Appendix C contains a summary of the symbols and abbreviations. A complete list of references and a subject index are given at the end of the book.

Summary

We assume that empirical studies always yield data regarding at least one character. Optimally, planning takes place prior to any study. Data are used to answer a specific question. Statistics as a scientific discipline provides the methods needed for this.

The diverse statistical methods that are recommended in this book and which can be used for answering the research questions posed by psychology as a science are often only practicable when using a computer. Therefore we refer to two software packages in this book. The program package **R** is both freely accessible and very efficient; that is why we continuously use **R** here. However, since in psychology the program package IBM SPSS Statistics is still preferred for statistical analyses most of the time, it will also be illustrated using the examples. The appropriate use of such packages is not trivial; that is why the necessary procedures will be demonstrated by the use of numerical examples. The reader can recalculate everything and practice their use.

The program package **R** can be used for the planning of a study, for the statistical analysis of the data and for graphical presentation. It is an adaptation of the programming language **S** that has been developed since 1976 by *John Chambers* and colleagues in the *Bell Laboratories* (belonging to *Alcatel-Lucent*). The functionality of **R** can be enhanced through freely available packages by everybody and at will, and also special statistical methods and some procedures of C and Fortran can be implemented. Packages that already exist are being made available in standardized archives (*repositories*). The most well-known archive to be mentioned here is CRAN (Comprehensive **R** Archive Network), a server network that is serviced by the **R** Development Core Team. With the distribution of **R**, the number of **R** packages has increased exponentially: whereas there were 110 packages available on CRAN in June 2001, there were 2496 in September 2010. **R** is available, free, for Windows, Linux and Apple. With few exceptions, there are implementations for all statistical methods in **R**. With the means of the recently built package `OPDOE` (see Rasch, Pilz, Verdooren & Gebhardt, 2011), it is possible, for the first time, to statistically plan studies or to calculate the optimal number of examination objects and also to successively collect and analyze data in **R**.

The program package **R** is available for free at http://cran.r-project.org/ for the operating systems Linux, MacOS X and Windows. The installation under Microsoft Windows is

initiated via the link 'Windows', from where the link 'base', which leads to the installation website, must be chosen. The setup file can be downloaded under 'Download R 2.X.X for Windows' (where X stands for the current version number). After executing this file, one is lead through the installation by a setup assistant. For the uses described in this book all the standard settings can be applied. SPSS as a commercial product must be acquired by purchase; normally universities offer inexpensive licenses for students. More on **R** can be found under www.r-project.org, and on SPSS under www.spss.com. In order not to unnecessarily prolong the explanation of the operational sequence in **R** or SPSS, we always assume that the respective program package, as well as the file that will be used, are already at hand and open.

In **R** the input window opens after starting the program; the prompt is in red: '>'. Here commands can be entered and run by pressing the *enter* button. The output is displayed in blue right below the command line. If the command is incomplete, a red '+' will appear in the next line in order to complete the command or to cancel the current command input by pressing the *Esc* button. An instruction sequence is displayed as in the following example:

```
> cbind(sub1_t1.tab, sub1_t1.per, sub1_t1.cum)
```

or also as

```
> cbind(sub1_t1.tab,
+       sub1_t1.per,
+       sub1_t1.cum)
```

or also as

```
> cbind(sub1_t1.tab,
+ sub1_t1.per,
+ sub1_t1.cum)
```

A special working environment in **R** is the `Workspace`. Several (calculation-) objects that have been created in the current session with **R** can be saved in there. These objects include results of calculations (single scores, tables, etc.) and also data sets. A workspace can be loaded with the sequence

```
File - Load Workspace...
```

For all the examples presented in this book the reader can download the `Workspace` '`RaKuYa.RData`' from the website www.wiley.com/go/statisticsinpsychology.

