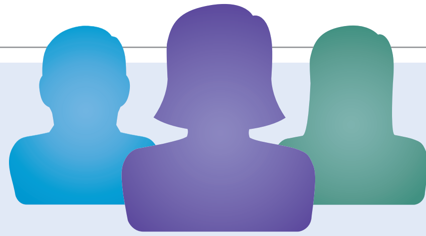
Eighth Edition

Understanding **Statistics in Psychology** with SPSS

Dennis Howitt
& Duncan Cramer

Pearson

Understanding Statistics in Psychology with SPSS



Understanding Statistics in Psychology with SPSS

Eighth edition

Dennis Howitt Loughborough University

Duncan Cramer Loughborough University

Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • São Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

PEARSON EDUCATION LIMITED
KAO Two
KAO Park
Harlow CM17 9SR
United Kingdom
Tel: +44 (0)1279 623623
Web: www.pearson.com/uk

First published 1997 (print)
Second edition published 2000 (print)
Revised second edition published 2003 (print)
Third edition published 2005 (print)
Fourth edition published 2008 (print)
Fifth edition published 2011 (print)
Sixth edition published 2014 (print and electronic)
Seventh edition published 2017 (print and electronic)
Eighth edition published 2020 (print and electronic)

ISBN: 978-1-292-28230-5 (print)
978-1-292-28233-6 (PDF)
978-1-292-28232-9 (ePub)

British Library Cataloguing-in-Publication Data
A catalogue record for the print edition is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Howitt, Dennis, author. | Cramer, Duncan, 1948- author.
Title: Understanding statistics in psychology with SPSS / Dennis Howitt, Loughborough University, Duncan Cramer, Loughborough University.
Description: Eighth edition. | Harlow, England ; New York : Pearson, 2020. | "First published 1997"—Title page verso. | Includes bibliographical references and index.
Identifiers: LCCN 2019059735 (print) | LCCN 2019059736 (ebook) | ISBN 9781292282305 (paperback) | ISBN 9781292282336 (ebook) | ISBN 9781292282329 (epub)
Subjects: LCSH: Psychometrics.
Classification: LCC BF39 .H74 2020 (print) | LCC BF39 (ebook) | DDC 150.1/5195—dc23
LC record available at https://lccn.loc.gov/2019059735
LC ebook record available at https://lccn.loc.gov/2019059736

A catalog record for the print edition is available from the Library of Congress

10 9 8 7 6 5 4 3 2 1
24 23 22 21 20

Front cover image: Scott Partridge/Moment Open/Getty Images

Print edition typeset in 9.5/12 pt Sabon LT Pro by SPi Global
Printed in Slovakia by Neografia

NOTE THAT ANY PAGE CROSS REFERENCES REFER TO THE PRINT EDITION

Brief contents

Contents

Companion Website

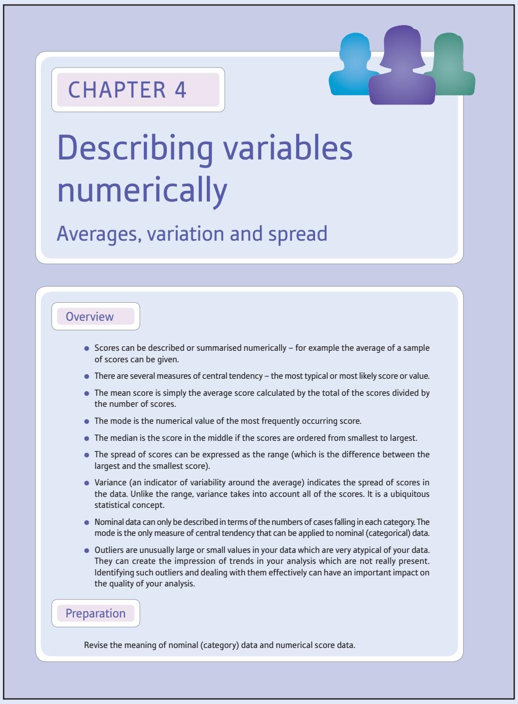
ON THE WEBSITE

For open-access **student resources** specifically written to complement this textbook and support your learning, please visit **go.pearson.com/uk/he/resources**

Lecturer Resources

For password-protected online resources tailored to support the use of this textbook in teaching, please visit **go.pearson.com/uk/he/resources**

Guided tour



CHAPTER 4

Describing variables numerically

Averages, variation and spread

Overview

- Scores can be described or summarised numerically – for example the average of a sample of scores can be given.
- There are several measures of central tendency – the most typical or most likely score or value.
- The mean score is simply the average score calculated by the total of the scores divided by the number of scores.
- The mode is the numerical value of the most frequently occurring score.
- The median is the score in the middle if the scores are ordered from smallest to largest.
- The spread of scores can be expressed as the range (which is the difference between the largest and the smallest score).
- Variance (an indicator of variability around the average) indicates the spread of scores in the data. Unlike the range, variance takes into account all of the scores. It is a ubiquitous statistical concept.
- Nominal data can only be described in terms of the numbers of cases falling in each category. The mode is the only measure of central tendency that can be applied to nominal (categorical) data.
- Outliers are unusually large or small values in your data which are very atypical of your data. They can create the impression of trends in your analysis which are not really present. Identifying such outliers and dealing with them effectively can have an important impact on the quality of your analysis.

Preparation

Revise the meaning of nominal (category) data and numerical score data.

110 CHAPTER 8 CORRELATION COEFFICIENTS: PEARSON'S CORRELATION AND SPEARMAN'S RHO

introduced the concept in Chapter 1 (Section 1.5) and devote a chapter (Chapter 11) to explaining it. Statistical significance indicates that your correlation, etc. is unlikely to be a fortuitous or fluke finding. That is, the correlation is large enough that it is unlikely to come from a population in which there is really a zero correlation. So you assume that your finding reflects a relationship that truly exists. If it is unlikely that your correlation is a fluke then the correlation is said to be statistically significant. This probability is usually set at .05 (i.e. 5%) or lower. However, the important point for now is to remember that statistical significance is invariably given with the value of the correlation coefficient.

We would write something like: 'It was found that musical ability was inversely related to mathematical ability. The Pearson correlation coefficient was −.90 which is statistically significant at the 5% level with a sample size of 10.' The statement about statistical significance will become clearer after you have studied Chapter 11.

If we follow the advice of the 2010 Publication Manual of the American Psychological Association (APA) we could write: 'Musical ability was significantly inversely related to mathematical ability, $r(8) = -.90$, $p < .05$.' The number in brackets after r is the sample size minus 2. This number is called the degrees of freedom and is explained in detail in Section 23.4. It is more usual in a statistical analysis to report degrees of freedom rather than sample size. Statistical significance is usually reported as a proportion rather than a percentage. Computer packages like SPSS give the exact significance level. The APA Publication Manual recommends that researchers give this exact significance rather than simply to indicate significance at the 5% or .05 level.

Box 8.1 Key concepts

Covariance

Many of the basic concepts taught in introductory statistics are relevant even at the advanced level. The concept of covariance is one of these. As we have seen, covariance is based on the deviation from the mean for the variable X multiplied by the deviation of the variable Y for each pair of scores. In other words, it is the top part of the Pearson correlation formula. The correlation coefficient is simply the ratio of the covariance over the largest value that the covariance could take for a particular pair of variables. That makes the correlation coefficient a standardised measure of covariance. But the term covariance crops up throughout this book in a number of different contexts. It is involved in ANOVA (especially the analysis of covariance) and regression, for example – lots of places, some of them unexpected.

One phrase that might cause some consternation when you first come across it is that of the 'variance–covariance' matrix. This is simply a table (matrix) which includes the variances of each variable in the diagonal and their covariances off the diagonal. This is illustrated for variables X, Y and Z in Table 8.3. The diagonal contains the variances but the other numbers are the covariances – each of these is presented twice because the covariance of X with Z is the same as the covariance of Z with X.

Similar matrices are produced for correlation coefficients. However, in this case the diagonal consists of 1.00s (the correlation of a variable with itself is always 1) and the off-diagonals have the correlation coefficients of each variable with the other variables.

Table 8.3 Variance–covariance matrix for three variables

	Variable *X*	Variable *Y*	Variable *Z*
Variable *X*	2.400	1.533	1.244
Variable *Y*	1.533	4.933	3.733
Variable *Z*	1.244	3.733	5.156

Clear overview

Introduce the chapter to give students a feel for the topics covered.

