CONTENTS

Preface: How to Use This Book ix

PART 1 1
1. Types of Research 3
   1.1 Introduction 3
   1.2 Hypothesis generating vs. hypothesis testing 4
   1.3 Case studies vs. group studies 7
   1.4 Description vs. explanation 8
   1.5 Non-experimental vs. experimental 9
   1.6 Process research vs. product research 11
   1.7 Longitudinal vs. cross-sectional 12
   1.8 Qualitative vs. quantitative 14
   1.9 In-situ/naturalistic research vs. laboratory research 15
   1.10 The approaches taken in this book 18

2. Systematicity in Statistics: Variables 19
   2.1 Introduction 19
   2.2 Why do we need statistics? 20
   2.3 Variables and operationalization 21
   2.4 Statistical dangers and limitations 24

3. Descriptive Statistics 27
   3.1 Introduction 27
   3.2 Statistics: Means versus relationships 27
   3.3 Describing datasets: Means and dispersion 29
   3.4 A different view on variability 35
   3.5 Frequency distributions 37

4. Statistical Logic 43
   4.1 Introduction 43
   4.2 Independent versus dependent variables 43
   4.3 Statistical decisions 45
   4.4 The sample and the population 52
   4.5 Degrees of freedom 55
   4.6 Checking assumptions 55
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.7 Alpha, beta, and effect sizes</td>
<td>57</td>
</tr>
<tr>
<td>4.8 Significance revisited</td>
<td>59</td>
</tr>
<tr>
<td>4.9 Steps in statistical studies</td>
<td>60</td>
</tr>
<tr>
<td>5. Assessing Relationships and Comparing Groups</td>
<td>63</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>63</td>
</tr>
<tr>
<td>5.2 Assessing relationships</td>
<td>63</td>
</tr>
<tr>
<td>5.3 Comparing two groups</td>
<td>73</td>
</tr>
<tr>
<td>6. Simple and Multiple Linear Regression Analyses</td>
<td>80</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>80</td>
</tr>
<tr>
<td>6.2 Assessing relationships and simple regression</td>
<td>80</td>
</tr>
<tr>
<td>6.3 Multiple linear regression</td>
<td>88</td>
</tr>
<tr>
<td>6.4 Assumptions for regression and (non-parametric) alternatives</td>
<td>92</td>
</tr>
<tr>
<td>7. Additional Statistics for Group Comparisons</td>
<td>95</td>
</tr>
<tr>
<td>7.1 Introduction</td>
<td>95</td>
</tr>
<tr>
<td>7.2 Comparing more than two groups: The ANOVA</td>
<td>96</td>
</tr>
<tr>
<td>7.3 Factorial ANOVA</td>
<td>101</td>
</tr>
<tr>
<td>7.4 Alternatives for comparing means in case of assumption violations</td>
<td>107</td>
</tr>
<tr>
<td>7.5 Effect sizes</td>
<td>109</td>
</tr>
<tr>
<td>8. Summary and Concluding Remarks</td>
<td>111</td>
</tr>
<tr>
<td>8.1 Introduction</td>
<td>111</td>
</tr>
<tr>
<td>8.2 Overview and summary of tests</td>
<td>111</td>
</tr>
<tr>
<td>8.3 Additional statistics</td>
<td>118</td>
</tr>
<tr>
<td>8.4 Concluding remarks</td>
<td>122</td>
</tr>
<tr>
<td>PART 2-R</td>
<td></td>
</tr>
<tr>
<td>Practicals in R/RStudio</td>
<td>127</td>
</tr>
<tr>
<td>Getting ready to start using R and RStudio</td>
<td>129</td>
</tr>
<tr>
<td>Why use R?</td>
<td>129</td>
</tr>
<tr>
<td>Download and start using R and RStudio</td>
<td>130</td>
</tr>
<tr>
<td>R PRACTICAL 1: Exploring R and RStudio and Entering Variables (Chapter 2)</td>
<td>132</td>
</tr>
<tr>
<td>Part A: Simple calculations, vectors, and scripts</td>
<td>132</td>
</tr>
<tr>
<td>Part B: Installing packages – R Markdown</td>
<td>137</td>
</tr>
<tr>
<td>Part C: Importing and looking at data</td>
<td>140</td>
</tr>
</tbody>
</table>
# CONTENTS

<table>
<thead>
<tr>
<th>R PRACTICAL 2: Descriptive Statistics (Chapter 3)</th>
<th>146</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part A</td>
<td>146</td>
</tr>
<tr>
<td>Part B</td>
<td>150</td>
</tr>
<tr>
<td>Part C</td>
<td>151</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R PRACTICAL 3: Calculations using R (Chapter 4)</th>
<th>154</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part A</td>
<td>154</td>
</tr>
<tr>
<td>Part B</td>
<td>161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R PRACTICAL 4: Inductive Statistics (Chapter 5)</th>
<th>162</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part A: Reading and listening</td>
<td>162</td>
</tr>
<tr>
<td>Part B: Social class</td>
<td>165</td>
</tr>
<tr>
<td>Part C: Gender and intelligence</td>
<td>167</td>
</tr>
<tr>
<td>Part D: Alpha – This assignment is slightly more advanced</td>
<td>171</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R PRACTICAL 5: Regression/Miscellaneous Assignments (Chapter 5/6)</th>
<th>173</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part A: Nativeness ratings and self-perceived L2 speaking skills</td>
<td>173</td>
</tr>
<tr>
<td>Part B: Age and writing scores</td>
<td>174</td>
</tr>
<tr>
<td>Part C: Instruction, age, and writing scores – This part is slightly more advanced</td>
<td>177</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R PRACTICAL 6: More Advanced Group Comparisons (Chapter 7)</th>
<th>181</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part A: Exposure and vocabulary scores</td>
<td>181</td>
</tr>
<tr>
<td>Part B: Instruction and writing scores</td>
<td>182</td>
</tr>
<tr>
<td>Part C: Subtitles and vocabulary learning</td>
<td>186</td>
</tr>
</tbody>
</table>

| R PRACTICAL 7: Exam Practice | 191 |

<table>
<thead>
<tr>
<th>PART 2-JASP</th>
<th>195</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practicals in JASP</td>
<td>197</td>
</tr>
<tr>
<td>Getting ready to start using JASP</td>
<td>197</td>
</tr>
<tr>
<td>Why use JASP?</td>
<td>197</td>
</tr>
<tr>
<td>Download and start using JASP</td>
<td>197</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>JASP PRACTICAL 1: Exploring JASP and Entering Variables (Chapter 2)</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part A: Importing and looking at data</td>
<td>200</td>
</tr>
<tr>
<td>Part B: Preparing a data file</td>
<td>202</td>
</tr>
</tbody>
</table>
CONTENTS

