Eighth Edition

Understanding Statistics in Psychology with SPSS

Dennis Howitt
& Duncan Cramer

Pearson

Understanding Statistics in Psychology with SPSS

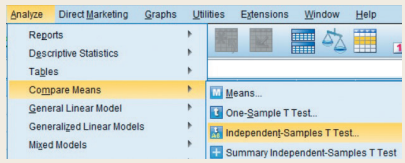
Interpreting and reporting the output

- The means tell you which group has the highest scores. Normally the highest scores mean a higher amount of the variable. In this example, emotionality was found to be higher for the children of two-parent families. The significance level should be noted as this tells you whether the difference is likely to be the product of chance.
- Reporting the results can following this pattern: 'It was found that emotionality was significantly higher, $t(20) = 2.81$, $df = 20$, $p < .05$, 95% CI [−6.82, −1.01], in the two-parent families ($M = 13.42$, $SD = 3.37$) than in the lone-parent families ($M = 9.50$, $SD = 3.10$).' CI stands for Confidence Interval and is discussed in Chapter 16.

	Family	Emotion
1	2	12
2	2	18
3	2	14
4	2	10
5	2	19
6	2	8
7	2	15
8	2	11
9	2	10
10	2	13
11	2	15
12	2	16
13	1	6
14	1	9

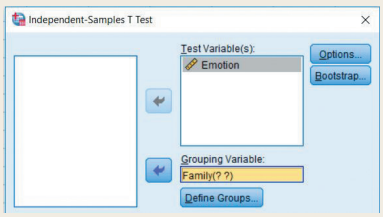
Reprint Courtesy of International Business Machines Corporation, © International Business Machines Corporation

SCREENSHOT 14.1 Part of data in 'Data View'



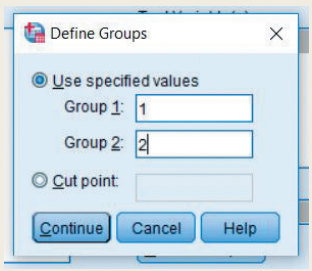
Reprint Courtesy of International Business Machines Corporation, © International Business Machines Corporation

SCREENSHOT 14.2 On 'Analyze' select 'Independent-Samples T Test... '



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

SCREENSHOT 14.3 Select variables for the analysis



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

SCREENSHOT 14.4 Define the two groups of scores

Group Statistics

	Family	N	Mean	Std. Deviation	Std. Error Mean
Emotion	1	10	9.50	3.100	.980
	2	12	13.42	3.370	.973

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Emotion	Equal variances assumed	.212	.650	-2.813	20	.011	-3.917	1.392	-6.821	-1.013
	Equal variances not assumed			-2.836	19.768	.010	-3.917	1.381	-6.800	-1.034

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

SCREENSHOT 14.5 Main output

CHAPTER 15

What you need to write about your statistical analysis

Overview

- A glance at reports in psychology journals shows that the reporting of outcomes of statistical analyses is brief and succinct.
- Things can look complicated because researchers often report (very similar) multiple statistical analyses in the same paper.
- The American Psychological Association recommends reporting the abbreviated name of the statistic you are using, the value of this test statistic, the degrees of freedom or the sample size as appropriate, the significance level, whether a one-tailed test of significance is used, confidence intervals, effect size and statistical power if possible. Many psychology journals including those of the British Psychological Society emulate this style.
- These are recommendations for professional publications but are appropriate for student reports at nearly every stage.

Preparation

You need to know about testing significance, from Chapter 11 onwards.

15.1 Introduction

A glance at any psychology journal shows that precious little space is devoted to reporting tests of significance and their outcomes, also that there are standard ways of reporting statistical analyses. These make it easier for even novice researchers to report their statistical findings in an acceptable fashion. Ways of reporting statistics have changed somewhat over the years and not all research journals use the same style but many do. The major influence on writing up research in psychology is the *Publication Manual of the American Psychological Association* (APA, 2010). Researchers wishing to publish their work in an APA journal, a British Psychological Society (BPS) journal, and in many other places need to conform to its style recommendations. As a student, you should check what is required by your department but it is likely to be APA style or a variant of it.

The current *Publication Manual* stresses the importance of not merely reporting information about the test of significance used but also (a) confidence intervals and (b) effect sizes. These are discussed in detail in Chapters 16 and 17, respectively. They are not particularly difficult ideas to understand and are fairly easily incorporated into your reports. Confidence intervals are given in most of the output of SPSS. Things are a little less satisfactory with effect sizes since SPSS does not provide too many of these though they are generally easily calculated using simple formulae where they cannot be obtained from SPSS.

What is a confidence interval? Briefly, in statistical significance testing the most probable value of the population parameter is given as a single figure based on the information we have from our sample or samples. So if the mean of the sample is 2.51 then we estimate the mean of the population as being 2.51, for example. This is known as a point estimate as it consists of a single value. However, this is only the best estimate and, as such, does not indicate anything about the probable range within which the population mean is likely to fall. We all know that samples usually vary somewhat from the population from which they came. A confidence interval (CI) is a range of scores within which a population parameter (such as the population mean) plausibly could fall. So a confidence interval might be 1.21 to 3.81 in this case. Confidence intervals are usually expressed as a percentage such as the 95% confidence interval. What this means is that the population parameter is likely to be within the stipulated confidence interval 95% of the time. It is possible to have confidence intervals for many of the statistics given in this book and SPSS usually calculates them for you routinely. More on confidence intervals and how they are calculated can be found in Chapter 16.