Key concepts

Offer guidance on the important concepts and issues discussed in the text.

Box 11.1 Focus on

Do correlations differ?

Notice that throughout this chapter we are comparing a particular correlation coefficient obtained from our data with the correlation coefficient that we would expect to obtain if there were no relationship between the two variables at all. In other words, we are calculating the likelihood of obtaining the correlation coefficient based on our sample of data if, in fact, the correlation between these two variables in the population from which the sample was taken is actually .0. However, there are circumstances in which the researcher might wish to assess whether two correlations obtained in their research are significantly different from each other. Imagine, for example, that the researcher is investigating the relationship between satisfaction with one's marriage and the length of time that individuals have been married. The researcher notes that the correlation between satisfaction and length of marriage is .25 for male participants but .53 for female participants. There is clearly a difference here, but is it a statistically significant one? So essentially the researcher needs to know whether a correlation of .53 is significantly different from a correlation of .25 (the researcher has probably already tested the significance of each of these correlations separately but, of course, this does not answer the question of whether the two correlation coefficients differ from each other). It is a relatively simple matter to do this calculation. It has to be done by hand, unfortunately. The procedure for doing this is described in Section 37.7 Comparing a study with a previous study.

11.4 Pearson's correlation coefficient again

Computer programs such as SPSS give exact significance levels for your correlation coefficient. Nevertheless, originally one would have used tables of the distribution of the correlation coefficient to find the significance level. Occasionally you still might need to consult such a table:

- For example, imagine that you are reviewing the research literature and find that one old study reports a correlation of .66 between two variables but fails to give the significance level, then what do you do? This sort of situation can occasionally happen since not every research paper is exemplary in its statistical analysis. Or you wish to check that there is not a typographical error for the given significance level then what do you do? SPSS will not be of help in these situations.
- What if you wanted to know the size of correlation which would be statistically significant for a given sample size? If, for example, you are expecting a small correlation of say .2 then how big a sample would be needed for this to be statistically significant? The only way to find out is to consult tables.

Since SPSS does not help here, in this section we will explain how significance levels may be obtained from tables. But you need to the size of the correlation coefficient and the sample size (or in some tables the degrees of freedom) involved.

The null hypothesis for research involving the correlation coefficient is that there is *no* relationship between the two variables. In other words, the null hypothesis states that the correlation coefficient between the two variables is .00 in the population (defined by the null hypothesis). So what if, in a sample of 10 pairs of scores, the correlation is .94 as for the data in Table 11.3? Do we accept or reject the null hypothesis?

Focus on

Explore particular concepts in more detail.

Explaining statistics 12.1

How the estimated standard error works

Table 12.3 is a sample of six scores taken at random from the population: 5, 7, 3, 6, 4, 5.

Table 12.3 Steps in calculating the standard error

X (scores)	X^2 (squared scores)
5	25
7	49
3	9
6	36
4	16
5	25
$\sum X = 30$	$\sum X^2 = 160$

Step 1 Using this information we can estimate the standard error of samples of size 6 taken from the same population. Taking our six scores (X), we need to produce Table 12.3, where $N = 6$.

Step 2 Substitute these values in the standard error formula:

$$\text{(estimated) standard error} = \frac{\sqrt{\dfrac{\sum X^2 - \dfrac{(\sum X)^2}{N}}{N-1}}}{\sqrt{N}} = \frac{\sqrt{\dfrac{160 - \dfrac{30^2}{6}}{6-1}}}{\sqrt{6}} = \frac{\sqrt{\dfrac{160 - \dfrac{900}{6}}{5}}}{2.449}$$

$$= \frac{\sqrt{\dfrac{160 - 150}{5}}}{2.449} = \frac{\sqrt{\dfrac{10}{5}}}{2.449}$$

$$= \frac{\sqrt{2}}{2.449} = \frac{1.414}{2.449} = 0.58$$

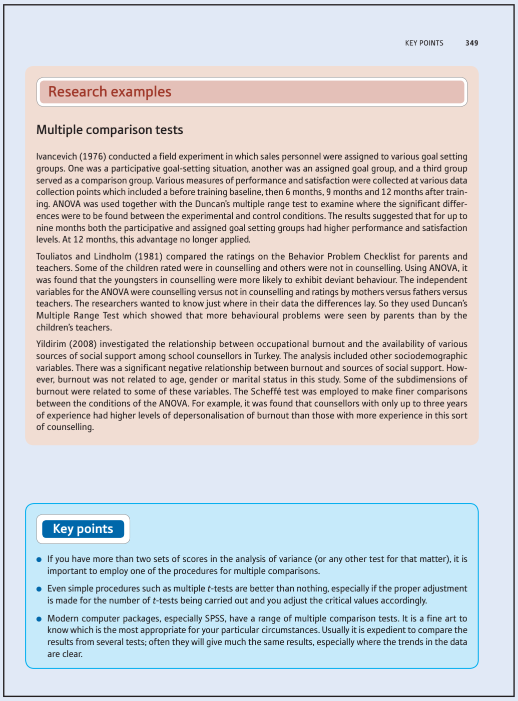
Note that this is the same value as that given by SPSS in Screenshot 12.5.

Interpreting the results

The standard error is 0.58. This is a measure of deviation of sample means from the population mean. It is a difficult concept to make concrete. Very roughly speaking, we could say that the standard deviation is the typical amount by which sample means deviate from the population mean. Some statisticians (e.g. Huck, 2009) dislike this sort of explanation though they offer no easily understood alternative for non-statisticians. It is possible to use a special mathematical distribution, the *t*-distribution, to indicate the proportions of sample means which lie between the population mean and any number of standard errors away from it. This is discussed in the following two chapters.

Explaining statistics

Take students through a statistical test with a detailed step-by-step explanation.



KEY POINTS 349

Research examples

Multiple comparison tests

Ivancevich (1976) conducted a field experiment in which sales personnel were assigned to various goal setting groups. One was a participative goal-setting situation, another was an assigned goal group, and a third group served as a comparison group. Various measures of performance and satisfaction were collected at various data collection points which included a before training baseline, then 6 months, 9 months and 12 months after training. ANOVA was used together with the Duncan's multiple range test to examine where the significant differences were to be found between the experimental and control conditions. The results suggested that for up to nine months both the participative and assigned goal setting groups had higher performance and satisfaction levels. At 12 months, this advantage no longer applied.

Touliatos and Lindholm (1981) compared the ratings on the Behavior Problem Checklist for parents and teachers. Some of the children rated were in counselling and others were not in counselling. Using ANOVA, it was found that the youngsters in counselling were more likely to exhibit deviant behaviour. The independent variables for the ANOVA were counselling versus not in counselling and ratings by mothers versus fathers versus teachers. The researchers wanted to know just where in their data the differences lay. So they used Duncan's Multiple Range Test which showed that more behavioural problems were seen by parents than by the children's teachers.

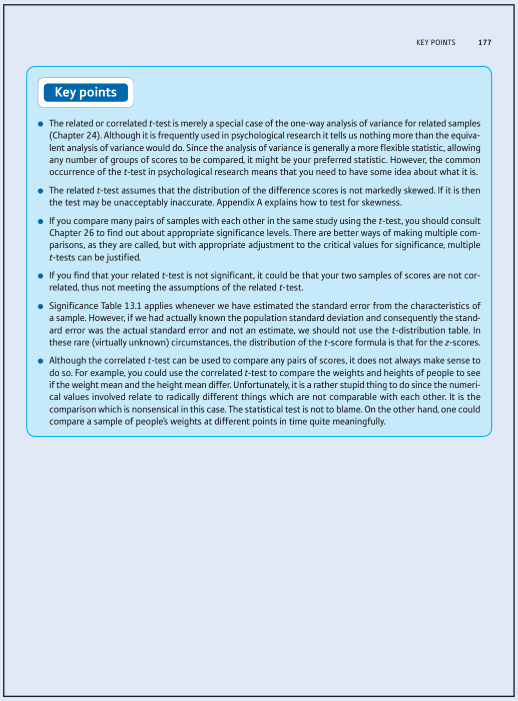
Yildirim (2008) investigated the relationship between occupational burnout and the availability of various sources of social support among school counsellors in Turkey. The analysis included other sociodemographic variables. There was a significant negative relationship between burnout and sources of social support. However, burnout was not related to age, gender or marital status in this study. Some of the subdimensions of burnout were related to some of these variables. The Scheffé test was employed to make finer comparisons between the conditions of the ANOVA. For example, it was found that counsellors with only up to three years of experience had higher levels of depersonalisation of burnout than those with more experience in this sort of counselling.

