

5 Coding, making the database, and reliability analysis

Having got your questionnaires and tests completed by as many people as you need, you can now enter the data onto a computer database, coding the information where necessary, and ensuring that the records of each person are not named. Then you will work out which are the best, most reliable, items in your scale and discard the rest. You will then have a ‘final’ scale consisting of these reliable items, on which you can work out a final total for each participant and produce descriptive statistics for this final scale (norms). This final scale will then be ready for validation work.

For each stage we describe what has to be done and present a general explanation of how to do it. [Appendix 5](#) provides examples of how to do it in IBM SPSS.

Coding, scoring and data entry

Coding

Responses that do not need coding

1. Most answers that are already in numerical form, such as *age*, and the *answers to scale items*, if these were given in a numerical form – for example +2, +1, 0, –1, –2, or 5, 4, 3, 2, 1. You do not need to do anything except enter them as they are onto the computer database.
2. Answers with a *yes/no* response format can also be entered directly onto the computer database: 1 for a ‘yes’ and 0 for a ‘no’. Similarly, if you were using a checklist format, answers to each item would be entered as 1 where an item had been checked, and 0 where it had not. (Strictly speaking, this *is* coding, but it is so simple that you can enter the coded responses straight onto the computer database.)
3. *Forced-choice* format answers, where your participants have to select one of several alternatives. You can enter 1 for items selected, and 0 for those not selected. Later, you can compute scores from this information.

Table 5.1 Examples of items that do not need coding

<i>Type of item</i>	<i>Example</i>	<i>Example of answer</i>	<i>What you enter on database</i>
Item requiring numerical answer	Age	23	23
Likert-type scale item	Liking for this chocolate (1 = very low, 2 = quite low, 3 = moderate, 4 = quite high, 5 = very high)	4	4
Yes/No or checklist	Mark words that apply to this chocolate:		
	Sweet	X	1
	Creamy	X	1
	Bitter		0
	Spicy		0
Forced-choice	Which one would you choose for your next snack:		
	This chocolate	X	1
	Another chocolate		0
	Another snack		0

Table 5.1 shows examples of answers that do not need coding and can be entered directly onto the computer database. If no answer has been given, leave the corresponding space in the database blank, to indicate missing data. There is no need to enter anything special to represent missing information, even though this used to be necessary for older statistics packages. Alphabetical information needs special treatment beyond the scope of this book. Therefore, as just indicated, translate ‘yes’ into ‘1’, and ‘no’ into ‘0’.

Information that needs coding

Other information will need to be coded so that you can do statistics. To speed things up and reduce errors, many researchers pre-code their questionnaires by placing the appropriate code next to each response option. Consider the following for example:

How much do you like this chocolate? (tick one answer) (Office use only)

Very much	5
Quite a lot	4
Moderately	3
Not very much	2
Not at all	1

The codes are put in an out-of-the-way column, perhaps with that useful heading 'Office use only'.

The following are types of information that need coding:

1. If your participants answered your scale items by checking one of several ordered responses, instead of writing a number: for example 'strongly agree', 'agree somewhat', 'neutral', 'disagree somewhat' or 'strongly disagree'. In such a case you simply translate the answers into a simple ordered number scale. In this case, it could be +2, +1, 0, -1, -2, or 5, 4, 3, 2, 1. It does not matter which, as long as you *consistently* apply the *same* scale to *each* item. You can use this procedure for a scale with any number of points: 3, 5, 7, 9 and 11 are the most common. You might wonder what to do about '*reverse-meaning*' items, where agreeing with such items suggests an 'opposite' set of beliefs or attitudes or feelings, compared to agreeing with other items. In such cases, you still give the same 'score' to 'strongly agree' and so forth. However, later, you will have to tell the computer to *reverse* the scoring of such items. [Appendix 5](#) offers instructions.
2. A *Visual Analogue Scale* (VAS) is a horizontal line across the page, on which the person is asked to place a mark to indicate their feelings or beliefs. [Figure 5.1](#) shows an example. As stated in [Chapter 2](#), responses to these scales are more time consuming to score than other types of response, so consider whether it is worthwhile to use this method.

The example suggests that the person tested perceived the chocolate as sweet and creamy, bland and rather boring. With a VAS, you score the answers by measuring the distance in centimetres or millimetres from the left-hand end of the line. This number is entered onto the database. When describing variables assessed on a VAS, use the word on the *right-hand* side of the scale. If necessary, reverse the scoring (by subtraction) once all the data are entered in the database.

3. If you have asked for background socio-demographic data from your participants, some of these (such as gender, occupation or group membership) will need to be translated into an appropriate number scale. You should be aware of what *type of number scale* (interval, ordinal or categorical) you are using because this will affect the kinds of statistics that are appropriate. Consult [Appendices 5.1 and 5.2](#) if you feel doubtful.

When to calculate total scores

If you are using a computer, there is normally no point in calculating people's total scores on the scale by hand. Assuming you are going to do a reliability analysis, the computer software will calculate the totals it needs. Following reliability analysis, you will discard some of the items on the

On each line, place a mark to indicate how this chocolate tastes:

For example, if you think the chocolate tastes very pleasant, put a mark very near the word pleasant, like this:

Pleasant x _____ Unpleasant

If you think the chocolate is quite unpleasant, place a mark quite near the word unpleasant:

Pleasant _____ x _____ Unpleasant

If you think the chocolate is neither pleasant nor unpleasant, put a mark near the middle:

Pleasant _____ x _____ Unpleasant

Mark each line to show what you thought of the chocolate:

Sweet x _____ Bitter
 Boring _____ x _____ Exciting
 Sharp _____ x _____ Creamy
 Bland _____ x _____ Tasty

Figure 5.1 Visual Analogue Scales (VASs), with instructions.

scale, and then you will be able to calculate total scores on your refined scale. You will be able to do this using the computer.

However, it is a good idea to reverse the scoring of any ‘reverse-meaning’ items as soon as possible after entering all the data, and to save these reversed scores. This chapter and its appendices describe how to do this.

So, total scores on the scale are not computed until *after* the data are entered, reverse-meaning item scores are reversed, reliability has been investigated, and non-cohesive items have been discarded.

Make a database

This can be done using SPSS, or another statistics package. If necessary, in the absence of a statistics package spreadsheet, a modern word-processing or spreadsheet package could be used. But you will eventually have to import the data into a statistics package for processing. If you are making the database using software that is different from that you will be using for statistics, make sure that your statistics package can use the database, or that

you can translate the database into a form that can be used by your statistics package. SPSS, for example, will be able to import Excel, Text and ASCII files, and even system files made for pre-Windows versions of SPSS.

With Text and ASCII files, you will have to work hard at defining the variables. It is much better to enter your data into the spreadsheet of the statistics package you will be using.

To enter data,

- Start with an ID number for each participant (person who worked on your test). This will be the first column on the spreadsheet.
- Name all the remaining columns with the variable names, such as age, gender, item1, item2, etc. If wished, you can take the opportunity to define what each code number represents, for example 0 = male, 1 = female.
- Enter the coded responses to your questions, including the answers to the main part of the questionnaire, which can be entered just as they are, if they are in numerical form. Data should normally be in numerical form; seek advice if you think you have to enter alphabetical information.
- Make sure you start a new line or row for each participant.
- Make sure that you have the right number of pieces of data for each participant. Thus each line of data should finish in the column of the last variable. This sounds obvious, but if you are doing a lot of data entry, or if you are inclined to rush and make mistakes, you may find that you have entered all the data for one participant, but you are a column too early or a column too late. This means that you will have to go back and check until you find where the error was made.
- If any answer is missing, leave the space in the spreadsheet blank. (Only if you are using word-processing software will you need to assign a number for all *missing values*. Decide on a number that is not otherwise going to be used [-1, 9 and 99 are popular for missing values], and put that whenever an answer is missing.)