JASP PRACTICAL 2: Descriptive Statistics (Chapter 3) 203
   Part A 203
   Part B 206
   Part C 206

JASP PRACTICAL 3: Calculations using JASP (Chapter 4) 208
   Part A 208
   Part B 214

JASP PRACTICAL 4: Inductive Statistics (Chapter 5) 215
   Part A: Reading and listening 215
   Part B: Social class 217
   Part C: Gender and intelligence 219
   Part D: Alpha – This assignment is slightly more advanced 223

JASP PRACTICAL 5: Regression/Miscellaneous Assignments (Chapter 5/6) 224
   Part A: Nativeness ratings and self-perceived L2 speaking skills 224
   Part B: Age and writing scores 225

JASP PRACTICAL 6: More Advanced Group Comparisons (Chapter 7) 228
   Part A: Exposure and vocabulary scores 228
   Part B: Instruction and writing scores 229
   Part C: Subtitles and Vocabulary Learning 232

JASP PRACTICAL 7: Exam Practice 236

References 239
Index 242
Packages used in 2-R 246
Functions used in 2-R 247
PART 1
1 TYPES OF RESEARCH

1.1 Introduction
The field of Applied Linguistics (AL) is a large one and this means that applied linguists are interested in many issues. Here is a random list showing a variety of topics:

• The effectiveness of early bilingual education: how effective is an early start?
• The relation between characteristics of linguistic input and language development in early bilingual children.
• Assessment of problems of elderly migrants learning Swedish.
• The lag in development in language proficiency of migrant children or children in deprived areas where a local dialect is spoken.
• The storage of multiple languages in our head: language selection and language separation (How do we keep our languages apart?).
• Can you ‘block’ a language while listening the way you do while speaking?
• The impact of learning a third language on skills in the first language (L1) and the second language (L2).
• The role of interaction in the language classroom: who is talking, what is the input?
• What is the impact of ICT on language learning?
• How can a threatened language be protected?
• Are Dutch people better at learning English than, for example, German learners?
• Why are prepositions in an L2 so difficult to learn?
• How can a forgotten language be reactivated?

This list could be extended for pages. A quick look at the Current Contents list of journals in Arts and Humanities, which shows the tables of contents of over 1000 journals, will make clear that the creativity has no limits and that even for journals focusing on second language development, the range of issues is breathtaking.
There are many topics of research, but the range of types of research in AL is much more limited. In this chapter we want to give a systematic overview of different types of research: what are relevant distinctions and how are different types related? There will be no single optimal method for all research topics, since each topic can be worked out in different ways. For your understanding of the research literature, it may be useful to become acquainted with the major categories, so you know what to expect and how to evaluate the use of a particular design. Also, some familiarity with research terminology and categorization will be helpful in finding relevant research on your own topic.

For clarity’s sake, we will make use of contrasting pairs of types of research, but it should be stressed from the outset that these contrasts are actually far ends on a continuum rather than distinct categories, and that the contrasts are all dimensions that may partly overlap.

1.2 Hypothesis generating vs. hypothesis testing

Testing and theory formation is a circular process. Theory must be based on empirical findings, and empirical studies can be used to test the theories formulated. One of the issues related to theory formation is to what extent a theory can be tested; that is, to what extent the theory can be used to set up hypotheses that can be tested empirically. Quite often, though, this is not a matter of either–or, but a sequential process. Often, theories have not been developed yet to the point that real, testable hypotheses have been generated. A current example could be the relation between the use of hand gestures and L2 development. We are only now beginning to see the importance of that type of research, and we are still looking for theories that may help us explain why people gesture the way they do (see for instance De Bot & Gullberg, 2010). So before we can test what explains cross-linguistic influence from L1 to L2 in gesturing, we first need to find out whether there are actually differences in gesturing between the two languages. Once we have established that, we can proceed to think about specific aspects,
such as the use of gestures with motion verbs. This means that for relatively
unexplored topics, we may first have to run some exploratory studies to
generate hypotheses that can be tested in subsequent research.

The next step in the research cycle is to test the hypotheses we have gen-
erated. In research reports, we often see phrases like ‘In this study, we test
the hypothesis that …’. However, the formulation of appropriate hypotheses
is not always obvious. For instance, if someone claims to ‘test the hypothesis
that after puberty a native level of proficiency can no longer be attained’,
then we may wonder what that actually means: is that true for every learner,
no matter what? If only one single individual can be found who can achieve
this, is the hypothesis then falsified? A hypothesis needs to be narrowed
down as far as possible to show what will be tested and what outcomes
count as support or counter evidence.

**ACTIVITY 1.2**

Formulating a research hypothesis

It is not easy to formulate a research hypothesis that is not too broad
and not too narrow. The more specific the hypothesis is, the better
the chance to test it successfully. The development of a research
hypothesis typically goes in stages. Consider a hypothesis like the
following:

‘Elderly people forget their second languages easily.’

The above sentence is not really a hypothesis, but rather a statement.
Which elderly people? Compared to what other groups? Do younger
people not forget their second languages? What does ‘easily’ mean
here? So we need to narrow that down:

‘Elderly people forget their second languages more quickly than
middle-aged people.’

Still, this is rather broad and some concepts are not clear, such as the
definition of ‘elderly’ and ‘middle-aged’. For the hypothesis this will
do, but in the description of the population the age range will have to
be made clear. Likewise, do you also want to include elderly people
suffering from dementia or other diseases? And do you want to test
every part of the language system? Maybe it is better to limit the
study to syntax, morphology, lexicon, or fluency. And do you want
to look at all second languages? How about the level of education,
which is likely to play a role? Narrowing the hypothesis down further
could result in something like:

‘Healthy elderly people forget words in their first second language more
quickly than education-matched middle-aged people.’
As we will see in Chapter 4, the careful formulation of hypotheses is crucial to our ability to find evidence for our observations. An important principle in finding evidence is the principle of falsification. Since it is logically impossible to prove that something is right, we will always try to prove that our hypothesis is wrong. In other words, the conventional common practice in research is to try and falsify hypotheses. When the formulation of a hypothesis is ambiguous, unclear, or too broad, it will not be possible to falsify a hypothesis.