Key points

- If you have more than two sets of scores in the analysis of variance (or any other test for that matter), it is important to employ one of the procedures for multiple comparisons.
- Even simple procedures such as multiple *t*-tests are better than nothing, especially if the proper adjustment is made for the number of *t*-tests being carried out and you adjust the critical values accordingly.
- Modern computer packages, especially SPSS, have a range of multiple comparison tests. It is a fine art to know which is the most appropriate for your particular circumstances. Usually it is expedient to compare the results from several tests; often they will give much the same results, especially where the trends in the data are clear.

Research examples

Demonstrate how the statistical tests have been used in real research.



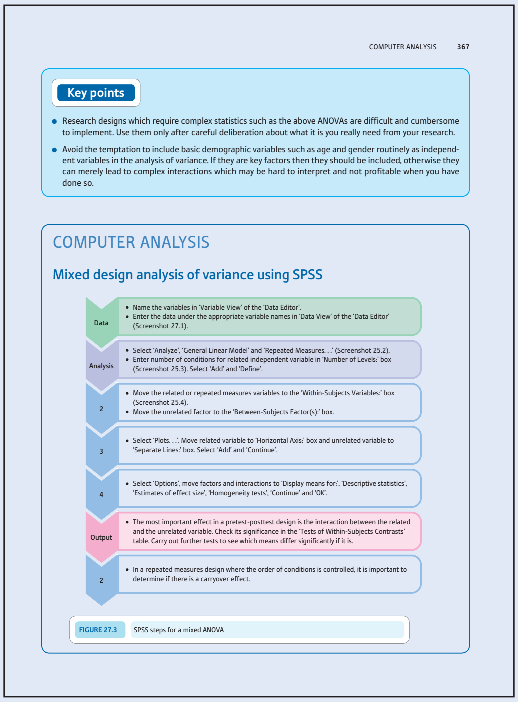
KEY POINTS 177

Key points

- The related or correlated *t*-test is merely a special case of the one-way analysis of variance for related samples (Chapter 24). Although it is frequently used in psychological research it tells us nothing more than the equivalent analysis of variance would do. Since the analysis of variance is generally a more flexible statistic, allowing any number of groups of scores to be compared, it might be your preferred statistic. However, the common occurrence of the *t*-test in psychological research means that you need to have some idea about what it is.
- The related *t*-test assumes that the distribution of the difference scores is not markedly skewed. If it is then the test may be unacceptably inaccurate. Appendix A explains how to test for skewness.
- If you compare many pairs of samples with each other in the same study using the *t*-test, you should consult Chapter 26 to find out about appropriate significance levels. There are better ways of making multiple comparisons, as they are called, but with appropriate adjustment to the critical values for significance, multiple *t*-tests can be justified.
- If you find that your related *t*-test is not significant, it could be that your two samples of scores are not correlated, thus not meeting the assumptions of the related *t*-test.
- Significance Table 13.1 applies whenever we have estimated the standard error from the characteristics of a sample. However, if we had actually known the population standard deviation and consequently the standard error was the actual standard error and not an estimate, we should not use the *t*-distribution table. In these rare (virtually unknown) circumstances, the distribution of the *t*-score formula is that for the *z*-scores.
- Although the correlated *t*-test can be used to compare any pairs of scores, it does not always make sense to do so. For example, you could use the correlated *t*-test to compare the weights and heights of people to see if the weight mean and the height mean differ. Unfortunately, it is a rather stupid thing to do since the numerical values involved relate to radically different things which are not comparable with each other. It is the comparison which is nonsensical in this case. The statistical test is not to blame. On the other hand, one could compare a sample of people's weights at different points in time quite meaningfully.

Key points

Each chapter concludes with a set of the key points to provide a useful reminder when revising a topic.



COMPUTER ANALYSIS 367

Key points

- Research designs which require complex statistics such as the above ANOVAs are difficult and cumbersome to implement. Use them only after careful deliberation about what it is you really need from your research.
- Avoid the temptation to include basic demographic variables such as age and gender routinely as independent variables in the analysis of variance. If they are key factors then they should be included, otherwise they can merely lead to complex interactions which may be hard to interpret and not profitable when you have done so.

COMPUTER ANALYSIS

Mixed design analysis of variance using SPSS

Data
- Name the variables in 'Variable View' of the 'Data Editor'.
- Enter the data under the appropriate variable names in 'Data View' of the 'Data Editor' (Screenshot 27.1).

Analysis
- Select 'Analyze', 'General Linear Model' and 'Repeated Measures. . .' (Screenshot 25.2).
- Enter number of conditions for related independent variable in 'Number of Levels:' box (Screenshot 25.3). Select 'Add' and 'Define'.

2
- Move the related or repeated measures variables to the 'Within-Subjects Variables:' box (Screenshot 25.4).
- Move the unrelated factor to the 'Between-Subjects Factor(s):' box.

3
- Select 'Plots. . .'. Move related variable to 'Horizontal Axis:' box and unrelated variable to 'Separate Lines:' box. Select 'Add' and 'Continue'.

4
- Select 'Options', move factors and interactions to 'Display means for:', 'Descriptive statistics', 'Estimates of effect size', 'Homogeneity tests', 'Continue' and 'OK'.

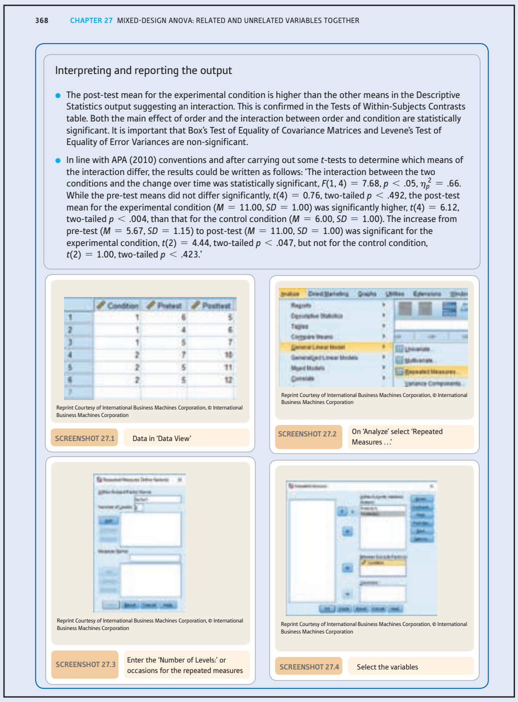
Output
- The most important effect in a pretest-posttest design is the interaction between the related and the unrelated variable. Check its significance in the 'Tests of Within-Subjects Contrasts' table. Carry out further tests to see which means differ significantly if it is.

2
- In a repeated measures design where the order of conditions is controlled, it is important to determine if there is a carryover effect.

FIGURE 27.3 SPSS steps for a mixed ANOVA

Computer analysis

Step-by-step advice and instruction on analysing data using SPSS Statistics are provided at the end of each chapter.



368 CHAPTER 27 MIXED-DESIGN ANOVA: RELATED AND UNRELATED VARIABLES TOGETHER

Interpreting and reporting the output

- The post-test mean for the experimental condition is higher than the other means in the Descriptive Statistics output suggesting an interaction. This is confirmed in the Tests of Within-Subjects Contrasts table. Both the main effect of order and the interaction between order and condition are statistically significant. It is important that Box's Test of Equality of Covariance Matrices and Levene's Test of Equality of Error Variances are non-significant.
- In line with APA (2010) conventions and after carrying out some t-tests to determine which means of the interaction differ, the results could be written as follows: 'The interaction between the two conditions and the change over time was statistically significant, $F(1, 4) = 7.68$, $p < .05$, $\eta_p^2 = .66$. While the pre-test means did not differ significantly, $t(4) = 0.76$, two-tailed $p < .492$, the post-test mean for the experimental condition ($M = 11.00$, $SD = 1.00$) was significantly higher, $t(4) = 6.12$, two-tailed $p < .004$, than that for the control condition ($M = 6.00$, $SD = 1.00$). The increase from pre-test ($M = 5.67$, $SD = 1.15$) to post-test ($M = 11.00$, $SD = 1.00$) was significant for the experimental condition, $t(2) = 4.44$, two-tailed $p < .047$, but not for the control condition, $t(2) = 1.00$, two-tailed $p < .423$.'