Since there are more data sets in our `Workspace`, the scores of single research units/persons have to be accessed by specifying the data set with a '`$`'; for example: `Example_1.1$native_language`. A useful alternative for the access is the

command `attach()`, which makes the desired data set generally available; for example: `attach(Example_1.1)`. To minimize repetition, in the instruction sequences given throughout the book, we assume that the `attach()` command has already been run and therefore the relevant data set is active. For some examples we need special **R** packages; they must be installed once via the menu `Packages - Install Package(s)...` and then loaded for every session in **R** with the command `library()`. The installation of packages is done via the menu

```
Packages - Install Package(s)...
```

In SPSS the desired data frame can be opened via File – Open – Data... after starting the program. Then we write the instruction sequence as in SPSS handbooks; for example like this

```
Analyze
        Descriptive Statistics
                Frequencies...
```

For all examples in the book the reader can find the data in the SPSS folder 'RaKuYa' on the website www.wiley.com/go/statisticsinpsychology.

For figures that are shown as the results of the calculations for the examples, we use either the one from SPSS or the one from **R**. Only if the graphs differ between **R** and SPSS will we present both.

It is the concept of this textbook to present illustrative examples with content – that can be recalculated – from almost all subject areas concerning the planning and statistical analysis of psychological studies. A lot of the methods described in this book will be demonstrated using one single data set in order to not have to explain too many psychological problems. This will be introduced in Example 1.1.

Example 1.1 The goal is to test the fairness of a popular natural-language intelligence test battery with reference to children with Turkish native language[1,2] (see Kubinger, 2009a[3]).

The following characters were observed per child (see Table 1.1 and the data sheet in Appendix A; then see, for **R**, the respective data structure in Figure 1.1, and for SPSS the screen shot shown in Figure 1.2).

[1] Fairness is a specific quality criterion of psychological assessment methods (tests). A psychological test meets the requirement of fairness if the resulting test scores don't lead to a systematic discrimination of specific testees: for example because of sex, ethnic, or socio-cultural affiliation; see Kubinger, 2009b).

[2] The data originally applied to German-speaking countries; however, there was no socio-political difference when the data in the following analyses were interpreted as relating to English-speaking countries and some ethnic-minority groups.

[3] Due to copyright reasons the original data had to be slightly modified; therefore no deductions regarding content can be drawn from the data found in the data sheet in the appendix.

Table 1.1 The characters and their names in **R** and SPSS (including coded values).

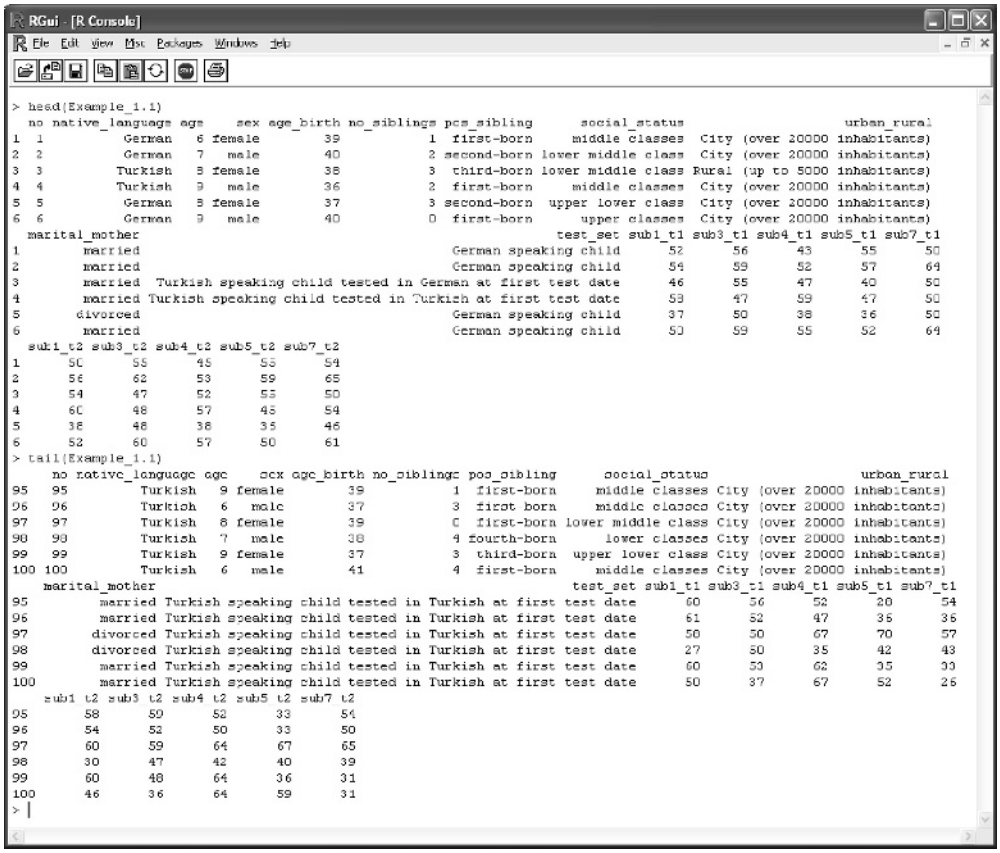
Name of the character	Name in **R**	Name in SPSS	Coded values
testee number	no	no	
native language of the child	native_language	native_language	1 = 'German' 2 = 'Turkish'
age of the child	age	age	
sex of the child	sex	sex	1 = 'female' 2 = 'male'
gestational age at birth (in weeks)[4]	age_birth	age_birth	
number of siblings	no_siblings	no_siblings	
sibling position	pos_sibling	pos_sibling	1 = 'first-born' 2 = 'second-born' 3 = 'third-born' 4 = 'fourth-born' 5 = 'fifth-born' 6 = 'sixth-born'
social status (after Kleining & Moore [1968] according to occupation of father/alternatively of the single mother)	social_status	social_status	1 = 'upper classes' 2 = 'middle classes' 3 = 'lower middle class' 4 = 'upper lower class' 5 = 'lower classes' 6 = 'single mother in household'
urban/rural	urban_rural	urban_rural	1 = 'city (over 20 000 inhabitants)' 2 = 'town (5000 to 20000 inhabitants)' 3 = 'rural (up to 5000 inhabitants)'
marital status of the mother	marital_mother	marital_mother	1 = 'never married' 2 = 'married' 3 = 'divorced' 4 = 'widowed'