If you get stuck, consult Field (2017) or Pallant (2016) or any other book on SPSS or computing for social scientists or psychologists, or ask a friend with relevant computing experience, or a computing adviser.

Remember to *save* the database as you go along, especially if you have a big database. Remember that SPSS has the undesirable feature that it does *not* automatically save your work as you go along. Also, make a backup copy in the cloud and/or on another device. Normally you would continue to use the same filename each time you save, over-writing the older version.

Check for accuracy

Once the database is complete, you should carry out one or more checks for accuracy, before you can start statistical analyses. [Appendix 5.3](#) discusses in more detail the following procedures:

- Scan the database for numbers that should not be there. For example a common ‘typo’ in data entry is of the type ‘25’, where you should have typed ‘2’ and ‘5’ in adjacent cells.
- If you are using a statistics package, ask for frequencies of variables and scan the results. When you scan, look for unlikely or impossible values, such as ‘2’ or ‘99’ on age, in a study where you know that all participants were aged 18–65, or ‘0’, ‘6’ or ‘25’ on a scale with a range of 1–5. Locate any error on the database, find the true value, and enter it.
- Finally, if possible, carry out a check on the printout of the database. This is easiest if working in pairs. One member of each pair looks at the original data, and the other looks at the database. One person reads out the data. It does not matter which person – take turns. The other person calls out when they spot a mismatch. Correct all mistakes on the computer database, and save the corrected version.

Reverse scoring and saving the database

You can now

- *Recode* any negatively worded (reverse-meaning) items
- *Deal with missing values*, if necessary (see the end of [Appendix 5.3](#) for details)
- *Save and back up your data file*

[Appendix 5](#) provides instructions for doing this using SPSS.

Selecting reliable items

Instructions for computing reliability appear in [Appendix 5](#).

The first reliability statistics

[Chapter 1](#) defines the main types of reliability. SPSS or other statistics packages with a reliability facility ([Chapter 1](#) lists some packages) will normally give you one reliability statistic (Cronbach’s alpha) by default, and you can specify others if needed.

If you want to understand more fully what the different reliability options mean, [Chapter 1](#) gives an outline. You should also study the statistics package handbook, or Field (2017), Pallant (2016) and/or a handbook of test construction such as that by Kline (2015) or Urbina (2014).

The first reliability statistic to look at is Cronbach's alpha, a widely used reliability coefficient. This is normally sufficient. This is the estimated correlation of the test with any other test of the same length with similar items (i.e. items from the same item universe). The square root of alpha is the estimated correlation of the test with true scores.

What are the *criteria of acceptability for reliability coefficients*? Nunnally (1978), Lance, Butts, and Michels (2006) and Urbina (2014), among others, suggest that 0.70 might be acceptable, but normally one should be aiming for 0.80 or above. If you have a scale with a small number of items, you are not likely to get reliability coefficients as high as this, and you may consider using a slightly lower criteria (of about 0.60) *if (and only if)*

- There is good evidence for validity
- There are good theoretical and/or practical reasons for the scale
- The scale is short (less than about 10 items)

Some test constructors develop tests that are long and *repetitious*. These features are likely to ensure high reliability coefficients. However, it is *not* very good practice to write tests in this way. Participants may get bored or suspicious if the same sort of question is asked over and over again, and improving internal consistency by these dubious means will not improve validity. It is better to settle for an alpha of around 0.60, given the conditions listed earlier.

However, there are legitimate steps you can take to improve the reliability of your scale. The first thing to try is 'cleaning up' your scale by weeding out those items that are lowering the internal cohesiveness. The section that follows (headed 'Improving reliability of the scale') describes how this can be done.

The K-R 20 (Kuder-Richardson 20) is a special case of the alpha coefficient, for items that have been dichotomously scored. You should not have to do anything to get this statistic. Your reliability facility should produce it automatically when there are only two values on a variable, instead of Cronbach's alpha. It can be interpreted in the same way.

Your reliability facility can compute other reliability coefficients (split-half, for example – see your menus or handbook for the full range of possibilities), but normally Cronbach's alpha, probably the default option, should be used. This is regarded by Cronbach (1951) and Nunnally (1978)

as the most important index of test reliability. The split-half reliability coefficient could be useful if you have only a small number of items in a scale or subscale, since in this situation you may not get a sufficiently high alpha.

If you think there might be some subscales within your overall scale, you can get alphas for the subscales, by computing alphas for just those items comprising each subscale. For instance, in our prayer example the scale was written with two kinds of items: those valuing prayer as a way of getting what one wants (instrumental), and those looking at prayer as a form of strengthening or inspiring communication with the divine (inspirational). You could look at the reliabilities of subscales like these. [Appendix 5](#) gives instructions for computing and improving reliability of scales, and the same procedures are followed for any subscales.

If your test or scale used a *forced-choice* format, you may need to compute the reliabilities of the different subscales covered by the choices offered to the participants. For example, if you wanted to find out people's habitual distress mode, you might present a number of items like this:

When someone pushes in front of me on line, I feel on the whole:

angry
tense
worthless

When I get disappointing news, such as doing worse than expected in a test, I feel on the whole:

depressed
worried
annoyed

In this example, people are asked to choose *one* answer. There are three hypothetical subscales (anger, depression and anxiety), and the score on each is the number of times a relevant emotion is chosen. You would need to look at reliabilities for all three subscales. However, if your forced-choice alternatives are just 'rubbish' apart from the one you are interested in, there is obviously no need to look at subscale reliabilities. For example, in tests of knowledge or ability, the alternatives to the correct answer do not form separate subscales, as in the following:

Churchill was:

a World War II General
a British Prime Minister
the place of a famous battle

Pearl Harbour was:
 the place of a famous battle
 a singer
 a jewellery shop

The score here is the number of correct items endorsed, and since there is no rhyme or reason in the alternatives, you do not treat them as subscales.

Improving reliability of the scale

There are five courses of action, and if you are lucky you may not have to follow these through past number one or two. If you are very unlucky, you will have to work your way through all five.

1. You may have high reliability coefficients and might decide to keep your scale (and any subscales) as it stands. No action is needed.
2. If your reliabilities are low, you must look at each item to see how it relates to all the others. This is worth doing even with high reliability if you want to improve the reliability still further. This is done by looking at item-total correlations. A reliability facility can give item-total correlations between each item and the total of the *other* items in the scale. Thus, the total is not contaminated by the contribution of the item in question. This indexes the *cohesiveness* of the scale.

If you have subscales, you should look at correlations between each item from each subscale with the relevant subscale totals in the same way as for the overall scale, simply using the items involved in the subscale.

You reject items with unsatisfactory item-total correlations.

Your reliability facility will probably tell you *the effect on the reliability coefficient of removing any given item*. This information will tell a similar story to the item-total correlation, and you may find it simpler and clearer to go straight for this information as a way of improving your scale. You can overlook the item-total correlations, or glance at them to check that they confirm that you have made the right decisions: items with *low* item-total correlations should be the ones whose removal leads to the *biggest* improvement in the reliability coefficient.

3. If the number of items in your scale (or subscale) is small, alpha is likely to be low even if they are quite strongly associated with each other. Try calculating split-half reliability, which involves the following steps:
 - Divide the items randomly into two groups of equal size
 - Calculate a total for each of these half-scales
 - Calculate the correlation coefficient between these half-scales

Note that a reliability facility will normally do these steps for you without you personally having to go through them.