Reprint Courtesy of International Business Machines Corporation, © International Business Machines Corporation

SCREENSHOT 27.1 Data in 'Data View'

Reprint Courtesy of International Business Machines Corporation, © International Business Machines Corporation

SCREENSHOT 27.2 On 'Analyze' select 'Repeated Measures ...'

Reprint Courtesy of International Business Machines Corporation, © International Business Machines Corporation

SCREENSHOT 27.3 Enter the 'Number of Levels:' or occasions for the repeated measures

Reprint Courtesy of International Business Machines Corporation, © International Business Machines Corporation

SCREENSHOT 27.4 Select the variables

SPSS screenshots

The guidance on how to use SPSS for each statistical test is accompanied by screenshots, so the processes can be easily followed.

Introduction

This is the eighth edition of *Understanding Statistics in Psychology with SPSS*. Hopefully, this is even more helpful to the student learning experience. One thing that has not changed which sets this book apart from others aimed at students: it continues to provide an accessible introduction to the wide range of statistics that are employed by professional researchers. Students using earlier editions of the book will by now often be well into teaching and research careers of their own. We hope that this edition will be permanently on their desks while they instruct their students how to do statistics properly. In the distant past, the abbreviation SPSS stood for Statistical Software for the Social Sciences. Although the official name of the latest release at the time of publication is IBM SPSS Statistics 25.0 we shall refer to it throughout this book as SPSS because it is shorter, most users refer to it this way and the first letter of the original acronym actually refers to Statistical and so to add Statistics again seems repetitive. For most users of SPSS, SPSS versions have changed little since SPSS 13 came out in 2005, so this book will also be suitable for those using these earlier releases. Real changes in SPSS only slowly emerge.

We have considered very carefully the need for instruction into how to compute statistics using SPSS and other computer programs. Our approach in this book is to provide the basic steps needed for the computation but we have added a number of screenshots to help the reader with the analysis. Students of today are very familiar with computers and many do not need overly detailed instructions. Too much detailed step-by-step instruction tends to inhibit exploration of the program – trying things out simply to see what happens and using one's intelligence and a bit of knowledge to work out what things mean. Students can become fixated on the individual steps and fail to get a complete picture of the process of doing statistics with SPSS or other computer programs. In the end, learning to use a computer program is quicker if the user takes some responsibility for their learning.

Much of our daily use of computers in general is on a trial and error basis (we don't need step-by-step instructions to use Facebook or eBay) so why should this be different for statistics programs? How many of us read instructions for the iPhone in detail before trying things out? Of course, there is nothing unusual about tying statistics textbooks to computer packages such as SPSS. Unfortunately, SPSS is not the complete answer to the statistical needs of psychologists. It simply does not do everything that students (and professionals for that matter) need to know. Some of these things are very simple and easily computed by hand if instructions are provided. Other things do require computer programs other than SPSS when procedures are not available on SPSS. We think that ideally psychologists should know the statistics which their discipline needs and not simply those that SPSS provides.

SPSS is very good at what it does but there are times when additional help is needed. This is why we introduce students to other programs which will be helpful to them when necessary. One of the most important features of SPSS is that it is virtually universally available to students for little or no cost thanks to site licensing agreements. Unfortunately, this is not true of all other commercial statistics software. For that reason we have suggested and recommended programs which are essentially free for the user. The Web has a surprisingly large amount of such software to carry out a wide range of statistical routines. A few minutes using Google or some other search engine will often be bountifully productive. Some of these programs are there to be downloaded but others, applets, are instantly available for calculations. We have added, at the end of each chapter, advice on the use of software.

This does not mean that we have abandoned responsibility for teaching how statistics works in favour of explaining how to press keys on a computer keyboard. Although we think it best that statistics are computed using statistics programs because the risk of simple calculation errors is reduced, it seems to us that knowing how to go about doing the calculations that computer programs will do for you leads to an understanding of statistics which relying on computers alone does not. So we have included sections entitled 'Explaining statistics' which are based on hand calculation methods which should help students understand better what the computer program does (more or less) when it is used to do that calculation. Statistical techniques, after all, are little more than the mathematical steps involved in their calculation. Of course, they may be ignored where this level of knowledge is not required.

The basic concept of the book remains the same – a modular statistics package that is accessible throughout to a wide ability range of students. We have attempted to achieve this while being as rigorous as possible where rigour is crucial. Ultimately this is a book for students, though its emphasis on statistics in practice means that it should be valuable to anyone seeking to familiarise themselves with the vast majority of common statistical techniques employed in modern psychology and related disciplines. Not all chapters will be useful to everyone but the book, taken as a whole, provides a sound basis for learning the statistics which professional psychologists use. In this sense, it eases the transition from being a student to being a professional.

For this edition, we have added a chapter on 'big data' and its analysis. Not only does this contrast with more traditional approaches to psychological research but it offers radically new data sets of importance to psychologists. We have also revised every chapter to improve readability and ease the study load wherever possible. Changes have been made to encourage the student to reflect on the special demands of studying the somewhat alien topic of statistics.

Dennis Howitt and Duncan Cramer

Acknowledgements

Working on a book with the team at Pearson is one of the most pleasurable experiences imaginable even when things do not run smoothly. We are in the debt to many people for their talents which contributed to the production of this edition. A few people warrant special mention.

For more years than usual in publishing we worked with Janey Webb who was exceptional in so many ways and is greatly missed now that she has departed from Pearson. Her support was constant and generous. Her official title changed over the years but ended as Publisher though we prefer to think of her just as Janey. We welcome Catherine Yates as the new Publisher. Goodbye also to Saraswati Banerjee who was Acquisitions Editor. She was always superb support and good wishes to her on her new path.

Akanksha Marwah was the Assistant Content Producer who took our manuscript pressed the metaphorical buttons needed to make ready for production. With a statistics book this is a task not to be envied, we guess, but somehow it worked out just fine

The Copy Editor was Marie Gill. She did excellent work imposing the text design on our manuscript. Remarkably she also found time to spot our mistakes and references we forgot. It was great to be working with her.

Kevin Ancient did the text design and the cover design. It is difficult to write anything new about Kevin's work as we have thanked him in numerous books and editions. His contribution has always been brilliant.

Proof reading is a magical skill to us as we have no talent at all for it. Wendy Telfer took on the task for this book. Anyone but a super human would see it as a living nightmare. As we type these acknowledgements we have crossed fingers for her sanity.

A good index makes any book better. The index compiled by Kim Stringer for this edition is unbeatable, we think. She has made the book infinitely easier to use.

Finally, we would be lost without Sweda (or Ms Sweda) who was the Editorial Project Manager for this edition. No we don't know what this means either. It seems to involve everything that needs to be done and more. She has been a rock throughout and deserves to be lavished with the greatest praise. It has been a pleasure working with her. We wish we could be as nice.

Dennis Howitt

Duncan Cramer

Publisher's acknowledgements

Text:

3–4 American Statistical Association: Gordon, S. (1995). A theoretical approach to understanding learners of statistics. *Journal of Statistics Education* [Online], *3*(3) http://www.amstat.org/publications/jse/v3n3/gordon.html, para 18; **5 Hogrefe Verlag:** Zimprich, D. (2012). Attitudes toward statistics among Swiss psychology students. *Swiss Journal of Psychology*, *71*, 149–155; **58 Ingenta:** Cetinkalp, Z. K. (2012). Achievement goals and physical self-perceptions of adolescent athletes. *Social Behaviour and Personality*, *40*, 473–480; **59 Taylor & Francis:** Otgaar, H., Horselenberg, R., van Kampen, R., & Lalleman, K. (2012). Clothed and unclothed human figure drawings lead to more correct and incorrect reports of touch in children. *Psychology, Crime & Law*, *18*, 641–653; **59 Taylor & Francis:** van Schaik, P., & Ling, J. (2012). An experimental analysis of experiential and cognitive variables in web navigation. *Human Computer Interaction*, 27, 199–234; **87 Taylor & Francis:** Contador, I., Fernández-Calvo, B., Cacho, L. J., Ramos, F., & López-Rolón, A. (2010). Non-verbal memory tasks in early differential diagnosis of Alzheimer's disease and unipolar depression. *Applied Neuropsychology*, *17*, 251–261; **87 Taylor & Francis:** Di Filippo, G., de Luca, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2006). Lexicality and stimulus length effects in Italian dyslexics: Role of overadditivity effect. *Child Neuropsychology*, *12*, 141–149; **117 The British Psychological Society:** Blom, D., van Middendorp, H., & Geenen, R. (2012). Anxious attachment may be a vulnerability factor for developing embitterment. *Psychology and Psychotherapy: Theory, Research and Practice*, *85*, 351–355; **155 The British Psychological Society:** Rohmer, O., & Louvet, E. (2012). Implicit measures of the stereotype content associated with disability. *British Journal of Social Psychology*, *51*, 732–740; **155 John Wiley & Sons:** Gannon, T. A., & Barrowcliffe, E. (2012). Firesetting in the general population: The development and validation of the Fire Setting and Fire Proclivity Scales. *Legal and Criminological Psychology*, *17*, 105–122; **155 Taylor & Francis:** Vallat-Azouvi, C., Pradat-Diehl, P., & Azouvi, P. (2012). The Working Memory Questionnaire: A scale to assess everyday life problems related to deficits of working memory in brain injured patients. *Neuropsychological Rehabilitation: An International Journal*, *22*, 634–649; **163 American Psychological Association:** Mercer, S. H., Harpole, L. L., Mitchell, R. R., McLemore, C., & Hardy, C. (2012). The impact of probe variability on brief experimental analysis of reading skills. *School Psychology Quarterly*, *27*, 223–235; **201 American Psychological Association:** Critcher, C. R., & Dunning, D. (2013). Predicting persons' versus a person's goodness: Behavioral forecasts diverge for individuals versus populations. *Journal of Personality and Social Psychology*, *104*, 28–44; **201 American Psychological Association:** Siy, J. O., & Cheryan, S. (2013). When compliments fail to flatter: American individualism and responses to positive stereotypes. *Journal of Personality and Social Psychology*, *104*, 87–102; **203 American Psychological Association:** Mitsumatsu, H. (2013). Stronger discounting of external cause by action in human adults: Evidence for an action-based hypothesis of visual collision perception. *Journal of Experimental Psychology: General*, 142, 101–118; **203 John Wiley & Sons:** Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*, 1762–1774; **213 Elsevier:** Huisman, A., van Houwelingen, C. A. J., & Kerkhof, A. J. F. M. (2010). Psychopathology and suicide method in mental health care. *Journal of Affective Disorders*, *121*, 94–99; **213 American Psychological Association:** Abeyta, A. A., Routledge, C., & Juhl, J. (2015). Looking back to move forward: Nostalgia as a psychological resource for promoting relationship goals and overcoming relationship challenge. *Journal of Personality and Social Psychology*, *109*, 1029–1044; **213 American Psychological Association:** Rubin, J., Wynn, J., & Moscovitch, M. (2016). The spatial scaffold: The effects of spatial context on memory for events. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *42*, 308–315; **221 American Psychological Association:** Meyer, M. M., Bell, R., & Buchner, A. (2015). Remembering the snake in the grass: Threat enhances recognition but not source memory. *Emotion*, *15*, 721–730; **221 American Psychological Association:** Zhang, Y., & Risen, J. L. (2014). Embodied motivation: Using a goal systems framework to understand the preference for social and physical warmth. *Journal of Personality and Social Psychology*, *107*, 965–977; **223 John Wiley & Sons:** Gervais, S. J., Vescio, T. K., & Allen, J. (2012). When are people interchangeable sexual objects? The effect of gender and body type on sexual fungibility. *British Journal of Social Psychology*, *51*, 499–513; **223 John Wiley & Sons:** Lautamo, T., Laakso, M. L., Aro, T., Ahonen, T., & Törmäkangas, K. (2011). Validity of the play assessment for group settings: An evaluation of differential item functioning between children with specific language impairment and typically developing peers. *Australian Occupational Therapy Journal*, *58*, 222–230; **254 John Wiley & Sons:** Hoicka, E., & Akhtar, N. (2012). Early humour production. *British Journal of Developmental Psychology*, *30*, 586–603; **266 McGraw-Hill Education:** Adapted and extended from Table I of R.P. Runyon and A. Haber (1989). *Fundamentals of behavioral statistics*. New York: McGraw-Hill; **296 John Wiley & Sons:** Frank, G. K. W., Roblek, T., Shott, M. E., Jappe, L. M., Rollin, M. D. H., Hagman, J. O., & Pryor, T. (2012). Heightened fear of uncertainty in anorexia and bulimia nervosa. *International Journal of Eating Disorders*, *45*, 227–232; **397 American Psychological Association:** APA (2010). *Publication Manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association; **398 Taylor & Francis:** Casidy, R. (2012). Discovering consumer personality clusters in prestige sensitivity and fashion consciousness context. *Journal of International Consumer Marketing*, *24*, 291–299; **439 Butler, C:** Butler, C. (1995a). Teachers' qualities, resources and involvement of special needs children in mainstream classrooms. Unpublished thesis, Department of Social Sciences, Loughborough University; **459, 461 American Psychological Association:** Butler, R. (1995b). Motivational and informational functions and consequences of children's attention to peers' work. *Journal of Educational Psychology*, *87*, 347–360; **462 The British Psychological Society:** Gibbs, S., & Powell, B. (2012). Teacher efficacy and pupil behaviour: The structure of teachers' individual and collective beliefs and their relationship with numbers of pupils excluded from school. *British Journal of Educational Psychology*, *82*, 564–584; **479 SAGE Publications:** Munford, M. B. (1994). Relationship of gender, self-esteem, social class and racial identity to depression in blacks. *Journal of Black Psychology*, *20*, 157–174; **479 SAGE Publications:** Data from Munford, M. B. (1994). Relationship of gender, self-esteem, social class and racial identity to depression in blacks. *Journal of Black Psychology*, *20*, 157–174; **495 John Wiley & Sons:** Wagner, U., & Zick, A. (1995). The relation of formal education to ethnic prejudice: Its reliability, validity and explanation. *European Journal of Social Psychology*, *25*, 41–56; **495 The British Psychological Society:** Kuhnle, C., Hofer, M., & Kilian, B. (2012). Self-control as predictor of school grades, life balance, and flow in adolescents. *British Journal of Educational Psychology*, *82*, 533–548; **496 American Psychological Association:** Lamoureux, B. E., Palmieri, P. A., Jackson, A. P., & Hobfoll, S. E. (2012). Child sexual abuse and adulthoodinterpersonal outcomes: Examining pathways for intervention. *Psychological Trauma: Theory, Research, Practice, and Policy*, *4*, 605–613; **497 American Psychological Association:** APA (2010). *Publication Manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association; **578 Prof. Dr. Axel Buchner:** G*Power, © Copyright 2010–2016 Heinrich-Heine-Universität Düsseldorf; **605 EBSCO Industries:** Bridges, F. S., Williamson, C. B., Thompson, P. C., & Windsor, M. A. (2001). Lost letter technique: Returned responses to battered and abused women, men, and lesbians. *North American Journal of Psychology*, *3*, 263–276; **605 American Psychological Association:** Tracey, T. J., Sherry, P., Bauer, G. P., Robins, T. H., Todaro, L., & Briggs, S. (1984). Help seeking as a function of student characteristics and program description: A logit-loglinear analysis. *Journal of Counseling*

Psychology, *31*, 54–62; **623 The British Psychological Society:** Griffin, B., & Hesketh, B. (2008). Post-retirement work: The individual determinants of paid and volunteer work. *Journal of Occupational and Organizational Psychology*, *81*, 101–121; **623 Elsevier:** Huisman, A., van Houwelingen, C. A. J., & Kerkhof, A. J. F. M. (2010). Psychopathology and suicide method in mental health care. *Journal of Affective Disorders*, *121*, 94–99; **623 Elsevier:** Kogan, S. M. (2004). Disclosing unwanted sexual experiences: Results from a national sample of adolescent women. *Child Abuse & Neglect*, *28*, 147–165; **640 Taylor & Francis:** Kenne, D. R., Boros, A. P., & Fischbein, R. L. (2010). Characteristics of opiate users leaving detoxification treatment against medical advice. *Journal of Addictive Diseases*, *29*, 283–294; **650 SAGE Publications:** Hao, J., & Ho, T. K. (2019) Machine learning made easy: A review of Scikit-learn Package in Python Programming Language. *Journal of Educational and Behavioral Statistics*, *44*, 348–361; **681–682 McGraw-Hill Education:** Adapted from Runyon, R.P. and Haber, A., (1989) *Table I of Fundamentals of behavioral statistics*, The McGraw Hill Companies Inc.