The discussion of hypothesis testing inevitably leads to the issue of generalizability. Generalizability refers to the extent to which findings in a study can be generalized to a population that is larger than the samples tested. In most cases it is unimaginable that all individuals of a given group can be included in a study. No study of Chinese learners of English will include all those millions of people. What is typically done is that we draw a sample from that larger population. There are different methods to do so; we will get back to these later (in Section 4.9). The most desirable approach is to have a so-called ‘representative’ sample, which means that all the variation in the larger population is represented in the sample tested. This ideal is hardly ever fully achieved, because it is very difficult to assess what makes a sample representative; we need to know all the traits that may be relevant and should be included in the sampling. No data will tell us exactly what the relevant traits are to draw a sample from the large population of Chinese learners of English. The best we can do is to guess and use common sense (and all the relevant research there is, of course) to define the sample.

Statistics will help us in assessing how representative a sample is and how reliable it is to make generalizations of the findings of a sample to a larger population. Several factors will play a role in this, but one of the most important ones is the size of the sample. Here too, we can use common sense. Observations about a sample of ten participants can barely lead to valid conclusions about a population of a million people. But fortunately, we can also apply statistics to help us decide on the most appropriate sample size for a study, as we will see later on in this book.

Hypothesis testing and generalization of the findings to a larger population form the core of statistical studies. Many scientists believe that no serious research can be done outside this framework. However, not all studies use groups to understand or explain observations.

‘More quickly’ is still a bit underdefined, but probably clear enough. Sometimes it may help to break the larger hypothesis down into a number of smaller ones in which more details can be provided.

For three of the topics of the list in Activity 1.1, formulate a clear and specific research hypothesis.
1.3 Case studies vs. group studies

Our discussion about hypothesis testing seems to lead to the conclusion that research should always be done with groups of people that make up representative samples, and that the larger the sample, the better it is. However, the comparison of groups of people is certainly not the only way to do research. In group studies, the similarity within the group is crucial. A group is selected with specific characteristics, for example Turkish undergraduate students doing a course on Academic Writing in English, and other differences are either ignored or controlled for by using background questionnaires and specific statistical techniques to cancel out such differences. But in many cases the grouping factor can be problematic, as there are more differences than similarities between the individuals within the group. Another problem for group studies occurs when the groups represent rather random selections of a continuous phenomenon, like ‘Old’ versus ‘Young’. Somewhere we have to draw the line to create the categories of Old and Young, which means that we will create artificial categories that do not really exist. In these cases it is better not to use a groups-approach to the study. Still, the comparison of groups of people that share a particular characteristic is one of the most frequently occurring ways of doing research.

Another way of doing research is by closely looking at the behaviour of single individuals. These studies are commonly referred to as case studies. Case studies are often used for qualitative analyses, in which a detailed description can be derived from several data of a single person or from the interaction of two people. For instance, in conversation analysis or discourse analysis, a detailed analysis is done of the sequence and the quality of the interaction between two language users. In this way, we can come to an understanding of how we organize our conversations linguistically, but also how language learning takes place by putting the utterances ‘under the microscope’. Detailed observations on the behaviour (including language use) of single individuals can help us to understand and illustrate theoretical principles. Case studies provide an opportunity to observe human behaviour in a real life context. Observations in case studies cannot be generalized to a larger population of similar individuals, but case studies can be used for the generalization to the theory. Theoretical principles can be falsified when not observed in the cases studied.

Although case studies are usually associated with qualitative research, quantitative case studies can be done as well. For instance, in corpus linguistics, the language profile of individual cases can be created by calculations of frequencies and distributions. Case studies can also be used to investigate an individual’s language use as it develops over time. By closely keeping track of the changes in an individual’s language use, we can understand how stages of development emerge and how the combination of events leads to certain types of change. Case studies in the time domain are frequently used in studies based on Complex Dynamic Systems Theory. The dynamic
interaction of factors over time are best investigated by closely concentrating on single cases (see Section 1.6 on Process vs. product research below). Some studies use multiple cases, and although in terms of numbers they may seem to be the same as a group, the approach is fundamentally different. In case studies, we typically find a holistic approach, which aims at trying to integrate as many aspects that are relevant for the individual case as possible.

### 1.4 Description vs. explanation

The discussion about generating hypotheses vs. hypothesis testing in the section above is closely related to the distinction between description and explanation. Before we can explain anything, we first need a good description. For instance, before we can explain why interaction in the classroom is beneficial for second language development, we need to describe what goes on in classrooms: who is actually saying what to whom, how complex is the language used, is what is said also understood, and is the language used correct or full of errors? A large part of the research we do in AL is descriptive. We describe processes of learning and teaching, giving different factors that play a role in these processes; we describe language policy programmes and their effectiveness; we describe the impact of learning environments on learning, and so on.

Many studies investigate the effect of X on Y. A randomly chosen issue of one of the leading journals in our field, *Studies in Second Language Acquisition* (2018, 40), reveals the following: there is a study on the effect of the gloss type (like a textual definition or picture) on the learning of words while reading (Warren et al., 2018), one on the effect of different types of instruction on L2 processing of causative constructions (Wong & Ito, 2018), and the effect of exposure frequency on vocabulary learning (Mohamed, 2018). All were eye tracking studies. All of these studies describe the effect of the manipulation of one variable on another one. But even if the effect studied is found (and it normally is, because journals do not usually publish null results), that is not an explanation. The only conclusion that can be drawn is that the change in Y co-occurs with a change in X. To really account for the co-occurrence observed, there must always be a rigorous and detailed theory behind the research, which provides theoretical explanations for the phenomena found.

