

# Quantitative research & analysis

## Selecting appropriate design and validated measures

- Begin with a literature review to find out which research designs are appropriate and to select valid, reliable measures. Knowing which designs and outcome measures are gold standard and using these where possible will mean that your data is comparable with current literature
- If there doesn't appear to be a gold standard, it can be helpful to ask practitioners and researchers which outcome measures they currently use
- When obtaining questionnaires to use in research, ensure you have the scale authors permission, do not alter the questionnaire without the authors permission, and ensure you are aware of how the scale should be scored according to the scale authors instructions

## Entering data into a database

- Write an ID code on the front sheet of the questionnaire pack and enter this number into the first column of your database. This way if there is ever doubt the questionnaire data was entered or scored correctly, you can go back and check
- Enter the data as it appears in the questionnaire pack with one participant per row, and one question per column
- Try to code any non-numeric data e.g. gender as numeric e.g. 1=male 2=female as more analysis can be done with numeric data. Keep a record of your coding within the database. In SPSS this can be done under 'values'.
- If you have missing data, leave the space blank or code missing data as an impossible score e.g. 999. SPSS has a function for this.

## Checking the data

- Before any scoring or analysis can occur, check the data is free of data-entry errors. This can be done quickly by looking at the frequency data per item. In SPSS this is under Analyse, Descriptives, Frequency. For example, if your scale is scored 1-5 then scores should only fall within this range and if a value of 55 emerged, you would know this was due to an error in data entry.

## Calculating questionnaire totals

- Using the scoring guide provided with the questionnaire, calculate the scale total scores and subscale scores.
- Some items on a questionnaire may be reverse-coded e.g. if all items on a scale ask about negative mood but there is an item on positive mood, it is likely that this positive item will need to be recoded to be included in the total score.
- Calculations can be performed quickly and more accurately by using syntax. There is a calculator for performing syntax in SPSS under Transform and then either Compute or Recode. You can copy and paste your calculations to a syntax file and save them there for later reference and later use.

- It is preferable to use a syntax file rather than calculating the total scores by hand as this can easily introduce errors in scoring. It is also important to keep a copy of your syntax file so if your data don't look right, you can refer back to your scoring and make sure it is correct.
- Syntax files are also useful for when you add more participants to a study which you have already analysed data for as you can re-run the syntax to calculate total scores for any additional participants. Just open the syntax file and click Run.

### **Cleaning the data**

- Explore the distribution of your data with histograms and/or measures of skewness and kurtosis. Does your data fall within normal ranges, do you have floor or ceiling effects?
- Sometimes data can be distorted by outliers. Use scatter-plots comparing one variable with another and see if there are any participants who are far away from the line of best fit. Another option is to create standard scores or Z-scores for each variable and see if any participant have especially large z-scores compared with others.
- Remove outliers but make sure you have sound statistical or theoretical reasons for doing so.
- If some variables remain non-normally distributed, consider log-transforming them for further analysis.

### **Analysing data**

- The type of analysis you choose will depend on your research design, the type of data collected e.g. categorical or continuous the normality of your data, how many participants you have, and what research question you would like to answer. Your analysis should test your primary hypothesis.
- If you have non-normally distributed data, small sample size or binary/categorical data it might be best to use non-parametric statistical tests which compare ranked data as opposed to mean scores.
- If however, you have good quality continuous data and want to compare means, then parametric analysis may be appropriate e.g. t tests, ANOVAs.
- If you want to assess the relationship between continuous variables then correlational analyses may be appropriate.
- If you want to look at causality then regression models may be appropriate.

### **Getting advice**

- It is advisable to explore with a literature review which types of analysis are most commonly used by other researchers conducting similar research. Once again this can allow your results to be compared with existing literature.
- It is also helpful to consult a statistician for advice on your analysis plan, the earlier this is done the better i.e. before applying for funding or commencing with ethics if the study is not funded.