Photo:

Chapter/Part Openers: Getty Images: Logorilla/DigitalVision Vectors/Getty Images; **23 Getty Images:** HishamIbrahim/Stockbyte/Getty Images

Screenshots:

Screenshots of IBM SPSS: Reprint Courtesy of International Business Machines Corporation, © International Business Machines Corporation; **521 Larry C. Lyons:** From The Meta-Analysis Calculator, http://www.lyonsmorris.com/lyons/metaAnalysis/index.cfm, Reproduced by permission of Larry C. Lyons; **573–580 Prof. Dr. Axel Buchner:** Screenshot of G*Power, © Copyright 2010–2016 Heinrich-Heine-Universität Düsseldorf.

CHAPTER 1

Why statistics?

Overview

- Students often approach learning statistics rather negatively. Everyone knows this but research demonstrates it too. Importantly, this leads to poor learning. Student culture tends to reinforce this already bad learning environment for statistics.
- The school environment is an especially important determinant of our attitudes to mathematics, which then impacts expectations concerning learning statistics.
- Mistakenly, some students believe that statistics is peripheral to professional psychology and other related careers. The truth is quite different. Professional psychologists rely on research based on quantitative methods and statistics to inform their work.
- Psychologists, practitioners included, are usually expected to carry out research as part of their work role.
- Knowledge-based practice characterises most modern professions which psychology graduates enter. So a good working knowledge of statistics is an advantage in the job market.
- Old and outmoded statistical ideas can make learning statistics unnecessarily difficult. Some of these ideas are unworkable. This contributes to a fog of confusion surrounding statistics. Unfortunately, textbook writers can be guilty of this.
- Null hypothesis significance testing receives far too much attention in psychology to the exclusion of more useful approaches. It is important to understand the much more varied contribution that statistics makes to psychological knowledge. There is growing dissatisfaction with aspects of statistics as practised by some psychologists.
- Few mathematical skills are needed to develop a good working knowledge of statistics. All but a few students have these skills. Even where these skills have got a little rusty, they can be quickly relearned by motivated students.

1.1 Introduction

For many psychology students the formula is simple: statistics = punishment. Statistics is 'sadistics'. Most would avoid statistics given the chance. This makes a very unpromising learning environment. And an unenviable teaching environment for the poor soul teaching such reluctant students. Student ratings for statistics modules are enough to bring tears to the eyes of all but the most hardened of professors and lecturers. Little could be less satisfactory. Couldn't statistics simply be omitted from psychology degrees? Well yes, but that is unlikely to happen. Statistics is central to most research that is studied on psychology courses. Surely many practitioners do much good without needing statistics? Even if this were once true, it is not so nowadays. Once research and practice were largely separate but modern practitioners combine practice with research. Wherever psychologists work, they are almost certain to have research as part of their job description. We are living in an information-based society and a great deal of this comes from statistical findings. So the bottom line is that some knowledge of statistics is professionally important – and not just in psychology.

Culturally, mathematics and consequently statistics are seen in a negative light. Remarkably, the average person has a poor opinion of statistics even without knowing much about what it involves. People groan when statistics is mentioned. They critique it with hackneyed phrases like 'you can prove anything with statistics' and 'lies, damned lies and statistics'. Statistics can be used misleadingly but that is not the usual intention. Of course, minor adjustments to a graph can distort whatever trend is found. A modest growth or decline in a graph may be presented as dramatic or calamitous simply by choosing what to show and how to show it. Nevertheless, statistics deserves greater respect than its reputation suggests. In contrast, the great majority of the general public have a positive attitude towards and interest in science (Castell, Charlton, Clemence, et al., 2014).

The word statistics derives from the Latin for state (as in nation). Statistics is the information collected by the State to help guide government decision-making. The government's appetite for statistical information is prodigious. Most areas of government planning are guided by statistical data – pay, pensions, taxes, health services, prisons, the police and so forth. Big supermarkets use it, charities use it, the health service uses it, industrialists use it – you name it and they probably use statistics though not quite the same statistics as psychologists. Sound statistical knowledge is fundamental to understanding, planning and analysing research. Nevertheless, students study psychology to learn about psychology – not statistics. However, the psychology they learn almost certainly involves statistics at some point. Of course there is qualitative research not involving statistics but this is a fraction of psychology's output. Statistics and psychology are intertwined.

Statistics, then, has a central role in psychology. It is not there to punish students – no matter how it feels. So why not try to see statistics as a sort of cuddly friend which will help you in all sorts of ways? We are serious here. Criticisms of the dominance of statistics in psychology are common, of course. As much as anyone else, we are against the mindless application of statistics in psychology for its own sake. Psychology may seem obsessed with a few limited statistical topics such as significance testing but this is to overlook the myriad of more far-reaching positive benefits to be gained from the proper application of modern statistical ideas. In recent years, the dominance of null hypothesis significance testing has been critiqued and calls for a new approach to statistics (Cumming & Calin-Jageman, 2017). We have incorporated these newer approaches into this text. Our view is that statistics provides a means of finding order in otherwise confusing data. Some of the various ways of doing this are illustrated in Figure 1.1.

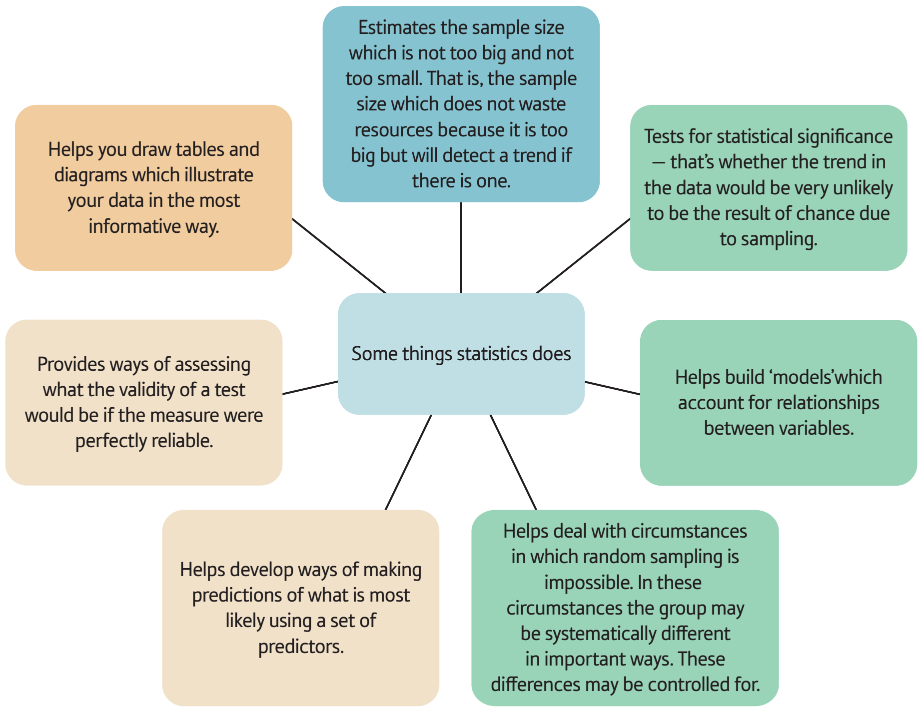


FIGURE 1.1 Some things statistics can do for the researcher

1.2 Research on learning statistics

Given society's endemic dislike of statistics, not surprisingly the research on psychology students and statistics makes generally depressing reading (e.g. Chew & Dillon, 2014). Trepidation and anxiety are characteristic responses to the prospect of studying statistics. Gordon (2004) talked to a large sample of Australian psychology students about their experience of statistics lectures. Three-quarters would not have studied it but it was compulsory. Statistics they saw as boring and difficult and felt that psychology and psychologists do not need it. They treated statistics as if it were a few mechanical procedures to be applied without understanding why. One student put it this way to Gordon (1995):

> I have a very pragmatic approach to university, I give them what they want. . . I really do like knowledge for knowledge's sake, but my main motivation is to pass the course. (paragraph numbered 18)

Even students who tried to master statistical methods and concepts had difficulty seeing its importance. Students who saw statistics as more personally meaningful in their studies said things like 'It would probably be useful in whatever job I do' (Gordon, 1995). As one might expect, the more positively orientated students did a bit better in statistics tests and examinations. Students with a negative orientation to statistics were generally not less able students and performed much the same as other students on other modules. But not seeing the point of statistics impacted their studies negatively. Figure 1.2 provides a broad classification of students in terms of how they see the relevance of statistics and their personal assessment of the discipline.