There is another sense of the concept of explanation, which is more statistical in nature. As we will see in Chapter 3 of this book, variation within individuals and groups vs. variation between individuals and groups is the essence of statistical procedures used to compare groups. We try to explain variation in variable A by looking at the impact of the systematic and controlled variation of variable B. For instance, we look at the variation in the acquisition of new L2 words (variable A) by manipulating the methods of teaching (variable B). If the experiment works, the variation in variable A is reduced, because we have taken out the variation that is caused by variable B. Suppose we teach two different methods to two different groups. If we look...
at the learning results of the group that used method 1 and the one that used method 2 together, we will find some learners who improved a lot, some less, and some not at all. If we look at the groups for methods 1 and 2 separately, we may find that for one method most learners have improved a lot, while for the other method learners improved only slightly or not at all. Even though there is a great deal of variation in the two groups taken together because there will be good and bad learners, there may be less variation within the groups than between the groups. In statistical terms, this is referred to as ‘variance explained’, variance being a specific type of variation. In this particular example, the variation between the groups can largely be explained by the different methods they were using. The goal in experiments is to explain as much variation as possible, as that will tell us to what extent we can explain a given effect. Again, this is not an explanation in the theoretical sense, but it is a description of the effect of one variable on another.

1.5 Non-experimental vs. experimental

Applied Linguistics has positioned itself as part of the social sciences and distanced itself from the humanities by adopting research techniques and paradigms based on the science model. In this model, quantitative empirical research and controlled experiments are often considered the only way to make progress. The aim is to decompose complex processes in parts that can be studied and manipulated experimentally. Experiments and statistical manipulations provide ‘hard’ evidence as compared to the ‘soft’ evidence evolving from the more interpretative research that dominates in the humanities. If one looks at the bulk of research as reported in the most prestigious journals and books in AL, the experimental approach is still dominant, but at the same time it is obvious that it is no longer seen as the only way to practise research. The choice of an experimental or a non-experimental approach largely depends on what a researcher wants to know. For a study of the organization of the bilingual lexicon or the perception of foreign accents by native speakers, experiments may be the logical choice. How human language processing takes place is not open to introspection and we can study this effectively through controlled experiments. Other aspects, such as non-instructed L2 development, we can study better through non-experimental techniques, such as observations and analyses of spontaneous speech.

In studying Second Language Development (SLD), a wide range of experimental techniques have been used, ranging from grammaticality judgments to lexical decision tasks. More recently, different neuroimaging techniques have been used that provide insight into brain activity while processing language. A detailed discussion of various techniques is beyond the scope of this book, but a good overview of different experimental techniques that have been used for the study of SLD can be found in Mackey and Gass (2005, 2016) and thorough discussions of different brain-imaging techniques have been provided by Abutalebi and Della Rosa (2008).
Sometimes there is a choice between using an experimental technique and a non-experimental technique. The choice for one or the other is determined by the aim of the study. An interesting example is the study of L2 pragmatic competence. Hendriks (2002) studied the acquisition of requests by Dutch learners of English. Her aim was to study the impact of power relations, social distance, and conversational setting on the use of politeness strategies in requests. She could have made recordings of requests in spontaneous conversations, but she wanted to study the systematic effects of each of these variables and their interactions. She would have needed a very large corpus of utterances to find sufficient examples of requests that differed in terms of power, social distance, and setting. Therefore, she decided to use an experimental technique that tries to mimic real life interaction while allowing for systematic variation of variables. The technique used was the Discourse Completion Task. In this task, a short description of a conversational setting is presented and the participant has to construct a sentence as a reaction. Here are two examples from Hendriks (2002, pp. 110–113):

**The living room**

You were in your room upstairs doing your maths homework, but you were not able to do the sums. You need some help. You go down to the living room where your dad is watching a documentary on television. What do you say to your dad?

.................................................................................................................................................................................................

**The supermarket**

You are standing in line at the checkout with a shopping trolley full of groceries. You are late for an important meeting. There is one man in front of you. What do you say to the man in front of you?

................................................................................................................................................................................................}

---

**ACTIVITY 1.3**

One of the subdivisions we can make with regard to types of research is the one between experimental and non-experimental.

- Which of the topics from the list in Activity 1.1 can best be investigated using experimental research and which with non-experimental research?
- Mention one advantage and one disadvantage of experimental research.
The use of such tasks allows for systematic variation of the variables, but it is of course not a natural setting or real conversation. Ideally, data from such controlled experiments should be validated through a comparison with real conversational data.

1.6 Process research vs. product research

A dimension that has been introduced relatively recently is the distinction between product-oriented research and process-oriented research. Most traditional research in the field of AL works with learning outcomes and background variables at one moment in time. This can be labelled as product research, as the conclusions relate to the outcomes of a process rather than the process itself. The process dimension, on the other hand, is founded in a research perspective that is rapidly gaining ground in developmental psychology and Applied Linguistics, and is based on Complex Dynamic Systems Theory (CDST) (see, for instance, De Bot et al., 2007). This perspective emphasizes the change of development over time and focuses on the dynamic interaction of factors affecting the language system, or the overall system in general, over time. CDST is thus interested in the process of development rather than in the eventual learning outcomes or products. The focus on the process has important consequences for the research choices that are made. For instance, a CDST approach takes into account that developmental processes are complex processes in which characteristics of the individual learner interact with the environment. CDST emphasizes the iterative nature of development, that is, each next step in development is based on the state of the complex system at the previous moment. And although one step ahead can be accurately predicted, each future step is increasingly difficult to predict, as all the factors that shape the system will also change over time. The implication of the iterations is that every learner’s developmental process is essentially unique. Inevitably, the complex and dynamic developmental process of an individual learner cannot be generalized to groups of individuals (populations), which means that process research is mostly done on cases. To make sense of the iterations, case studies in a CDST framework usually include a large number of observations over time, from about 30 observations in some studies to over 1000 in other studies. In this way, individual patterns of variability can tell us in detail how a language learner developed and what factors may have contributed to shaping each step in the developmental process.

Since language learning is essentially a process, it will be obvious that process-oriented case studies provide a valuable contribution to our understanding of learning. At the same time we need general tendencies, not only because of educational policy reasons, but also because we need information on the likelihood that a given factor will have an impact on the product of development, so that we can include it in the study of individual developmental patterns. This creates a research cycle in which we move from factors
that seem to have an effect at the individual level to testing that effect on a larger sample to get an estimate of its strength and then back to the individual level again to study the impact in more detail. An example could be the role of motivation: its effect may be suggested when we look at the learning process of an individual learner who indicates why she was motivated to invest time and energy in learning a language at some moments in time and not at all at other moments. To know more about the potential strength of such a factor, we may then do a study on a larger sample of similar learners that we compare at one point in time (a product study). With that information, we can go back to individuals and see how it affects them. The general pattern will be less detailed and typically will not give us information about change of motivation over time in the same way that an individual case study can.