What is effect size? An effect size is just that – the size of the correlation or difference found in your study. So, in an experiment, the effect size would be the difference between the mean score in the experimental condition and the mean score in the control condition. In another study, it might be the size of the correlation between two variables. When analysing data, among the most important things to look at is the size of the effect. It would seem very obvious to include the effect size in your reports but it can be overlooked if one is too focused on statistical significance testing. So means and standard deviations for, say, the experimental and control groups need to be in the report as should be the size of the correlation coefficients that you have calculated. Without this, the reader will struggle to appreciate the importance of your findings. There is nothing difficult about effect sizes except that sometimes they are reported in standardised form to allow easy comparison between studies. Actually the correlation coefficient is a good measure of effect size in itself but others include Cohen's *d* and many more. The calculation of effect size is discussed in Chapter 17.

These are the basics. This chapter concentrates on reporting the fundamental features of your statistical analysis and mainly concentrates on statistical significance. Confidence intervals and effect size can be added in once you have studied Chapters 16 and 17.

15.2 Reporting statistical significance

At a minimum, the following should be mentioned when reporting statistical significance:

- The size of the effect.
- The statistical distribution used (F, chi-square, r, z, t, etc.).
- The degrees of freedom (df). Alternatively, for some statistical techniques you may report the sample size (N).
- The value of the calculation (e.g. the value of your z-score or your chi-square).
- The probability or significance level. In older journal articles you will find the following used: 'not significant', 'not sig.' or '*ns*'. However, statistical calculations are now done almost exclusively on computers and it is most usual now to give the exact probabilities for your significance test. These are provided routinely by SPSS.
- If you have a one-tailed hypothesis then this should be also mentioned. Otherwise a two-tailed hypothesis will be assumed by the reader.

As your statistical skills develop, you might wish to add confidence intervals and standardised measures of effect size.

15.3 Shortened forms

In research reports, comments such as the following are to be found:

- The hypothesis that drunks slur their words was supported ($t = 2.88$, degrees of freedom $= 97$, $p < .01$).
- The null hypothesis that drunks do not slur their words more than sober people was rejected ($t = 2.88$, $df = 97$, $p = .01$).
- The hypothesis that drunks slur their words was accepted, $t(97) = 2.88$, $p = \ < .003$, 1-tail.

Each of these says more or less the same thing. The symbol t indicates that the t-test was used. The symbol $<$ indicates that your probability level is smaller than the given value. That is, the test is statistically significant at better than the reported level of .01. Sometimes, the degrees of freedom (df) are put in brackets after the symbol for the statistical test used, as in $t(97) = 2.88$. All of the above examples are statistically significant at the .0 or 5% level and so the null hypothesis is rejected.

The following are examples of what might be written if the hypothesis was not supported by your data:

- The hypothesis that drunks slur their words was rejected ($t = 0.56$, degrees of freedom $= 97$, $p > 0.05$).
- Drunks and sober people did not differ in their average rates of slurring their speech ($t = 0.56$, $df = 97$, not significant).
- The hypothesis that drunks slur their words was rejected, $t(97) = 0.13$, $p = .45$, *ns*, 1-tail.

All of these mean much the same. The symbol $>$ means that your probability is greater than the listed value. Notice that each of these examples in some way states that the finding is not significant ($p > 0.05$, *ns*, and not significant all mean the same).

The significance level determines whether we accept or reject the statistical null hypothesis. The significance level is usually set at .05 or 5% but it is not unusual to see the .01 or 1% levels used. If the probability is equal or less than this then we reject the null hypothesis in favour of the alternative hypothesis. Computers do far more complex calculations than are practicable by hand. In particular, they work out the exact probability for your statistical test. So instead of writing that a difference is significant at the 5% level we can give the exact significance, such as .037. Our significance level for rejecting the null hypothesis still remains at 5%. The current American Psychological Association (APA) style is to report the exact significance. One possible objection to this is that it gives a false sense of precision to the statistical findings. Statistical significance can become a holy grail in statistics, supporting the view 'the smaller the probability the better'. We have seen the variability that is possible in randomly selected data so we should be very cautious about assuming that a significance level of .003 is really better in some sense than a significance level of .006. On the other hand, the 5% is a somewhat arbitrary but traditional criterion questioned by some. It has been suggested that by giving the precise significance it is absolutely clear whether the significance level was, say, .049 or .051. Pedantically, .049 would be statistically significant but .051 would not be, though the difference is miniscule.

Remember, statistical significance is important but the size of the effect is arguably more so. A significant result with a strong trend is the ideal which is not obtained simply by exploring the minutiae of probability. If you are using exact probabilities, then make it clear what significance level you are using to reject the null hypothesis. This significance level is known as alpha so you could write 'Throughout the analysis an alpha level of .05 was used for all statistical tests.' The reader may have assumed this but it is best to be absolutely clear. It would not be common but the alpha criterion could be varied.

15.4 APA (American Psychological Association) style

The *Publication Manual of the American Psychological Association* sets out the ways in which manuscripts should be typed to be considered for publication in the Association's journals. The latest version is the sixth edition and was published in 2010. It is claimed that about 1000 journals worldwide use APA style. The recommendations for reporting statistics seem to be relatively straightforward although there are a lot of them. The main ones are:

- Generally, report numbers to no more than two decimal places.
- Probability or significance values are the exception to this. These may be reported to three decimal places. APA style asks for exact probabilities to be given down to the .001 level of significance. So if your significance level starts with .000 (e.g. 0003) then report this as $p < .001$. (Giving significance levels such as .000 is confusing because it is not the same as zero.)
- Leading zeroes (i.e. zeroes before the decimal point) should not be used for numbers which cannot be more than 1.00 such as correlation coefficients. For example, correlations should not be reported with a leading zero such as 0.671 but as .671.