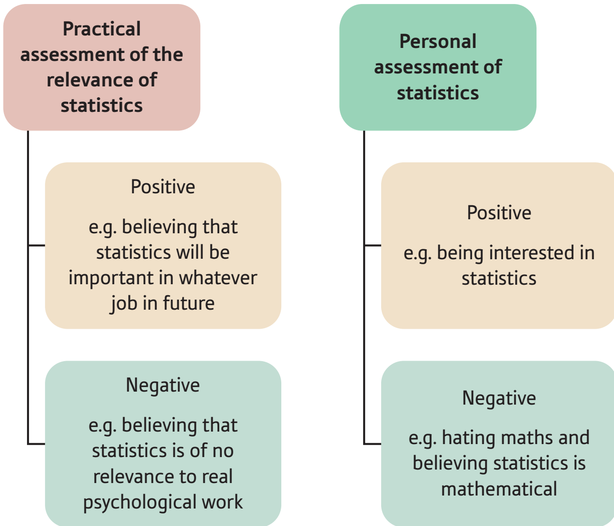


FIGURE 1.2 Responses of students to statistics according to Gordon (1995)

1.3 Why is learning statistics difficult?

Students tend to blame themselves for finding statistics a rather difficult subject. It makes them feel inadequate. Conscientious students should not blame themselves. Nor should they blame their lecturer for being a bad teacher. The villain of the piece is statistics itself quite patently. It has been argued that statistics with its primary concern with probability has fundamental unresolved issues which make learning difficult. Certainly, the evidence is that people do not think probabilistically at all well (Aksentijevic, 2015). Another problem is the fact that most statistical concepts are defined by a mathematical formula rather than in words. This makes it difficult to explain statistics in everyday language. Attempts to do so may be a little misleading but, worse still, criticised for being wrong. Statistics is a unique and distinctive way of thinking (Ben-Zvi & Garfield, 2004; Ruggeri, Dempster & Hanna, 2011). It has its own language and concepts. Some would say that many statistical ideas are decidedly odd. Grasping the statistical way of thinking and learning to speak statistical language takes some effort. Students in all sorts of disciplines struggle somewhat with statistics, it is not just psychology students. Statistical thinking is a different way of thinking. It also requires a rather different way of learning than most modules on a psychology degree. You can't afford to miss statistics lectures, much as this seems desirable, since statistical knowledge builds gradually in crucial steps. Miss any of these out and you are in trouble.

Learning statistics involves many emotional factors. University staff, for the most part, recognise that teaching statistics is made difficult because of students' anxieties, beliefs and negative attitudes concerning the subject (Schau, 2003). University life can be an experience full of emotion, and emotion affects learning – especially learning statistics. Real tears are shed. One student told Gordon (1995), 'I was drowning in statistics' – affective and extreme words but real. Being at university and studying statistics follows a long period of personal development through schooling (and for some at work). Personal histories, experiences, needs and goals are reflected in our strategies for coping with statistics (Gordon, 2004). These influence the way that we think about our learning processes and education more generally. Beliefs such as 'I'm no good at maths' will impact on our response to statistics.

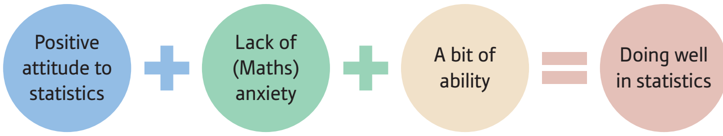


FIGURE 1.3 Formula for doing well in statistics based on research findings

In other words, students often bring baggage to learning statistics which seriously interferes with their studies. Issues concerning their mathematical ability are common. Some students incorrectly assume that their poor maths skills make statistics too hard for them. This view is reinforced by those departments which require good maths grades for admission. Given other time pressures, such students may adopt avoidance tactics such as skipping lectures rather than putting the time into studying statistics. Procrastination over one's work is known to be related to variables reflecting problems with statistics (Onwuegbuzie, 2004). Furthermore, every statistics class has its own culture in which students influence each other's attitudes to learning statistics. A class dominated by students antagonistic to statistics is not a good learning environment. Acting silly, talking in class or plagiarising the work of other students just does not help.

Whether mathematical ability is important to making a good statistics student is doubtful. There is, however, evidence that believing that good maths ability is needed to do well in statistics is undesirable. For example, Siew, McCartney and Vitevitch (2019) suggest on the basis of their research that students with high levels of anxiety about statistics should be specifically taught that there is no evidence of such a relationship. This aside, research strongly indicates that three factors – anxiety, attitudes and ability (see Figure 1.3) – are involved in learning statistics (Lalonde & Gardner, 1993). A negative attitude towards statistics is associated with poorer performances in statistics to some extent. Anxiety plays its part primarily through a specific form of anxiety known as mathematics (math) anxiety. Mathematics anxiety is common among psychology students. Those with higher levels of mathematics anxiety tend to do worse in statistics. This is more important in this context than trait or general anxiety which is when someone has a generally anxious personality in all sorts of situations. To be sure, mathematical ability is associated with better test and examination results, but not to a major extent. Poor mathematical ability has its influence largely because it is associated with increased levels of mathematical anxiety. That is, in itself, poor mathematical ability is not primarily a cause of worse results.

If more research evidence is needed, using a formal measure known as the Survey of Attitudes toward Statistics, Zimprich (2012) showed that attitudes towards statistics are made up of four components:

- Affect: How positive or negative a student is about statistics such as liking the subject.
- Cognitive competence: A student's beliefs about their ability and competence to do statistics
- Value: Attitudes concerning the relevance and usefulness of statistics such as using statistics in everyday life.
- Difficulty: The student's views about how difficult or easy statistics is.

Every one of these were interrelated, as one might expect. They also correlated with actual achievement in statistics. These attitudes were much more important than actual maths ability in terms of how well students do in statistics. In other words, how a student feels about statistics has a far more tangible effect on their performance in statistical tests and examinations than their mathematical ability. The lack of relationship between maths and statistics ability applies also to other aspects of studying for a psychology degree. Bourne (2018) investigated whether various mathematical abilities were related to performance on a British psychology degree. There was little relationship between maths ability and performance on the research component of the degree. One exception was that good graphical abilities were related to marks on aspects of research methods but only in the first year of study. Otherwise there was no relationship.

Along with others, we would argue that the level of mathematical ability needed to cope with the mathematical part of statistics is not great – fairly minimal in fact. Mostly, though, the statistical analyses you need are available on SPSS and other statistics programs. Generally speaking, rarely will you need to do calculations by hand and then these are usually simple. Often you will find websites which will calculate the few things which SPSS does not do. Diligence in making sure that your data has been correctly entered plus some knowledge of the appropriate statistical analysis are the important things. Some basic mathematics is helpful in that numbers and symbols can be daunting, at first. Statistics is a maths-based discipline and its concepts are generally defined by formulae rather than words. So if you are good at understanding mathematical formulae then this is an advantage, though far from necessary. Professional researchers differ widely in their mathematical skills and many do not regard themselves as at all mathematical. Yet they have learned to use statistics appropriately and intelligently, which is very much the task facing students. You need to understand the purpose of a statistical test and why it was developed, understand a little about how it works, know when to use it and most of all be able to make sense of the computer output. Maths is peripheral for the most part.

If you understand the concepts of addition, subtraction, multiplication and division then you have the basics for coping with statistics. You may not always get the right answer but the important thing is that you understand what these mathematical operations do. What might you need beyond this? Probably just the following:

- You need to understand the concept of squaring (that is multiplying a number by itself).
- You need to understand the concept of square root (the square root of a number is that number which when multiplied by itself gives the original number).
- It is good too if you understand negative numbers – such as that when multiplying two negative numbers you get a positive number but when you multiply a positive number by a negative number then the result is a negative number. A short time spent trying out positive and negative calculations on a calculator is a good way to refresh yourself of these basics.
- It is preferable if you understand the underlying principles or ‘rules’ governing mathematical formulae as these are used in statistical formulae, but if you don’t, your computer does.

Little else is necessary – if you know what a logarithm is then you are in the ultra-advanced class. All in all, the requirements are not very demanding. Anything that has been forgotten or never learned will be quickly picked up by a motivated student. Not all lecturers will share this opinion. Nevertheless, the overwhelming majority know that students can really struggle with statistics for any number of reasons. So they provide teaching which serves the needs of all students taking the psychology programme, not just the maths-able ones. Interestingly, research suggests that the more approachable the statistics lecturer is the less statistics anxiety manifested by their students (Tonsing, 2018).