**ACTIVITY 1.4**

In a CDST approach, we want to take into account that developmental processes are complex processes in which characteristics of the individual learner interact with the environment. Every learner is unique, and her developmental path will be affected by the internal structure of her system and her interaction with the environment. Inevitably there will be variation between individual learners.

- Order the items in the list in 1.1 on page 3 according to the degree to which the topic could be easily investigated from a CDST perspective.

Now that we have distinguished process-based research from product-based research, we can see that we need both of these dimensions to create a full picture. This means we should be balancing two ways of looking at research: on the one hand we need attention for individual developmental patterns revealing individually unique processes, but on the other hand we need to find general tendencies for various purposes. It may be true that for some learners there is no impact of the first language when learning the second, and the variation in cross-linguistic influences between individuals may be considerable, but what remains is that *in general*, the first language does play a role in learning a second language.

### 1.7 Longitudinal vs. cross-sectional

Longitudinal research is research in which individuals’ development over time is studied. Most studies of children growing up bilingually are examples of longitudinal research: the child is typically video/audio-recorded at regular intervals over a longer period of time, sometimes more than three
years, and transcripts of the recordings are analysed with respect to relevant aspects, such as mean length of utterance and lexical richness. But also other types of development can be studied longitudinally: Hansen et al. (2010) looked at the attrition of Korean and Japanese in returned missionaries who typically acquired the foreign language up to a very high level, used it in their work as missionaries, but hardly ever used it after they returned. This study is unusual, because it is longitudinal with only two moments of measurement in 10 years. In many longitudinal studies there are more moments of measurement with smaller time intervals. Longitudinal studies often take long; even three-year data collection periods may be too short to cover a significant part of the developmental process. And funding agencies are not very keen on financing projects that take more than four or five years to generate results. Therefore, the number of longitudinal studies is small, but those projects (like the European Science Foundation study on untutored L2 development in different countries (see Klein & Perdue, 1992; Becker & Carroll, 1997)) have had a major impact on the field.

Because of the time/money problem that characterizes longitudinal studies, many researchers use cross-sectional designs. In cross-sectional research, individuals in different phases of development are compared at one moment in time. For the study of the development of morphology in French as an L2, a researcher may compare first, third, and fifth graders in secondary schools in Denmark. Rather than follow the same group of learners for four years as they progress from first to fifth grade, different groups in the three grades are compared at one moment in time.

Both longitudinal designs and cross-sectional designs can have their problems. In longitudinal studies, the number of participants is generally very small because a large corpus of data is gathered on that one (or very small group of) individual(s). Large numbers of participants would make both the data collection procedure and the processing and analysis of the data extremely demanding and time-consuming. Small numbers, on the other hand, mean that the findings may be highly idiosyncratic and difficult to generalize. As we discussed in 1.6, this may not be a problem in studies that use the uniqueness of the individual’s development as the central issue, as is normally the case in CDST approaches to language development. Another problem of longitudinal studies is subject mortality, that is the dropping out of subjects in a study. With each new measurement, there is a risk of subjects dropping out, and the longer and more demanding the study, the higher the risk of drop-out. An additional problem is that in such studies drop-out is typically not random, but selective or biased: in a study on acquisition or attrition, subjects that do not perform well will be more likely to lose their motivation and drop out than more successful ones, leaving a biased sample that is even less generalizable.

Cross-sectional designs can also be problematic, because the assumption that the three groups that are compared behave like one group tested longitudinally may not be true. There may be specific characteristics of different age groups, such as changes in the school curriculum, natural disasters,
changes in demographic trends, and changes in school population, that can make the three groups very different. One solution for this so-called cohort effect is to take more than one cohort, so rather than only testing grades one, three, and five in year \( x \), also testing grades one, three, and five of the next year or cohort. If the findings for the two cohorts are similar, it is assumed that the groups do not behave atypically. Some studies try to get the best of two worlds by combining longitudinal and cross-sectional designs: in research on ageing such a cross-sequential design has been used frequently, and it was also used by Weltens (1989) and Grendel (1993) in their studies of the attrition of French as a foreign language in the Netherlands.

1.8 Qualitative vs. quantitative

The discussion of qualitative versus quantitative studies would take another book to discuss in sufficient detail. It has been one of the main rifts in the social sciences, including AL, in the past decades. It looks as if the fiercest controversy is over now, but the different communities still view each other with distrust. For a long time a researcher had to be in one or the other community, but now it seems acceptable to take an elective stance and use more qualitative or more quantitative methods depending on the type of research question one wants to answer. Following Mackey and Gass (2005, p. 400), the two approaches can be defined as follows:

**Qualitative:** Research in which the focus is on naturally occurring phenomena, and data are primarily recorded in non-numerical form.

**Quantitative:** Research in which variables are manipulated to test hypotheses, and in which there is usually quantification of data and numerical analyses.

From these definitions it follows that the two approaches differ fundamentally in epistemological terms and with respect to the research methods used. Qualitative research is holistic, trying to integrate as many aspects that are relevant into one study. It is also by definition interpretative and therefore in the eyes of its opponents ‘soft’. In qualitative research a number of techniques may be used, such as diaries of learners, interviews, observations, and introspective methods such as think-aloud protocols (see Mackey & Gass, 2005, Chapter 6, and Brown & Rodgers, 2002, Chapters 2–4 for discussions of various methods). One of the main problems is the lack of objectivity in those methods: in all these methods, the researchers interpret what is going on, and some form of credibility can only be achieved through combinations of data (triangulation) and the use of intersubjective methods, in which the interpretations of several ‘judges’ are tested for consistency. All of this may not satisfy the objections raised by hard-core quantitativists. For them, only objective quantitative data are real ‘hard’ data. Such data are claimed to be

objective; for example, there is little a researcher can change or interpret in the latencies in reaction-time experiments. The starting point of quantitative research is that the entire world is one big mechanism and by taking it apart and studying its constituent parts we will in the end understand the whole machine. Qualitative researchers criticize this approach as, as explained by Leo van Lier, ‘one cannot peel away the layers of an onion to eventually get to the real onion within’ (van Lier, 2004, p. 43). And that, according to qualitative researchers, is exactly what quantitative researchers tend to do. They try to explain complex phenomena by calculating the influence of a limited number of factors.