[4] The gestational age is the age of the (unborn) child counted from the day of supposed fertilization.

Table 1.1 (*Continued*)

Name of the character	Name in **R**	Name in SPSS	Coded values
test setting	test_set	test_set	1 = 'German speaking child' 2 = 'Turkish speaking child tested in German at first test date' 3 = 'Turkish speaking child tested in Turkish at first test date'
Everyday Knowledge, 1st test date (*T*-Scores)[5]	sub1_t1	sub1_t1	
Applied Computing, 1st test date (*T*-Scores)	sub3_t1	sub3_t1	
Social and Material Sequencing, 1st test date (*T*-Scores)	sub4_t1	sub4_t1	
Immediately Reproducing – numerical, 1st test date (*T*-Scores)	sub5_t1	sub5_t1	
Coding and Associating, 1st test date (*T*-Scores)	sub7_t1	sub7_t1	
Everyday Knowledge, 2nd test date (*T*-Scores)	sub1_t2	sub1_t2	
Applied Computing, 2nd test date (*T*-Scores)	sub3_t2	sub3_t2	
Social and Material Sequencing, 2nd test date (*T*-Scores)	sub4_t2	sub4_t2	
Immediately Reproducing – numerical, 2nd test date (*T*-Scores)	sub5_t2	sub5_t2	
Coding and Associating, 2nd test date (*T*-Scores)	sub7_t2	sub7_t2	