1.4 The importance of understanding research designs

Broadly speaking, different research designs require different statistical techniques. So a fundamental requirement is that you appreciate the different kinds of research design and what they can achieve. Statistical problems in research are often fundamentally research design problems. You really do have to formulate your research question, your hypotheses and your research design carefully for the statistical analysis to fall into place. If you don't have confidence that your research design is capable of addressing your research question, statistics cannot redress the problem. As you study statistics, you should gradually see ever more clearly that most statistical techniques are appropriate for particular research designs. For example, analysis of variance statistics are ideally suited to analysing randomised experiments.

Every degree course will give you a grounding in research methods and how research is done. But such knowledge will not translate directly into an ability to do research. This is developed through practical or lab classes in which you experience the process of doing research. Although research skills build up quite slowly over the course of your degree, these skills are little or nothing to do with mathematics. They are about the application of logic and thought to the research process. That is why being comfortable with your ability to design valid research comes before being able to choose suitable statistics. If you are confused about your research question, your hypotheses and your research design, it follows that you will be confused about the appropriate statistical analysis. So statistical analysis takes a minor role compared to the more general research skills involved in a quantitative study. You will find that once you can identify the type of research design you are using, the task of choosing appropriate statistical techniques for its analysis becomes almost self-evident.

1.5 Positive about statistics

It is clear that having a negative attitude towards statistics, although somewhat understandable, is counterproductive in many ways. Many of your fellow students will feel much the same. Professional researchers often make little claim to having any statistical expertise (Aksentijevic, 2015) but they still do excellent work. Nevertheless, it is important to deal with bad study techniques in relation to statistics such as avoiding or day dreaming through lectures. So how is it possible to be positive about statistics? Having an appreciation of what statistics contributes will help. Just why did statistics become so important in modern research when for centuries people did experiments and other research without statistics procedures such as significance testing? One of the most well-known statistical techniques used by psychologists is the *t*-test (see Chapters 13 and 14) or the Student *t*-test as it is also known. For decades, psychology students have learned to do the *t*-test. Student was the pen name of William Gosset who had studied chemistry and mathematics at university. He was employed by the Guinness Brewery in Dublin as a 'bright young thing' in the 1890s.

Quality control was important to the company. Obvious practical problems would follow if every bottle or barrel of beer had to be tested perhaps to see if the alcoholic strength was constant. One solution would have been to use just a small number of samples. Gosset worked on the extent of error likely to occur when sampling is employed and developed a mathematical way of calculating the likely error. For example, if a sample of just 10 bottles is taken, to what extent are these likely to mislead quality controllers about the alcoholic strength of the product in general?

Of course, you will never know from a sample exactly what the error will be but Gosset was able to estimate the likely error from the variability within the sample of bottles. Put into a formula, this is the idea of standard error which plagues many students on introductory statistics courses for countless decades. The *t*-test is based on standard error. Gosset had laid down a systematic basis for research to use samples rather than everything. Think about it: if it had not been for Gosset's innovation then you would spend your lifetime carrying out your first research study simply because you need to test everyone or everything (the population).

So rather than considering William Gosset as some sort of alien, would it not be better to see him as one of the statistical cuddly friends we mentioned earlier?!

Is it statistically significant?

The point of Gosset's revolutionary ideas is probably easy to see when explained in this way. But instead students are introduced to what to them are rather complex formulae and the question 'Are your findings statistically significant?' The question 'Is it significant?' is one of the fixations of many psychologists – the question probably sounds like a mantra to students when they first begin to study psychology. So intrusive is the question that for most students, statistics in psychology is about knowing what test of statistical significance to apply to their data. A test of statistical significance merely addresses the possibility that a trend that we find in our sample might have occurred by chance when there is no trend in reality. That is, how likely is it that the trend is simply the result of a fortuitous selection of a sample which appears to show a trend? (A trend might be, say, athletes scoring more highly on a measure of personal ambition than non-athletes or a relationship between a measure of ability to speak foreign languages and a measure of sociability.) But significance testing is only a small part of statistics, which provides a whole range of tools to help researchers (and students) address the practical problems of data analysis. Over the years, psychology has been criticised for the way statistics is used. A central focus of the criticism is the fixation of significance testing.

What sample size do I need?

Gosset's focus on small samples begs the question of how small a sample can be used. There would be something perverse about planning research which involved a sample size so small that our findings could never be statistically significant. But that is done inadvertently all of the time simply because researchers (including students) do not address the question of sample size properly. Often the advice given to those asking what sample size to use is that they should get as big a sample as they can. But this is a crude way of going about deciding sample size. Even the smallest trend will be statistically significant if the sample size is large enough. However, there is little point in using large samples when smaller ones would be adequate. The optimum sample size depends on the size of the effect the researcher thinks is worthwhile investigating, the statistical significance level required and the risk of not supporting the hypothesis when it is in fact true that the researcher is prepared to take. There are conventional values for the latter two but the researcher may wish to vary these.

There are no objective criteria which tell us what potential size of effect is worth studying which apply irrespective of circumstances. It might appear obvious that research should prioritise large trends but it is not as simple as that. In medical research, for instance, there are examples of very small trends which nevertheless save lives. Taking aspirin has a small effect on reducing the risk of heart attacks but saves lives in aspirin takers compared with a control group. The size of a trend worth the research effort therefore depends on what is being considered. A pill which prevents cancer in 10% of people would be of more interest than a pill which prevents flatulence in 10% of people, for example. So if a researcher designs a study which has a sample size too low to establish a statistically significant trend then this would be more worrisome in the case of the cancer cure than the flatulence cure. Chapter 39 explains how to go about deciding sample size in a considered, rational way. This area of statistics is known as *statistical power analysis*. So the apparently simple question of the sample size needed is rather more complex than at first appears.

This is not the place to give a full overview of the role of statistics in psychological research. It is important, though, to stress that statistics can help research in many ways. This is hardly surprising since statisticians seek to address many of the problems which researchers face in their quantitative research. Now this book is just about as comprehensive as understandable statistics texts get but not everything that statistics can do is represented. Nevertheless, you will find a great deal which goes far beyond the issue of statistical significance. Take, for example, factor analysis (Chapter 33). This is not at all about statistical significance but a way of finding or identifying the basic dimensions in your data. So, for example, many famous theories of personality and theories of intelligence have emerged out of factor analysis – for instance, the work of Hans Eysenck (Eysenck & Eysenck, 1976). This suggests that extraversion, neuroticism and psychoticism are the major underlying dimensions or components of personality on which people differ. There is no way that a researcher can simply look at their data, which can be enormously complex, and decide what its underlying structure is. It is not possible to identify extraversion, neuroticism and psychoticism simply by looking at the scores from a 50-item questionnaire that has been completed by 2000 participants. Nevertheless, statisticians (and psychologists) developed factor analysis as a method of doing just that.

Statistics also has a very important role in model building. This sounds complicated but it isn't too difficult. A model is simply a proposed set of relationships between variables. So, for instance, the relationships shown in Figure 1.3 between various characteristics of students studying statistics and their achievement in tests and examinations is a sort of model. Statistics addresses just how well the data fits the proposed model – there may be other characteristics of students to be considered in addition to those in Figure 1.3 to fully account for how well students do in statistics. The researcher may propose models but, equally, statistical techniques can also suggest them.

Some of the other things which statistics can help you with include:

- Is the trend that I have just found in my data big or small?
- Does this line of research show potential for further development?
- Are the measures that I am using sufficiently reliable and valid to detect a trend that I am interested in?
- Is it possible to amalgamate a number of variables into a single, more readily understood one?
- Can I eliminate competing explanations of my findings so as to give more credence to my hypothesis?
- How best can I present my data graphically in order to visually present my findings to an audience at a conference?
- Can I combine the findings of different studies so as to have a good idea of the typical findings of past research?

Statistics is just one aspect of the decision-making processes which underlie psychological research. It should not dominate a researcher's thinking exclusively. It is not even the most important part of research. But without it your decision-making may be sub-optimal.

1.6 What statistics can't do

Years of experience teaching statistics means, of course, that we were the statistics doctors that students having problems with their data analysis data came to – or even got sent to. These encounters vary widely. Some students appear to want help but really they are seeking confirmation that their ideas for their analysis are correct or that they have understood their data correctly. Yet others have designed their research so badly that either it is difficult to analyse at all or it is difficult to analyse using the statistics that the student knows at this point. A few students cannot relate what they learned in statistics lectures to their own research.