Another problem with the experimental and quantitative approach is that it is not always clear what participants in such experiments actually do. There is a substantial set of studies on the recognition of pseudo-homophones (like English ‘coin’, a piece of money, and French ‘coin’, which translates to corner in English). The list of words to be recognized typically consists of many regular words with some of these pseudo-homophones interspersed. The researcher’s hope is that the participants will not notice these words and become aware of the fact that they are special, because that could have an effect on their strategies in processing. To what extent participants actually do notice the trick is often unclear. Participants in such experiments are typically psychology students who have to take part in many different types of experiments and who have accordingly become quite clever in detecting the trick.

ACTIVITY 1.5

- Which of the topics in the list in Section 1.1 can best be investigated using longitudinal research and which can best be investigated using cross-sectional research?
- Mention one advantage and one disadvantage of cross-sectional research.
- How can the cohort effect be avoided?

1.9 In-situ/naturalistic research vs. laboratory research

In-situ or naturalistic research refers to research that studies a phenomenon in its normal, natural setting and in normal everyday tasks, while laboratory research refers to both isolating a phenomenon from its normal setting and to the use of data that are an artifact of the procedures used. Laboratory research aims at finding ‘pure’ effects that are not tainted by the messiness of everyday life. In such studies the grammaticality of sentences in isolation is tested through grammaticality judgments, or the process of lexical access is studied using reaction-time experiments. Experimental laboratory research
has reached extremely high standards, mainly through a very successful experimental psychology tradition in North America and Western Europe. Therefore, that type of research stands in high regard, also in AL. The counter movement that advocates a more qualitative and naturalistic approach has long been marginalized and has created its own subculture and its own journals and societies. Their main argument is that reductionist research has no ecological validity in that it does not really tell us what reality looks like. Researchers in the laboratory tradition have problems countering this argument, because their research does not always lead to the kind of deeper insight it is supposed to bring. The gap between the methods used and the reality it claims to inform us about has become so wide that even researchers themselves may have problems showing the relevance of what they do. To give an example, there is a large body of research on word recognition, mostly using the lexical decision paradigm in which participants are presented with letter strings on a computer screen and asked to indicate as quickly as possible whether the letter string is a word in a given language or not. At the beginning of this field of research, it was assumed that word recognition data would inform us about the process of normal reading, but over time, researchers working on word recognition have developed their own sets of questions that have basically no link to the process of normal reading.

Researchers using naturalistic data claim that their research is more ecologically valid because it focuses on the tasks in their normal setting. Up to a point this is probably true, but with the use of different introspective methods, they may also have crossed the line and adapted methods that may create their own type of data that are as far removed from reality as the reaction times and error rates of word recognition researchers. The validity of introspection has been questioned and the core of the problem is sublime expressed in the title of Klein’s (1989) review of Kasper and Faerch on introspective methods in L2 research: ‘Introspection into what?’.

ACTIVITY 1.6

Here is the summary of an article from the journal Applied Linguistics (Webb & Kagimoto, 2011). What do you think are the main characteristics in terms of research types of the research reported on?

Learning Collocations: Do the Number of Collocates, Position of the Node Word, and Synonymy Affect Learning?

Stuart Webb & Eve Kagimoto
Victoria University of Wellington

This study investigated the effects of three factors (the number of collocates per node word, the position of the node word, synonymy)
We have presented various types of research by giving a number of more or less opposite characteristics of research types. In Table 1.1, the list is repeated again, and a reflection on the characteristics on the left-hand side and on the right-hand side will make clear that the characteristics on either side are related and co-occur: quantitative research is often experimental, based on groups, and laboratory research, while qualitative research is often based on case studies in a naturalistic setting aimed at description more than explanation. This is not to say that there is no mixing possible; many researchers nowadays are eclectic and take what suits them without caring too much about what different communities of researchers may say about this. This typology may help the interpretation of research and to structure your own research.

Table 1.1 Types of research

<table>
<thead>
<tr>
<th>Hypothesis generating</th>
<th>Hypothesis testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Explanation</td>
</tr>
<tr>
<td>Non-experimental</td>
<td>Experimental</td>
</tr>
<tr>
<td>Process research</td>
<td>Product research</td>
</tr>
<tr>
<td>Longitudinal</td>
<td>Cross-sectional</td>
</tr>
<tr>
<td>Case studies</td>
<td>Group studies</td>
</tr>
<tr>
<td>Qualitative</td>
<td>Quantitative</td>
</tr>
<tr>
<td>In-situ research</td>
<td>Laboratory research</td>
</tr>
</tbody>
</table>
1.10 The approaches taken in this book

We have elaborated on the necessity of looking at individuals on the one hand, and generalizing over larger groups on the other hand. Only with these complementary research methods will we be able to advance in science and bring different theories closer together. To go into all possible ways of analysing data will be too much for an introductory book on methodology and statistics. In order to do group studies, you will need to know how to advance in structuring your research and how to analyse the data you get. We often encounter linguists who think of an interesting study to carry out, but who forget to think about how to analyse the data and then get stuck; likewise, some report on statistics but have forgotten an important step, which has rendered their results unreliable. The aim of this book is to make you aware of how to construct a quantitative study in an analysable way, and which tests you can carry out to check your hypotheses. Although we focus on group studies, we will keep stressing that there is also the possibility to look at individual data. If you want to learn more about how to analyse data from a longitudinal CDST perspective, you can consult Verspoor et al. (2011): *A dynamic approach to second language development: methods and techniques.*
INDEX

A
Alpha error (α) 47–49
Alternative hypothesis (H₀/H₁/H₂) see Hypothesis
Analysis of Covariance see ANCOVA
Analysis of Variance see ANOVA
ANCOVA 89, 120, 226
ANOVA 95–96, 179
  contrasts 121, 188
  mixed 108–109, 119
  multivariate 119
  one-way 96–98, 117, 184, 230–231
  repeated measures 109, 119
  types (I, II, III, IV) 188–189, 233–234
  Welch’s 108, 184, 231
Assumptions 55–56