4. A scale may have poor overall cohesiveness, and this may be because there are subscales (factors) that you had not suspected. Try factor or principal components analysis (see the section on this topic later in this chapter). If the results make sense and produce a factor (or factors) on which items with high loadings relate to the construct(s) you are assessing, then items with high loadings (0.3 or 0.4 and above) are retained.

Note that this method may give you two or more factors, suggesting the existence of two or more subscales. For each subscale, retain and score only those items with high loadings on the relevant factor. Note that if an item has a negative loading, its score will have to be reversed.

Strictly speaking, if you have identified two or more subscales using factor or principal components analysis, you do not need to confirm this by doing a reliability analysis of the subscales.

5. The last and saddest resort is to salvage any items that seem worthwhile, write some more items, and start testing again. However, as discussed, avoid the all-too-common failing of generating a lot of repetitious items.

Descriptive statistics (norms) for the final scale

You define your final scale by listing only those items with satisfactory item-total correlation and obtain Cronbach's alpha (or other reliability coefficient), the number of cases, mean, range and variance of the final scale, and of any subscales.

Some would prefer to divide the scale mean by the number of items in the scale, to give an *item mean*. The advantage of this is that if the number of items in the scale is ever changed at a later date, then comparisons involving the item mean would still make some sense.

If you have decided to use *item weightings*, then answers to each item must be multiplied by the relevant weighting before the mean score on the scale is determined.

The production of *standardised scores* is beyond the scope of this book, and Kline (1999) or Urbina (2014) may be followed if these are desired.

Summary of steps for data entry and reliability

1. Code data where needed.
2. Enter data.
3. Check data.

4. Recode where necessary (for 'reverse-meaning' items).
5. Save data file, making at least one backup copy.
6. Compute the reliability coefficient (usually Cronbach's alpha) for the scale, and for any subscales. If satisfactory, you can go straight to stage 9.
7. If this is unsatisfactory (less than 0.70–0.80) examine item-total correlations (scale cohesiveness).
8. Remove items with unsatisfactory item-total correlations.
9. Produce descriptive statistics (norms) for the final scale, containing only those items with satisfactory item-total correlations: number of cases, mean, range and variance (or standard deviation), and reliability coefficients.

If this fails to produce a scale with a satisfactory reliability coefficient, you could attempt a factor (principal components) analysis, as described later and in [Appendix 5.6](#).

If ever describing any scale with dubious reliability, you should state clearly that the scale does not meet criteria for reliability.

Factor and principal-components analyses

If you are developing a scale for professional use or major research purposes, some psychometricians would regard factor analysis as an important or even essential procedure in the construction of psychological scales and tests. If, however, you have obtained satisfactory reliability or you are not applying the scale in a major way, or both, you could skip this bit. As discussed earlier in this chapter, you might consider this analysis as a way of identifying possible subscales, if your measure has poor internal consistency.

The underlying theory is a specialist aspect of statistics, and there is a good deal of controversy surrounding the applications of factor analysis. [Chapter 1](#) refers to relevant discussions.

For those taking first steps in scale construction, you could consider using factor analysis if you thought there were several factors in your scale, but you were not sure which items were contributing to them. Exploratory factor analysis will tell you the answers to these questions, and so, too will principal components analysis. [Chapter 1](#) gives a brief description of factor and principal components analyses and explains the difference between them, advising you to be clear about which you are doing – they are sometimes confused with each other, and they will give similar results. It is suggested that principal components analysis may give clearer answers than factor analysis. [Appendix 5.6](#) offers brief instructions for carrying out factor or principal components analysis.

If you do a factor (or principal components) analysis, you need to examine which items load heavily on each factor. Common criteria are as follows:

- Normally, loadings of 0.4 and above, or
- Loadings of 0.3 and higher where there are few or no high loadings, for example, and where it would make sense in naming the factors to include items with slightly lower loadings

Name the factor according to the items that load heavily on it. To do this, make a list of high-loading items, and look for a common feature. Note that *negative* loadings indicate that the item is negatively associated with the factor.

You do not have to accept and use all the factors. A rough guide to deciding which factors to accept is as follows:

- The first one to four factors will probably account for quite a lot of variance each, and after that there may be a sudden drop. A likely type of scenario would be for factor 1 to account for about 13%, factor 2 about 8%, and then factors 3 and 4 to account only for a mere 3% or 4% each. Look for such a drop, and use it as a cut-off point in accepting which factors to use. SPSS will normally drop factors that account for insufficient variance.
- Only accept factors accounting for, say, about 8% or more of the variance. Thus, in this example, only consider using factors 1 and 2.
- Only accept factors that make sense in terms of the constructs you are assessing. Of course, a reasonable amount of variance should be accounted for. This means that if a factor accounts for a lot of variance, and it does not make sense or is not of interest to you – then you do not have to use it.

If you do a factor analysis, you may want to consider whether to use *item weightings*. Weighting the items means that an item with high weighting contributes more to the score than an item with low weighting. The main advantage of weighting is that it could give a more sensitive scale, possibly with improved validity. The main disadvantage is that calculation of scale totals and norming data is more complicated and time consuming. Responses to each item have to be multiplied by the item weighting (of course this is normally done by the computer). You could certainly consider using weighted scores if you find the results of your validity analyses are disappointing. There are several ways of developing item weightings: factor loadings can be used as item weightings, or the factor

analysis facility on a statistics package will calculate factor scores (consult your statistics package manual if necessary). If you are thinking of using item weightings, a more advanced manual on test construction should be consulted.

A simpler application of factor analysis, which has some of the advantages of item weighting without the disadvantages, is to *use the factor loadings as criteria* for retaining items in the scale. Decide whether to use a loading of 0.3 or 0.4 as a cut-off, and accept only items with this loading or higher in your scale. This may improve the sensitivity and reliability of your scale, without the extra computational labour of item weightings. However, unless you had an undiscovered subscale structure embedded in your scale, which only factor analysis could reveal, this method is unlikely to give better results than a reliability (item) analysis.

If the results of factor analysis look useful, and you need more information than is provided in this book, you could consult Field (2017) for a more detailed guide on doing factor analysis. Tabachnick and Fidell (2013) present a much more detailed account of factor analysis, including a comparison of relevant statistics packages.

Summary

This chapter describes how to get a final, reliable scale from the pool of items that your participants have responded to. This involved coding your questionnaire, making a database and using a reliability analysis to see how strongly each item, in turn, relates to all the other items in the scale (or subscale). Those that do not relate well are removed, and the result should be a consistent, reliable scale.

Factor and principal components analysis are briefly outlined, as is item weighting.

Appendix 5: Reliability

Notes on statistics packages

The specific instructions for computing given in these appendices are for IBM SPSS for Windows users. Full instructions are given for carrying out the necessary operations to enter and check data, and to compute and improve reliability.

IBM SPSS for Windows is frequently updated. As of this writing, SPSS 26 is the most modern, but this will almost certainly not be the case by the time you read this. Fortunately the changes made with each new edition of SPSS generally consist of the addition of new statistical tests. The general menus and output formats remain constant, except for a change in

formatting in SPSS 7.5 and later versions. This means that these instructions should be appropriate for any later editions of SPSS.

Loewenthal (1996) gives instructions for calculating reliability under circumstances without a suitable statistics package.