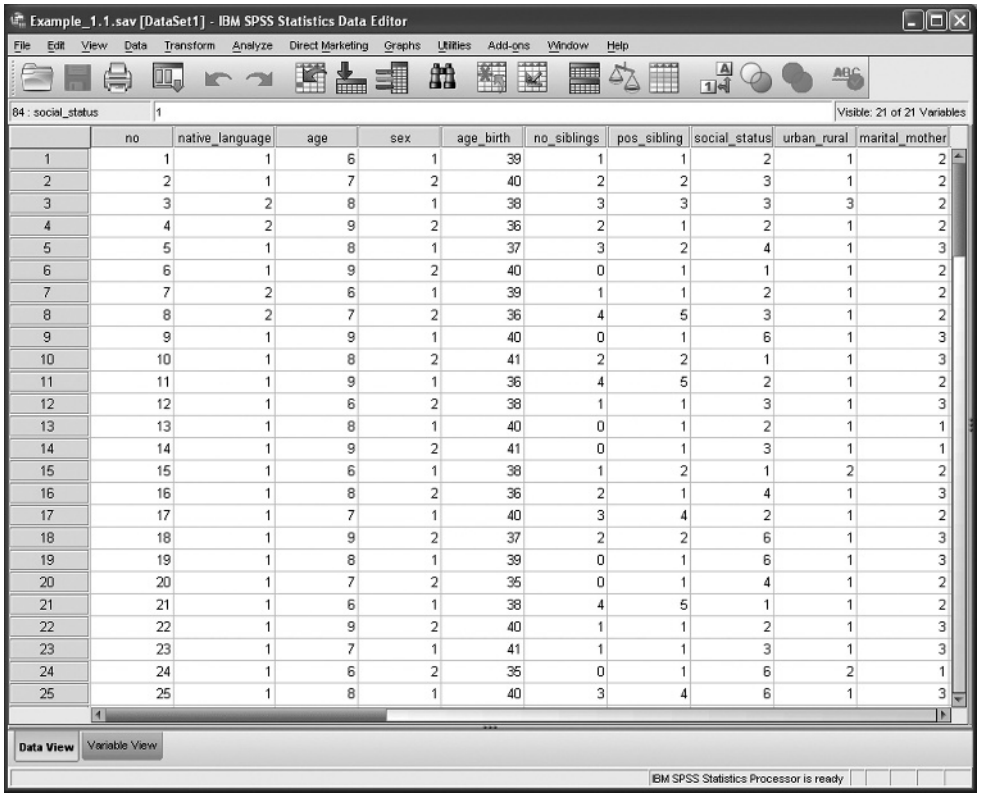
[5] Test scores are generally standardized to a certain scale; *T*-Scores are a very common method of standardization.




```
> head(Example_1.1)
  no native_language age    sex age_birth no_siblings pos_sibling      social_status                    urban_rural
1  1          German   6 female        39           1  first-born     middle classes  City (over 20000 inhabitants)
2  2          German   7   male        40           2 second-born lower middle class  City (over 20000 inhabitants)
3  3         Turkish   8 female        38           3  third-born lower middle class Rural (up to 5000 inhabitants)
4  4         Turkish   9   male        36           2  first-born     middle classes  City (over 20000 inhabitants)
5  5          German   8 female        37           3 second-born  upper lower class  City (over 20000 inhabitants)
6  6          German   9   male        40           0  first-born      upper classes  City (over 20000 inhabitants)
  marital_mother                                              test_set sub1_t1 sub3_t1 sub4_t1 sub5_t1 sub7_t1
1        married                                    German speaking child      52       56       43       55       50
2        married                                    German speaking child      54       59       52       57       64
3        married  Turkish speaking child tested in German at first test date      46       55       47       40       50
4        married Turkish speaking child tested in Turkish at first test date      53       47       59       47       50
5       divorced                                    German speaking child      37       50       38       36       50
6        married                                    German speaking child      50       59       55       52       64
  sub1_t2 sub3_t2 sub4_t2 sub5_t2 sub7_t2
1      50      55      45      55      54
2      56      62      53      59      65
3      54      47      52      55      50
4      60      48      57      45      54
5      38      48      38      35      46
6      52      60      57      50      61
> tail(Example_1.1)
    no native_language age    sex age_birth no_siblings pos_sibling      social_status                    urban_rural
95  95          Turkish   9 female        39           1  first-born     middle classes City (over 20000 inhabitants)
96  96          Turkish   6   male        37           3  first born     middle classes City (over 20000 inhabitants)
97  97          Turkish   8 female        39           0  first-born lower middle class City (over 20000 inhabitants)
98  98          Turkish   7   male        38           4 fourth-born      lower classes City (over 20000 inhabitants)
99  99          Turkish   9 female        37           3  third-born  upper lower class City (over 20000 inhabitants)
100 100         Turkish   6   male        41           4  first-born     middle classes City (over 20000 inhabitants)
    marital_mother                                              test_set sub1_t1 sub3_t1 sub4_t1 sub5_t1 sub7_t1
95         married Turkish speaking child tested in Turkish at first test date      60       56       52       20       54
96         married Turkish speaking child tested in Turkish at first test date      61       52       47       35       35
97        divorced Turkish speaking child tested in Turkish at first test date      50       50       67       70       57
98        divorced Turkish speaking child tested in Turkish at first test date      27       50       35       42       43
99         married Turkish speaking child tested in Turkish at first test date      60       53       62       35       33
100        married Turkish speaking child tested in Turkish at first test date      50       37       67       52       25
    sub1_t2 sub3_t2 sub4_t2 sub5_t2 sub7_t2
95       58      50      52      33      54
96       54      52      50      33      50
97       60      59      64      67      65
98       30      47      42      40      39
99       60      48      64      36      31
100      46      36      64      59      31
>
```