B
Bar graph see Barplot
Barplot 71, 167, 219
Bayesian statistics 60, 197
Beta error (β) 47–50
Between-subjects 108–109
Boxplot 29, 32, 150, 187, 205–206, 233

C
Case studies 7–8, 11–12
Categorical variable see Variable
Causal modelling 66
Central tendency 30, 151, 204
CDST see Complex Dynamic Systems Theory
Chi-square (χ²) 69–73, 113, 115, 166–167, 218–219
CI see Confidence Interval
Cohen’s d 78, 170, 222
Cohen’s f 185
Cohort effect 14–15
Collinearity see Multicollinearity
Complex Dynamic Systems Theory 7, 11–13, 18, 35–36, 122
Confidence Interval (CI) 51, 59, 65–66, 99–100, 169, 216, 221
Console window 130–131
Constant see Intercept
Construct 21–24
Contingency table see Cross-tabulation
Continuity correction see Yates correction
Continuous variable see Variable
Contrasts see ANOVA, contrasts
Control variable see Variable
Correlation 63–67, 163–164, 216
Correlation coefficient 64
Covariates 120
Cramer’s V 73, 167, 219
Critical values 51–52, 66
Cronbach’s Alpha 67, 171–172, 223
Cross-sectional research see Research
Cross-tabulation 70, 165, 217–219

D
Degrees of freedom (df) 55
Dependent samples t-test see t-test
Dependent variable (DV) see Variable
Descriptive research see Research
Descriptive statistics 27, 147–148, 198, 203–204
Dispersion 31–33, 151
Dunn test 108

E
Ecological validity see Validity
Editor window 130–131
Effect size 57–59, 67–69, 78, 109–110
Equality of variances see Homogeneity of variance
Error (ε) 83
Error bars 105, 221
Eta-squared (η²) 109–110, 185, 231
Expected frequency (FE) 72, 166, 218
Experimental research see Research

F
Factor analysis 118–119
Factorial ANOVA see ANOVA
<table>
<thead>
<tr>
<th>Index Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors</td>
<td>24, 44</td>
</tr>
<tr>
<td>Falsification (principle of)</td>
<td>6, 48</td>
</tr>
<tr>
<td>Frequency</td>
<td>See Chi-square distribution 37–39</td>
</tr>
<tr>
<td></td>
<td>polygon 39–40</td>
</tr>
<tr>
<td></td>
<td>tally 37</td>
</tr>
<tr>
<td>Frequentist statistics</td>
<td>60, 197</td>
</tr>
<tr>
<td>F-value</td>
<td>88, 98</td>
</tr>
<tr>
<td>G</td>
<td>G*Power 225</td>
</tr>
<tr>
<td></td>
<td>Gaussian 39, 122</td>
</tr>
<tr>
<td>Generalization</td>
<td>6, 13, 36, 55, 124</td>
</tr>
<tr>
<td>Generalized additive mixed models (GAMMs)</td>
<td>121–122</td>
</tr>
<tr>
<td>Goodness of fit</td>
<td>87–88</td>
</tr>
<tr>
<td>Grouping variable</td>
<td>See Variable</td>
</tr>
<tr>
<td>H</td>
<td>Histogram 37–38, 153, 157–158, 207</td>
</tr>
<tr>
<td>Homogeneity of variance</td>
<td>56, 76–77</td>
</tr>
<tr>
<td>Homoscedasticity (of variance)</td>
<td>56, 92–94, 175–176, 226</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>4–6, 48–49, 59, 61</td>
</tr>
<tr>
<td></td>
<td>alternative (H₀/H₁/H₂) 48–49, 52, 61</td>
</tr>
<tr>
<td></td>
<td>null (H₀) 48–49, 52, 59, 61</td>
</tr>
<tr>
<td></td>
<td>testing 4, 6, 156</td>
</tr>
<tr>
<td>I</td>
<td>Independent samples t-test see t-test</td>
</tr>
<tr>
<td></td>
<td>Independent variable (IV) see Variable</td>
</tr>
<tr>
<td>Inductive statistics</td>
<td>27, 43, 162, 215</td>
</tr>
<tr>
<td>Inferential statistics</td>
<td>63</td>
</tr>
<tr>
<td>In-situ research</td>
<td>See Research</td>
</tr>
<tr>
<td>Integer variable</td>
<td>See Variable</td>
</tr>
<tr>
<td>Interval variable</td>
<td>See Variable</td>
</tr>
<tr>
<td>Intercept (b0)</td>
<td>82–83, 93</td>
</tr>
<tr>
<td>Interquartile range (IQR)</td>
<td>31–32, 45</td>
</tr>
<tr>
<td>Introspection</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Kruskal-Wallis H test 107–108, 184, 230</td>
</tr>
<tr>
<td></td>
<td>Kurtosis 41–42, 153, 156–158, 207, 210, 212–213</td>
</tr>
<tr>
<td>L</td>
<td>Laboratory research see Research</td>
</tr>
<tr>
<td>Least squares approach</td>
<td>86</td>
</tr>
<tr>
<td>Level of significance</td>
<td>See Significance Testing</td>
</tr>
<tr>
<td>Levels</td>
<td>23, 96</td>
</tr>
<tr>
<td>Levene’s test</td>
<td>76–77, 159, 211</td>
</tr>
<tr>
<td>Line graph</td>
<td>37–38, 221</td>
</tr>
<tr>
<td>Linearity</td>
<td>56, 92, 122</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>See Regression</td>
</tr>
<tr>
<td>Longitudinal research</td>
<td>See Research</td>
</tr>
<tr>
<td>Lower quartile</td>
<td>31–32</td>
</tr>
<tr>
<td>M</td>
<td>Main effect 102–104</td>
</tr>
<tr>
<td>Mann-Whitney U-test</td>
<td>77, 110, 159, 171, 212, 221</td>
</tr>
<tr>
<td>Mean</td>
<td>29–30</td>
</tr>
<tr>
<td>Meaningful(ness)</td>
<td>58, 62, 123</td>
</tr>
<tr>
<td>Means analyses</td>
<td>27, 95–96, 116</td>
</tr>
<tr>
<td>Median</td>
<td>30–32, 147, 204</td>
</tr>
<tr>
<td>Min-Max graph</td>
<td>36</td>
</tr>
<tr>
<td>Mode</td>
<td>30–31, 147–148, 204</td>
</tr>
<tr>
<td>Monotonic</td>
<td>57</td>
</tr>
<tr>
<td>Multicollinearity</td>
<td>92–93, 119, 179</td>
</tr>
<tr>
<td>Multiple regression</td>
<td>See Regression</td>
</tr>
<tr>
<td>N</td>
<td>Naturalistic research see Research</td>
</tr>
<tr>
<td>NHST</td>
<td>See Null Hypothesis Significance Testing</td>
</tr>
<tr>
<td>Non-Constant Error Variance test</td>
<td>176</td>
</tr>
<tr>
<td>Nominal variable</td>
<td>See Variable</td>
</tr>
<tr>
<td>Non-parametric statistics</td>
<td>43, 55–57</td>
</tr>
<tr>
<td>Null hypothesis (H₀)</td>
<td>See Hypothesis</td>
</tr>
<tr>
<td>Null Hypothesis Significance Testing (NHST)</td>
<td>59</td>
</tr>
<tr>
<td>Numeric variable</td>
<td>See Variable</td>
</tr>
<tr>
<td>O</td>
<td>Observed value (F₀) 72</td>
</tr>
<tr>
<td>Observer’s paradox</td>
<td>62</td>
</tr>
<tr>
<td>Omega-squared (ω²)</td>
<td>110, 185, 189, 231, 234</td>
</tr>
<tr>
<td>One sample t-test</td>
<td>See t-test</td>
</tr>
<tr>
<td>One-tailed testing</td>
<td>52</td>
</tr>
<tr>
<td>One-way ANOVA see ANOVA</td>
<td></td>
</tr>
<tr>
<td>Operationalization</td>
<td>21–25, 43, 61</td>
</tr>
<tr>
<td>Ordinal variable</td>
<td>See Variable</td>
</tr>
</tbody>
</table>