Other statistics packages for reliability analysis include

- SYSTAT
- SAS/STAT
- Analyse-it

Appendix 5.1: Types of data, coding, and statistics

You need to be aware what kind of number scale you are using when you code your data. The type of number scale will affect the choice of statistics that are appropriate. Consult the guide that follows (and, if necessary, [Chapter 5](#)), or consult a suitable book (such as Coolican, 2019, de Vaus, 2013, or Gravetter & Wallnau, 2016) or an appropriate person if you feel unsure on any of the points.

- *Interval data:* Age and other numerical data (such as number of children, years of education and the like) can be entered straight onto the computer. It is called interval data because you know the size of the interval between two numbers – the difference between three and two is the same as the difference between two and one. If this sounds like a strange thing to be concerned about, be assured that it *is* important. It is quite likely that some of the data you deal with will not be interval data: other types of data are described later, and in [Table A5.1](#). Interval data are suitable for *parametric* statistical analysis including analysis of variance (ANOVA), Pearson correlation, *t*-test, and multiple regression analysis.
- *Ordered (or ordinal) data:* Here the data may look like numerical data, but there is a difference. With ordinal data you can only be sure that one number is greater than another, but you cannot be sure by how much. Supposing three brands of chocolate are ranked in order of preference, with brand X first (rank = 1), brand Y second (rank = 2) and brand Z third (rank = 3). We know that 1, 2 and 3 are in order of magnitude, but we do not know anything about the *size* of the intervals between them. Brand X may be liked very much, much better than brand Y. However, brand Y is only a bit nicer than brand Z, but this is not reflected in rank-order data. Ordinal data may be used in *non-parametric* statistics, suitable for ordinal data, such

Table A5.1 Different types of number scales: Interval, ordinal and nominal

Type of Scale		Interval	Ordinal	Nominal
	Type of snack	Rating for nutrition (1–7)	Order of preference	Taste (sweet = 1, spicy = 2, bland = 3)
Data from one participant	Chocolate	7	5	1
	Pizza	4	3	2
	Crisps	4	4	2
	Nuts	1	1	3
	Muesli bar	2	2	1
Descriptive statistics	—	Range, means, standard deviations	Range	Frequencies, cross-tabs
Inferential statistics		Parametric: e.g. ANOVA, <i>t</i> -test, Pearson correlation, multiple regression	Non-parametric for ordinal data: e.g. Kruskal–Wallis ANOVA, Spearman correlation	Non-parametric for nominal data, e.g. chi-square, loglinear analysis

Abbreviation: ANOVA, analysis of variance.

as Mann-Whitney U test, Kendall's tau, Kruskal-Wallis one-way ANOVA and Jonkheere's trend test. In some cases it may be useful to treat the data as categorical, particularly as dichotomous (see later), but never treat it as interval data. Taking the most likely scenarios, you can use ordinal data in the non-parametric tests of association listed previously. Ordinal data may also be used in cross-tabs and chi-squares, though you may have to collapse some categories to get a result that can be interpreted. Ordinal data should not be used in tests suitable only for parametric data, such as ANOVA, or most other forms of multivariate analysis.

- *Nominal (or categorical) data*: Here, you do not even know that one number is greater or less than another number. Numbers are just used to indicate *category differences*. In Table A5.1 there are three categories for food taste: mainly sweet, mainly spicy and bland. If you are going to code these, you should represent them by numbers (because computer statistical packages do not like and cannot normally use alphabetical data). You can see that the numbers 1, 2 and 3 are not ordered. It can

be *very* advantageous to have just *two* categories, rather than more than two. This is called *dichotomous* data. Statistics and interpretation become much clearer and simpler, and important statistics are possible with dichotomous data. Notably, statistics packages *can* run correlations and reliability statistics on dichotomous data. The techniques are called, respectively, point biserial correlation, and the Kuder–Richardson KR–20 reliability coefficient. It is also allowable to include variables involving dichotomous data as independent variables in a multiple regression analysis. Unlike categorical data with more than two categories, dichotomous data *are* ordered. The moral of all this is to try and reduce categorical data to dichotomies, where it is sensible to do so. You do not have to do this when you code the data; you can make lots of categories and then collapse them into two when you are computing (see the instructions for recoding in [Appendix 5.3](#)).

- *Rating scales*: What about ratings? Are they interval or ordinal scales? If I rate my liking for chocolate on a 5–point scale, and brand X gets a 5, brand Y gets a 4, and brand Z gets a 3, is the difference between the 5 and the 4 the same as the difference between the 4 and the 3? Generally, *rating scales may be treated as interval scales*, unless you suspect the contrary. Likert (1932) suggested this when he developed his best-selling method of scale construction involving rating. Likert’s (cunning) suggestion was that the intervals in such rating scales should be *equal-appearing*.

Appendix 5.2: An example of coding

Here is a simple coding scheme for the beginning of the questionnaire shown in [Chapter 2](#); lots of others could have been devised. The answers to the ‘test’ part of the questionnaire (the questions on prayer, in the example) do not need any special coding. They can be entered onto the computer database just as they were given. Most of the answers to the first part of the questionnaire, which deals with background, social-demographic, information need to be coded before they can be entered onto the computer. Some coding is provided in which some variables like marital status, details of children, occupation and religious membership have been reduced to simple 0/1 dichotomies, to make statistics more straightforward, but you may wish to use a more elaborate categorisation and reduce it later. If you do use a more elaborate categorisation, look in [Table A5.3](#) for instructions on how to recode.

[Table A5.2](#) shows the answers, and their coded versions, of a hypothetical participant on this first part of the questionnaire.

Table A5.2 Sample questionnaire

QUESTIONNAIRE ON PRAYER

We are studying people's views on the uses of prayer.

Your answers to the following questions would be very helpful. Your answers will be confidential and anonymous, identified only by a code number. You need not answer any questions that you would prefer to leave unanswered.

Thank you.

Date _____ (Do not code)

Your age _____ (Enter as given)

Male/Female _____ (1 = male, 0 = female)

Current marital status (Circle one): (Code 1 if currently in stable relationship: married, engaged or cohabiting, 0 otherwise)

Married (1)

Engaged (1)

Single (0)

Cohabiting (1)

Divorced (0)

Widowed (0)

Separated (0)

Number of children, if any _____ (Enter as given)

Their ages _____

(Code 1 if any under 18 years, 0 otherwise)

Your occupation _____

(Code 1 if earning, 0 otherwise)

If married, your spouse's occupation _____

(Code 1 if earning, 0 otherwise)

Do you regard yourself as spiritual? _____ (Yes = 1, No = 0)

Do you belong to any church, mosque or synagogue? _____ (Yes = 1, No = 0)

How often do you attend? (Circle one)

Daily

Weekly

Monthly

Occasionally

Never

(daily = 4, weekly = 3, monthly = 2, occasionally = 1, never = 0)

How often do you pray? (Circle one)

Daily

Weekly

Monthly

Occasionally

Never

(daily = 4, weekly = 3, monthly = 2, occasionally = 1, never = 0)

(Continued)

Table A5.2 (Continued) Sample Questionnaire

How often do you study religious texts? (Circle one)

- Daily
- Weekly
- Monthly
- Occasionally
- Never

(daily = 4, weekly = 3, monthly = 2, occasionally = 1, never = 0)

Table A5.3 Sample coding

ID 1

Your age 22

Male/Female male 1

Current marital status (underline one):

- Married
- Engaged
- Single
- Cohabiting
- Divorced
- Widowed
- Separated 0

Number of children, if any 0

Their ages _____ 0

Your occupation student 0

If married, your spouse's occupation _____

Do you regard yourself as spiritual? Yes 1

Do you belong to any church, mosque or synagogue? No 0

How often do you attend? (underline one)

- Daily
- Weekly
- Monthly
- Occasionally
- Never

1

How often do you pray? (underline one)

- Daily
- Weekly
- Monthly
- Occasionally
- Never

0

(Continued)

Table A5.3 (Continued) Sample coding

How often do you study religious texts? (underline one)

Daily
 Weekly
 Monthly
 Occasionally
 Never
 1

Appendix 5.3: Making and checking a database in SPSS for Windows

Entering data

Notes:

1. *Important:* Make sure that you save your data *frequently*, and back up. SPSS does not automatically save as you go along, and if there is any kind of glitch (e.g. you exit by mistake, you lose power) everything you have done will be lost, unless it was saved. A good idea is to back up as you go along.
2. You do not have to enter data into SPSS if you are working on a device without SPSS. Data may be entered into another programme such as Excel, and then imported into SPSS.