Figure 1.1 Representation of the data structure of Example 1.1 in **R**.

In order to illustrate some statistical procedures we need other examples regarding content, but the data for these examples will not be found in Appendix A due to space limitations; however they are provided in the aforementioned `Workspace` and SPSS folders respectively. For the recalculation of the examples as well as for later calculations with the reader's own data, we will also provide the **R** instruction sequences, so that they don't have to be typed out. They can be found on the website www.Wiley.com. For beginners in **R** these are simply listed in order in a PDF file; for those readers already experienced in the use of **R** they are in a syntax editor for **R**; that is, `Tinn-R` (www.sciviews.org/Tinn-R/).



	no	native_language	age	sex	age_birth	no_siblings	pos_sibling	social_status	urban_rural	marital_mother
1	1	1	6	1	39	1	1	2	1	2
2	2	1	7	2	40	2	2	3	1	2
3	3	2	8	1	38	3	3	3	3	2
4	4	2	9	2	36	2	1	2	1	2
5	5	1	8	1	37	3	2	4	1	3
6	6	1	9	2	40	0	1	1	1	2
7	7	2	6	1	39	1	1	2	1	2
8	8	2	7	2	36	4	5	3	1	2
9	9	1	9	1	40	0	1	6	1	3
10	10	1	8	2	41	2	2	1	1	3
11	11	1	9	1	36	4	5	2	1	2
12	12	1	6	2	38	1	1	3	1	3
13	13	1	8	1	40	0	1	2	1	1
14	14	1	9	2	41	0	1	3	1	1
15	15	1	6	1	38	1	2	1	2	2
16	16	1	8	2	36	2	1	4	1	3
17	17	1	7	1	40	3	4	2	1	2
18	18	1	9	2	37	2	2	6	1	3
19	19	1	8	1	39	0	1	6	1	3
20	20	1	7	2	35	0	1	4	1	2
21	21	1	6	1	38	4	5	1	1	2
22	22	1	9	2	40	1	1	2	1	3
23	23	1	7	1	41	1	1	3	1	3
24	24	1	6	2	35	0	1	6	2	1
25	25	1	8	1	40	3	4	6	1	3

Figure 1.2 Part of the data view of Example 1.1 in SPSS.

References

Kleining, G. & Moore, H. (1968). Soziale Selbsteinstufung (SSE): Ein Instrument zur Messung sozialer Schichten [Social Self-esteem (SEE): An Instrumnet for Measuring the Social Status]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *20*, 502–552.

Kubinger, K. D. (2009a). *Adaptives Intelligenz Diagnostikum - Version 2.2 (AID 2) samt AID 2-Türkisch* [Adaptive Intelligence Diagnosticum, AID 2-Turkey Included]. Göttingen: Beltz.

Kubinger, K. D. (2009b). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens* (2nd edn) [Psychological Assessment – Theory and Practice of Psychological Consulting]. Göttingen: Hogrefe.

Rasch, D., Pilz, J., Verdooren, R. L., & Gebhardt, A. (2011). *Optimal Experimental Design with R*. Boca Raton: Chapman & Hall/CRC.