Copyrighted material – 9781352007817
INDEX

P
Package 129, 137
Paired samples t-test see t-test, dependent samples
Parametric statistics 43, 55–57
Pearson product-moment correlation see Pearson r
Pearson r 64–65, 111–112, 114, 163–164, 216
Percentile scores 39–40
Phi (φ) 73, 167, 219
Plots and files window 130–131
Population 6, 43, 52–55
Post-hoc 99–100, 184, 231
Power 53, 58, 123, 174, 185, 225
Predictor variable see Variable
Principle of falsification see Falsification (principle of)
Process research see Research
Product research see Research
p-value 50–51, 100, 176, 227

Q
Q-Q plot 176, 226–227, 230
Qualitative research see Research
Quantitative research see Research
Quasi-experimental research see Research

R
Range 31–32
Ratio variable see Variable
Regression
analysis 80
line 83, 85
logistic 94
mixed-effects 121
model 82–86
multiple 88–92, 115–116, 120–121, 178
simple 80–88, 115–116, 175, 226–227
Regression coefficients 83, 91, 116, 176–177, 180, 226–227
Regression towards the mean 25
Relationships 27–28, 63–67, 114
Reliability 25
Repeated measures ANOVA see ANOVA
Replication crisis 24–25, 60, 123
Representative sample see Sample
Research
cross-sectional 12–14
cross-sequential 14
descriptive 8
experimental 9–10, 43–45, 53
in-situ 15–16
laboratory 15–16
longitudinal 12–14, 18, 36
naturalistic 15–16
process 11–12
product 11–12
qualitative 14–15
quantitative 14–15
quasi-experimental 43, 53
Researcher expectancy 62
Residual sum of squares (RSS) 179
Residuals 84–85, 87, 92–94, 175–176, 226–227
r² 69, 78, 88, 109, 169–170, 221
R-squared
adjusted 92
multiple 88, 92

S
Sample 6–7, 43, 52–55, 61–62
Sample size 6, 53, 58
Scale 22–24
Scatterplot 20, 27–28, 64, 148–149, 205
Script 135–136
SD see Standard Deviation
Self-selection 61
Shapiro-Wilk test 56, 118, 157–158, 211–213
Significance 50–51, 59–60
Skewness 41–42, 153, 156–158, 207, 210, 212–213
Slope (b1) 82–83, 90
Split-half correlations 67
Standard deviation (SD) 32–35
Standard Error (SE) 54
Statistical model 30
Stratum 53
Student t-test see t-test, independent samples
Subject expectancy 62
Subject mortality 13
Subsets 152, 155, 157
Sum of squared differences 86
Sum of Squares of the Model (SSM) 109
Sum of Squares of the Total (SST) 109

T
Test-retest reliability 67
Triangulation 14
Tukey Honest Significant Differences (HSD) 99–100, 184, 231

\( t \)-test 73–75
  - dependent samples 77, 117, 160–161, 213–214
  - independent samples 73–75, 117, 156, 159–160, 210, 213
  - one sample 75, 161, 214
Welch's 76–77, 118, 160, 211
\( t \)-value 74, 78, 87

Two sample \( t \)-test see \( t \)-test, independent samples

Two-tailed testing 52

Two-way ANOVA see ANOVA, factorial
Type I error see Alpha error (\( \alpha \))
Type II error see Beta error (\( \beta \))

Upper quartile 31–32

Validity 21, 25
  - ecological 16
Variable 21–24
  - categorical 23–24
  - continuous 23, 89, 120
  - control 45
  - dependent (DV) 43–45, 61
    - functions 43–45
    - grouping 44
  - independent (IV) 43–45, 61
  - Integer 24, 143–144
  - interval 23–24
nominal 23–24
numeric 24, 134, 143–144
ordinal 23–24
predictor 44, 80–82, 120
ratio 23–24
scale 22–24
types 24
Variance (\( s^2 \)) 34, 57
Variance explained 9, 69
Vector 133–135

Welch-Satterthwaite correction 76
Welch's
  - adjustment 76–77
ANOVA see ANOVA
\( t \)-test see \( t \)-test

Wide format 160
Wilcoxon rank-sum test see Mann-Whitney U-test
Wilcoxon signed–rank test 77
Within–subjects 109, 119
Workspace window 130–131

XY-graph 64

Yates correction 166, 218

\( z \)-score 35, 155–156, 209