To enter data in SPSS for Windows, follow these steps:

- *Enter SPSS:* Click on Start/Programs/SPSS for Windows. An empty data entry spreadsheet should appear, with the highlight in the top left-hand cell. The highlight can be moved to this position by pressing Ctrl + Home simultaneously.
- *Name variables:* This can be done at any time, but it will help to keep things clear if you do this before you start entering data. Each column will contain data for each variable: move the highlight to *any* cell in the column, select Data/Define variable, click, type the variable name in the top (highlighted cell of the) dialogue box that appears, and press 'Enter'. You can change the variable names at any time, using this same method.
- *Type in the data:* What you type will appear in a small dialogue box above the spreadsheet itself, and just under the menu bars. Any errors made in this dialogue box can be easily corrected using the method you

prefer – the ‘Delete’ keys for small errors, and highlight and re-type for longer pieces of material. Press ‘Enter’ (bent arrow key) to put each piece of data in the highlighted cell. It is usually easiest to type in the data moving horizontally across the spreadsheet. Enter the participant’s ID number in the left-hand cell of each horizontal line, and then use the right-arrow key to move to each new cell. Move to a new line for each new case (participant). It is important to let each *row* contain the data for each participant, and each *column* to contain the data for each variable. This is important because that is the way SPSS will read the information when it does any statistics.

- *Missing data:* Just leave the relevant cell(s) empty. A single full-stop point (.) should appear in any cell with missing data.
- *Correcting data:* Move the highlight to the cell needing correction, type in the correct entry, and press the ‘Enter’ (bent arrow) key.
- *To copy or move sections of data:* Use click/drag to highlight the section to be moved. Select Edit, and click on Copy (or Ctrl C) or Cut (or Ctrl X) (as appropriate). Move the cursor to the cell where the section is to be pasted, and click Select Edit/Paste (or Ctrl V).
- *Value labels:* If you wish, you can use the ‘Change Settings’ area in the dialogue box to define the variable: give a full description of the variable, and the value labels. Click on ‘Change setting’, then ‘Labels’. For example, under ‘Variable label’ you could describe the variable ‘married’ as ‘whether now-married or not’. Under ‘Value labels’, you can record that now-married = 1, not-now-married = 0, as follows:
 - Enter ‘1’ under ‘value’.
 - Type ‘now-married’ by ‘value label’.
 - Click on ‘Add’.
 - Enter ‘0’ under ‘value’.
 - Type ‘not-now-married’ by ‘value label’.
 - Click on ‘Add’.
 - Click on ‘Continue’.
 - Click on ‘OK’.

It is useful to record these value labels if you are working on a database that is going to be used and re-used a few times. Sometimes the value labels will appear on the database, rather than the values themselves. Thus you might see words like ‘Male’ and ‘Female’ in the data cells, rather than numbers. SPSS will still perform statistics on the numbers themselves. You may want to change the entries from the value labels to the values themselves (or vice versa). Click on the little icon that looks like a luggage label. It should be the second from the end, on the right.

- *Recoding*: There are two reasons why you might need to do some recoding. First, if you have any *reverse-meaning* items, for example ‘I hate chocolate’ on a scale assessing liking for chocolate, or ‘Prayer is a waste of time’ on a scale measuring the perceived usefulness of prayer. There are good reasons for including some items like this, as described in [Chapter 2](#). Second, if you have a categorical variable with lots of values, and for reasons described in [Chapter 5](#) and in [Appendix 5.1](#), you decide it would be helpful to *recode it as a dichotomous variable*. To *recode reverse-meaning items*, use the *recode into same variables* option:
 - Select Transform/Recode/Into Same Variables.
 - Select the variables to be recoded (e.g. item 4, item 6) and move them into the ‘Numeric Variables’ box on the right, by clicking on the right-pointing arrowhead.
 - Click on Old and New Values.
 - Suppose your participants had been asked to rate how much they agree with each statement on a scale from 1 to 5, where 5 means strong agreement, and 1 means strong disagreement. For the reverse-meaning items, you will need to change all the 5s into 1s, all the 4s into 2s, all the 2s into 4s, and all the 1s into 5s. Three, the mid-point, stays as it is. So, in the Old and New Values screen:
 - Enter 5 into the Old Values box.
 - Enter 1 into the New Values Box.
 - Click on Add. 5→1 should appear in the box headed Old→New.
 - Repeat these three steps for 4→2, 2→4 and 1→5. (Remember, it is very important to enter *all* the reverse-scoring for the variables concerned on the one occasion, and remember that changing ‘5’ to ‘1’, or ‘2’ to ‘−2’ does just that (for the variables you have selected). It does *not* change ‘1’ to ‘5’, or ‘−2’ to ‘2’).
 - Click on Continue.
 - Click on OK.

A similar procedure is followed for *recoding categorical variables with many values into dichotomous variables*, except that here it is safer to *recode into a new (different) variable*, so that you retain the original information for possible future use. So to recode into a dichotomy,

- Select Transform/Recode/Into Different Variables.
- Select the variables to be recoded (e.g. RelGroup) and move them into the ‘Numeric Variables→Output Variables’ box on the right, by clicking on the right-pointing arrowhead.

- Now you must give a name to the new (output) variable(s). Type a suitable name into the Output Variable Name box. For example, RelGroup was coded as 0 if no campus group was joined, and 1 or 2 according to *which* campus religious group was joined. We now want to recode to a variable showing *whether* a campus religious group was joined at all. So the new variable could be called 'joined', and all the religious groups concerned (only two in this case) can be recoded as '1', to represent the fact that the individuals concerned joined a religious group. Under 'Label' you can enter a description of the new variable, such as 'Whether joined campus religious group'.
- Click on Old and New Values.
- If there is only one value to be changed, it can be entered individually, into the top box under 'Old Value'. However, it may be quicker to use one of the 'Range' options if there is a longish string of consecutive numbers to be changed. In the example under consideration, you just need to change '2' into '1', so just enter 2 into the top box.
- Enter '1' in the New Value box.
- Because you are recoding into a new variable, you also have to tell SPSS to copy all the old values that are remaining unchanged. So enter 0 into the Old Value box, select Copy Old Value/Add. Enter other unchanged values in the same way. If there are any missing values in the variable, select System Missing under Old Value, and Copy Old Value followed by Add, under New Value.
- Click on Continue.
- Click on OK.
- *Save data:* Select File/Save As. A Save As Data File dialogue box will appear. Type in the filename, e.g. religion.sav. If you want to save on a floppy disc, click on [-A-] in the field headed Save in. Otherwise select an appropriate directory. In the bottom of the dialogue box are options for the format to be used. Normally the default will be SPSS *.sav. This saves the data in SPSS for Windows format. If for any reason you wish to select another format, do so. Select OK and click. Note that you must use the filename extension .sav when naming and saving data files in SPSS for Windows.
- *To retrieve data:* Select File/Open Data. A dialogue box will appear. Either type in the filename (e.g. a:prayer.sav), or click on the little down-pointing arrow to the right of the Look In box. This will display a choice of drives or directories. Double-click on [-A-] (or the appropriate drive, and then if necessary the appropriate directory) in the Look In box. All files on your floppy disc ending in .sav will be

listed. Select the file to be retrieved and click. Check that the filename in the Name box is indeed the file you want, select OK and click. The data file may now be edited (or analysed).

- You can also choose to import data from Excel, SYSTAT or another source by selecting File/Import Data followed by your selection from the list of possible types of data.

Checking the database

Note: At any time, you can click on *Edit* at the top of the SPSS screen, and choose *Options*. This will allow you to choose how your results are displayed – for example the type of table your results are displayed in, and whether the variables appear in alphabetical order, or their order in the database. It is not urgent or even essential to change the display type, you can just keep going with the default option. But you might want to experiment with the display options, or use the relevant section of Pallant's manual.

Before starting statistical analysis, you should carry out the following checks for accuracy. We have described three checks, starting with the quickest and easiest, but if you carry out the second check, there is no real need for the first, and if you carry out the third check, there is no real need for the first two. If you have labelled your values, and these labels are showing up, rather than the numbers you need to check, place the cursor anywhere in the column concerned, and click on the little icon that looks like a luggage label, near the end on the right.

- First, a *preliminary scan* ('eyeballing'). This may only enable you to detect errors that would be detected anyway, in the second and third checks. But it is always a good idea to 'eyeball' your data. It is embarrassing to find out – when it is too late to do anything about it – that there are errors that could have been detected if you had just glanced over your material. So, glance over your database. It is usually easier to do this if the database has been printed out. *To print*, with the database on the screen, either click the Print icon, or select File, then Print. Look quickly over the database for things that should not be there, for example ages like 3025, or values like 53, when you should have typed '30' and '25', '5' and '3' in separate cells. Note any errors, and check your original data for the correct entry. Then go back to the computer database and correct the entries. *To correct errors* on the computer database, move the cursor to the cell concerned, and type in the correct entry. This will not appear in the cell itself, but in the dialogue box immediately under the tool bars. When you press 'Enter', the corrected entry should then

appear in the highlighted cell. *Save* the new, corrected version of the database, including at least one backup copy.

- The second check of the database involves asking for frequencies of each variable and scanning the results. This may show entries that have no business being there but which do not stand out visually, so you might not have spotted them on the first check. Examples would be an age like 16 (when all your participants were over 18), or a rating of 6 on a scale ranging from 1 to 5. To obtain frequencies in SPSS:
 - With the database on the screen, select *Analyze*.
 - Select *Descriptives*, then *Frequencies*.
 - Move *all* variables into the *Variables* box, then click on *OK*.

Scan for unlikely or impossible values, such as ages like 2 or 99 when all participants were aged 18–60, or values like 0, 6, 25 or 3.341 on a rating scale involving only the whole numbers from 1 to 5. Locate all errors on the database – this is easier to do on a paper database. You should then look at your original records, note the correct entry, and when all corrections have been made on the paper database, amend the computer database as described earlier, and save. Some sample outputs are shown in [Table A5.4](#). You might like to look at them to see if you can spot possible errors.

The ‘answers’ to the error-spotting exercise in [Table A5.4](#) are as follows: age has two values of 2, and one of 451. These would need to be located and corrected. Religious affiliation has one value of 12, and Q1 has one value of 0 and one of 11.

- The third method of checking the database will uncover any errors that cannot be detected using the first and second methods – for example entering 2 when 3 should have been entered, for a rating on a 1–5 scale. This third method is best done working in a pair. One member of the pair looks at the original data, and the other at a printout of the database. One person reads out their data, and the other calls out when they spot a mis-match. Correct all errors on the paper print-out, then correct the electronic version, save it and back it up. As stated, if you are able to carry out this type of check (and ideally, you should), then the first two checks are not strictly necessary.

Replacing missing data

Before starting reliability analysis, you might need to give some thought to the question of missing data. Some people may not have answered all questions, and missing data can be a nuisance because SPSS normally

Table A5.4 Locating errors from SPSS frequency data: Age, religiously affiliated and prayer

		Frequency	Percentage (%)	Valid Percentage (%)	Cumulative Percentage (%)
Age					
Valid	2.00	2	2.0	2.0	2.0
	18.00	4	4.0	4.0	6.0
	19.00	4	4.0	4.0	10.0
	20.00	2	2.0	2.0	12.0
	21.00	2	2.0	2.0	14.0
	22.00	6	6.0	6.0	20.0
	23.00	10	10.0	10.0	30.0
	27.00	2	2.0	2.0	32.0
	28.00	2	2.0	2.0	34.0
	29.00	10	10.0	10.0	44.0
	31.00	1	1.0	1.0	45.0
	33.00	3	3.0	3.0	48.0
	34.00	15	15.0	15.0	63.0
	35.00	10	10.0	10.0	73.0
	36.00	1	1.0	1.0	74.0
	37.00	1	1.0	1.0	75.0
	38.00	1	1.0	1.0	76.0
	39.00	1	1.0	1.0	77.0
	40.00	1	1.0	1.0	78.0
	42.00	2	2.0	2.0	80.0
	45.00	6	6.0	6.0	86.0
	46.00	2	2.0	2.0	88.0
	53.00	1	1.0	1.0	89.0
	55.00	1	1.0	1.0	90.0
	56.00	7	7.0	7.0	97.0
	57.00	1	1.0	1.0	98.0
	59.00	1	1.0	1.0	99.0
	451.00	1	1.0	1.0	100.0
	Total	100	100.0	100.0	
Total	100	100.0			
RelAffil (whether religiously affiliated = 1, or not = 0)					
Valid	0.00	40	40.0	40.0	40.0
	1.00	59	59.0	59.0	99.0
	12.00	1	1.0	1.0	100.0
	Total	100	100.0	100.0	
Total	100	100.0			

(Continued)

Table A5.4 (Continued) Locating errors from SPSS frequency data: Age, religiously affiliated and prayer

		Frequency	Percentage (%)	Valid Percentage (%)	Cumulative Percentage (%)
Q1 (ratings of 1–5)					
Valid	0.00	1	1.0	1.0	1.0
	2.00	19	19.0	19.2	20.2
	3.00	30	30.0	30.3	50.5
	4.00	28	28.0	28.3	78.8
	5.00	20	20.0	20.2	99.0
	11.00	1	1.0	1.0	100.0
	Total	99	99.0	100.0	
Missing System		1	1.0		
Total	100	100.0			

excludes *all* cases with even *one* piece of missing data from just *one* of the variables included in any analysis. Here are some suggestions to consider:

- If there are very few pieces of missing data, you may ignore the ‘problem’. Some information will obviously not be used in all analyses, but if this involves just one or two cases this may be a trivial loss.
- If you find that much of the missing information is from *one* (or two, or a few) variable only, you may be able to exclude that variable from the analysis. For example if only 60% of participants answered one of the items on a scale, but all other items had been answered by nearly everyone, it would be advisable to exclude that item from the scale. There is certainly something suspicious about an item that nearly half the participants would not or could not answer, and you might want to look at it to see if you can work out why. Almost certainly it should not be included in further work. Or there might be a variable like age or religion which some people have not given any information. You can still use the information that has been given in some analyses, for example in offering descriptive statistics for the sample. Remember to indicate how many (or what proportion) people did not give this piece of information.
- In some circumstances, it may be acceptable to replace missing values, with made-up values. This should only be done if it would make sense

to replace missing values in this way. The most common replacement strategy is to replace missing values with the mean for that variable. This would *not* make sense for variables like age, or religious affiliation. It might make sense if you have, say, 30 participants who have each left out one answer, each to a different item on a scale. If you have only a hundred participants, you will lose a large amount of data from your analyses, and you might try replacing missing values with the mean on the variable concerned. To do this in SPSS,

- Select Transform.
- Select Replace Missing Values.
- Select the variables that have missing values that you want to replace. Move them, by clicking on the right-pointing arrowhead, into the New Variables box.
- The default option is to replace missing values with the series mean (mean for that variable).
- Click OK.
- SPSS will put a new variable in your database, identifying it by a suffix of the form ‘_1’.
- You can try analyses using these new variables, instead of the old ones with missing values. If you do this, remember to report that you replaced missing values with series means, identifying the variables for which you did this.

It is not however very good practice to use items with replaced missing values in a reliability analysis. It is better to use the original data. But there may be some analyses that you carry out later – in which you might want to compare total scores on your scale with other factors – in which the replacement of missing values is defensible.

Note regarding normalcy and homogeneity: If your statistical analyses involve interval data, you may feel you need to worry about whether your data are normally distributed and the variance is homogenous – or somebody else may suggest this. Provided your sample is reasonably large, normalcy of data is not generally a concern, and homogeneity of variance need not be of concern unless you have some obvious outlying data (values in your data that are way different from most). If you are concerned about normalcy or homogeneity of variance, Field (2017) offers a clear explanation of what is involved in exploring these issues, while Pallant (2016) offers concise instructions for computing. Both authors however reassure that the statistical tests normally used are robust, and still produce meaningful results if assumptions of normalcy and homogeneity of variance are not totally correct.

Appendix 5.4: Reliability analysis in SPSS for Windows

This section describes the steps needed to compute Cronbach’s alpha (a coefficient of internal consistency) using SPSS for Windows.

- Retrieve your data if necessary. Select File/Open/Data. Type the filename in the dialogue box (e.g. a: prayer.sav), or select the appropriate drive, directory and file, and click OK.
- Select Analyze/Scale/Reliability Analysis.
- Enter all items in your scale (or subscale) into the [items] box. Select from the box on the left, click, and click on the right-pointing arrow. This will enter the selected items into the [items] box.
- Select required model. Normally, the default option, Alpha, should be used.
- Select OK.

A sample output is shown in [Table A5.5](#).

[Table A5.5](#) is not difficult to interpret. It first tells you how many cases were included in the analysis, and how many items. Check that you are happy with the number of cases included in the analysis. If there are many missing, check for missing data and consider the suggestions made earlier, for dealing with missing data. Then check that the right number of items are there, i.e. the number you intended to include. If that is wrong, you need to go back to the items box in the reliability analysis (just select Analyze/

Table A5.5 SPSS reliability analysis output

<i>Reliability</i>			
<i>Scale: Prayer</i>			
<i>Case Processing Summary</i>			
		<i>N</i>	<i>%</i>
Cases	Valid	94	94
	Excluded ^a	6	6
	Total	100	100
<i>Reliability Statistics</i>			
<i>Cronbach's Alpha</i>		<i>Number of Items</i>	
0.6111		8	

a Listwise deletion based on all variables in procedure.

Scale/Reliability Analysis again). Check and rectify the contents. Finally, the statistic you have been working so hard for, Alpha. In [Table A5.5](#) it is not particularly high – as discussed in [Chapter 5](#), 0.8 and above is good, and 0.7 and above is satisfactory. A value of 0.6111 is not very satisfactory, but it is not hopeless. If you have a result like this, you could turn to the next section and follow the steps described to improve alpha.

This section has not dealt with reliability indices other than Cronbach's alpha. If you have reason to believe that split-half reliability, Guttman coefficients or other options would be in order, these can be computed by altering the model in the Model box in the lower left of the Reliability Analysis window.

Appendix 5.5: Improving reliability

Reliability is improved by looking at the item-total correlations of each item and rejecting those with low item-total correlations. Alternatively, SPSS will tell you the effect on alpha, of rejecting any given item from the scale. In practice, it makes no difference whether you are guided by the item-total correlation information, or by the information about the effect on alpha. To obtain these pieces of information,

- Select Analyze/Scale/Reliability.
- Move the items in the scale into the Items box, as described earlier (if they are not already there).
- Click on Statistics, in the upper right corner of the Reliability Analysis window.
- A new window will appear. Select Scale if item deleted – click on the adjacent box and a tick (check) should appear.
- Continue.
- OK.

A sample output is shown in [Table A5.6](#).

Look at the column headed 'corrected item-total correlation'. This tells you how well the item correlates with the others. If you look in the output presented, you see that the lowest item-total correlations are for items 6 and 8. Item 8 should certainly go, and we might consider throwing out 6. Correlations of the order of 0.15 or less could definitely mean that item should be deleted, unless you are desperately short of higher correlations. The column headed 'alpha if item deleted' tells you what the alpha coefficient would be if that item were gotten rid of. As you can see, the effect of throwing out the less cohesive items is to raise alpha, and vice versa.

Table A5.6 SPSS reliability analysis output: Item-total statistics

Item-Total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Alpha if Item Deleted
Q1	21.5000	30.3535	0.4611	0.5521
Q2	21.8000	28.1414	0.5735	0.5168
Q3	22.1000	31.0000	0.6576	0.5436
Q4	21.9000	31.4444	0.4047	0.5674
Q5	22.0000	29.1515	0.5285	0.5328
Q6	22.2000	33.4545	0.1341	0.6200
Q7	22.0000	29.6970	0.6268	0.5286
Q8	21.5000	23.6869	0.1026	0.8109

Reliability Statistics	
Cronbach's Alpha	Number of Items
0.6111	8

The first two columns tell you what would happen to the scale mean and variance if each item were deleted.

You can ‘throw out’ items with unsatisfactory item-total correlations, and finish your computing as follows:

- Analyze/Scale/Reliability analysis.
- If all items are in the Items box (they will be if you are doing all the computing in one session), select the items you wish to discard (e.g. Q8, Q6), click on the left-pointing arrowhead.
- If no items are in the Items box, select those you wish to include (e.g. Q1–5, Q7), and click on the right-pointing arrowhead.
- OK.

A sample output is shown in [Table A5.7](#).

If alpha is now satisfactory, you can ignore the rest of the output. [Table A5.7](#) shows an alpha well over 0.8, which is satisfactory. If alpha were still low, you should look at the ‘Alpha if item deleted’ to see if there is scope for any further improvement. If any of the alphas in this column are higher than the alpha at the bottom of the output, then you could try removing the relevant item(s). However, if the alphas in the left-hand column are all lower than the alpha at the bottom of the output – this is the case in

Table A5.7 SPSS reliability analysis output after removing selected items

<i>Item-Total Statistics</i>				
	<i>Scale Mean if Item Deleted</i>	<i>Scale Variance if Item Deleted</i>	<i>Corrected Item-Total Correlation</i>	<i>Cronbach's Alpha if Item Deleted</i>
Q1	15.2000	14.9091	0.6604	0.8486
Q2	15.5000	13.5859	0.7248	0.8383
Q3	15.8000	16.1212	0.8153	0.8349
Q4	15.6000	15.3939	0.6625	0.8481
Q5	15.7000	15.3636	0.5384	0.8728
Q7	15.7000	15.3636	0.7167	0.8399

<i>Reliability Statistics</i>	
<i>Cronbach's Alpha</i>	<i>Number of Items</i>
0.8693	6

Table A5.7 – then there are no further improvements possible, at least not in terms of removing items and recomputing alpha. Some suggestions are made in [Chapter 5](#) for procedures to be followed if reliability coefficients are unsatisfactory.

If alpha is satisfactory: compute statistics for the scale:

- Select Analyze/Scale/reliability Analysis.
- With the final selection of ‘good’ items in the Items box, click on Statistics (box in upper right of screen).
- You can remove the tick/check from Scale if item deleted.
- Tick/check Descriptives for: Scale.
- Continue.
- OK.

A sample output is shown in [Table A5.8](#).

These statistics tell you the scale mean, variance, standard deviation and number of items (variables) in the scale. This information should be given in any description of the scale.

If you wish, you can ask for *item means*. Select Analyze/Scale/Reliability Analysis/Statistics/Descriptives for Item. The advantage of this is that if you (or anyone else using your scale) vary the number of items in the scale, comparisons with individual item means would still make some sense.

Table A5.8 Scale statistics: SPSS output after removing selected items (extract)

Mean	Variance	Standard Deviation	Number of Items
18.700	21.2222	4.6068	6

Appendix 5.6: Factor and principal components analyses in SPSS for Windows

This part of [Appendix 5](#) gives a very brief introduction to the use of factor analysis (FA) and principal components analysis (PCA). This might be considered if you have failed to obtain a satisfactory coefficient of reliability, and you are wondering if there might be two or more subscales embedded in your measure. FA or PCA could identify such subscales. They would emerge as factors (or components) accounting for a reasonable proportion of variance, and on which several of your items could have high loadings.

Factor analysis is a method of examining associations between associations. It finds a small number of underlying dimensions from a larger number of variables, by consolidating the variance. In the case of a test or scale, it will enable you to group items that seem to be assessing the same ‘factor’, by examining the ‘loading’ of all items on each factor and selecting those with high loadings.

Factor analysis proceeds by first extracting factors, and then if more than one factor emerges, a rotation is carried out to give a clearer picture. PCA is not actually FA (it uses regression analysis), but it usually gives similar answers and is often used instead of FA. Our illustrations follow this option. We use a popular method of rotation called *varimax*. Note that you are supposed to have at least three times as many participants as variables to get a meaningful result from FA or PCA, so our little hypothetical example, with 100 participants and eight variables (items), does qualify.

- Select Analyze/Dimension Reduction/Factor. A Factor Analysis dialogue box will appear.
- Enter all items in your scale into the [variables] box. Select from the box on the left, click, and click on the right-pointing arrow. This will enter the selected items into the [variables] box.
- A PCA will be carried out by default. If for any reason you prefer another option, select this by clicking on Extraction. Tabachnik and Fidell (2013) or some other authoritative source would guide on the selection of another extraction method. In reporting this analysis, be clear that it was PCA and not truly FA.

- It is advisable to choose a rotation method, and varimax is the most commonly used method: click on Rotation/Varimax. If you wish for more discussion on rotation methods, consult Tabachnik and Fidell, or other suitable authority.
- Continue.
- OK.

Part of a sample output is shown in [Table A5.8](#). There will be more in the output, but this can usually be ignored. Just concentrate on the bits headed Total Variance Explained and Component Matrix.

The *eigenvalue* of a factor or principal component indicates how much variance (in the original variables) is accounted for. The convention is to discount factors or principal components with eigenvalues of less than one. This is what has happened in the first part of [Table A5.9](#). Eight components have been extracted, but only three have eigenvalues greater than one, so the rest of the output concentrates on those. The upper part of [Table A5.9](#) contains a repetition of information about the first (three) components, with further information about what happened to the variance explained after rotation. [Table A5.9](#) tells us that the first factor (component) accounts for over 40% of the variance, and the first three factors between them account for nearly 80% of the variance. This means that these factors are important.

If there is more than one factor (with an eigenvalue greater than one), then the factors should be rotated, and the component factor matrix reflects the result of this rotation. This matrix tells us how the different items on the questionnaire load on each factor (component). The loading is the correlation between that item and the factor. We can use this information to label the components (see later).

There will be more to the output than described, but hopefully enough is indicated of the important features of FA and PCA.

Interpreting factor loadings and naming factors: The experts on FA may turn a bit coy when it comes to talking about this bit. The awful truth is that after a massive amount of number crunching (by SPSS), your job is to

- Look at which items have high loadings (above about 0.4) on each factor.
- Decide what they have in common.
- Dream up a plausible name for each factor in the light of this.
- Drop any factors that account for small amounts of variance. (What constitutes ‘small’ can vary somewhat, but this would normally be below about 8%–10% of variance.)
- The outcome of this stage is generally what actually gets remembered, by you and by anyone interested in your research; there is no direct computing or number crunching, just your intuition.

Table A5.9 SPSS factor and principal components analysis: Part of the output

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	Variance (%)	Cumulative (%)	Total	Variance (%)	Cumulative (%)	Total	Variance (%)	Cumulative (%)
1	3.813	47.664	47.664	3.813	47.664	47.664	3.373	42.160	42.160
2	1.415	17.683	65.347	1.415	17.683	65.347	1.590	19.877	62.037
3	1.156	14.456	79.803	1.156	14.456	79.803	1.421	17.765	79.803
4	0.877	10.956	90.759						
5	0.474	5.927	96.686						
6	0.159	1.993	98.679						
7	9.014E-02 ^a	1.127	99.806						
8	1.551E-02 ^a	0.194	100.000						

Component Matrix

	Component		
	1	2	3
Q1	0.787	0.177	0.412
Q2	0.811	0.100	−0.141
Q3	0.889	−8.214E-02 ^a 9.754E-02 ^a	
Q4	0.749	0.623	4.245E-02 ^a
Q5	0.678	−0.278	0.508
Q6	0.164	0.838	0.321
Q7	0.829	0.433	−0.187
Q8	0.104	0.100	0.748

Note: Extraction Method: Principal Component Analysis. Three components extracted.

a E-02 indicates overflow of numbers; because the value is so small, this can be ignored.

Look at the items in the scale in relation to the loadings in the component (factor) matrix. For example, looking at the items in the hypothetical prayer scale, in relation to [Table A5.9](#):

1. It is important to pray when you need help.
2. There are better routes to understanding life's mysteries than praying.

3. Praying gives comfort.
 4. Regular contemplative prayer is important.
 5. It is foolish to believe that prayers get answered.
 6. People who get inspired by prayer are kidding themselves.
 7. Prayer puts things in perspective.
 8. Prayer is a waste of time.
- All items except 6 and 8 loaded heavily on the first factor (component). We might suggest that this factor reflects general favourability towards prayer.
 - Items 4, 6 and 7 loaded on the second factor. This factor might reflect a belief in the positive cognitive effects of prayer, particularly contemplative prayer. So we might label this second factor belief in the value of contemplative prayer.
 - Items 1, 5 and 8 loaded heavily on the third factor. This factor looks as if it reflects belief in the instrumental uses of prayer.

You might run separate reliability analyses on any subscales identified in this way, but this is not strictly necessary. FA or PCA is normally considered sufficient evidence of the structure of scales and subscales.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>