Pearson *r* analysis. When you do a Spearman's rho, you preferably use the Greek symbol ρ (or write out rho), and when you do a Kendall's tau, you preferably use the Greek symbol τ (or write out tau).

Apart from the *r*-value and the *p*-value, you will see that there is another value mentioned in between brackets behind the *r*. This number refers to the degrees of freedom. In a correlation analysis, you can simply calculate the degrees of freedom by taking the number of participants minus 2. The lower and upper limits of the 95% Confidence Interval can be taken directly from the output.

Whenever you are reporting on statistics in a research paper, it is not very common to include entire JASP outputs in your results section apart from the scatterplot. The results from the table are usually reported in the text, as in the above sentence. Conventionally, charts are included for significant results only.

Note that statistical notations have to be in italic. When you add notes in your JASP report, there is the option of making text italic.

Part B: Social class

A sociolinguistic researcher wants to find out if there is an association between the use of 'haven't got' versus 'don't have' and social class (two levels) (we saw this data already in Chapter 5.2.3 in Part 1). Social class is determined by asking people what they do for a living. The use of 'haven't got' or 'don't have' is determined by asking subjects to rephrase sentences (for instance: Jim is jobless). Note that the actual numbers in the dataset are slightly different from the ones discussed in Chapter 5!

Questions:

1. List the variables included in this study and, for each variable, say what its function is (dependent, independent, etc.) and its type (nominal, ordinal, interval).
2. How would you formulate H_0 and H_a?
3. Which statistical test could be used?
4. There are two ways in which data can be entered for our analysis. The first and usually most common option also for other datasets is when the data

Table J.3 Contingency table dividing the number of respondents according to social class and the use of '*don't have*' and '*haven't got*'

	low social class	high social class
haven't got	70	59
don't have	64	31

Table J.4 Example of data in a long format

Participant	Variable 1	Variable 2
1	NO	YES
2	NO	NO
3	YES	NO
Etc.		

are organized for each individual case. An example of this type of data organization would be the format in Table J.4, which is generally referred to as long format (such as the data used in Part A of this practical).

The second option is when you only have the total frequencies for each of the cells in the contingency table, such as the one in Table J.3.

For this practical, we will use the raw data in the long format. These can be found in the file 'Prac4B_data.csv'. The data file consists of values of 1s and 2s. For *social class*, 1 is high and 2 is low social class. For the *reply*, 1 is 'haven't got' and 2 is 'don't have'. Open the file in JASP.

5. Go to *Frequencies* > *Contingency Tables* and move one variable to *Rows* and the other to *Columns*. To check the expected values, select *Expected* under *Cells*. Under *Statistics*, you want to deselect the χ^2 (chi-square) symbol, since we are not interested in that yet.

   Although we are dealing with a non-parametric test, we do have to check some assumptions before conducting the actual test. One assumption is that every subject only contributes to one of the cells, which can normally be checked by comparing the number of subjects to the total of all cells. In this particular example, you can assume that this one has been met.

   Secondly, as mentioned in Section 5.2.3, in a 2x2 table, none of the *expected* frequencies in the table should be lower than 5. Do note that in a larger table, the expected counts must be at least 1, and no more than 20% of the cells are allowed to be less than 5. If your expected cell frequencies are below 5, you should look at the outcome based on the χ^2 *continuity correction,* which is also known as the Yates correction. With samples smaller than about 30, it is advised to use the *Likelihood ratio* test.

   You should be able to find the expected values in the output provided. Has this assumption been met?

6. Go to *Frequencies* > *Contingency Tables* again, and move the variables again. At this point we want to make sure that the χ^2 symbol is selected and that *Expected* is deselected. It is also useful to select the different *Percentages* (*Row*, *Column*, and *Total*), as these will give us the relative frequencies in each cell.

   The actual results of the chi-square test can be found in the second table of the output. Can you reject the null hypothesis?

7. What is the effect size? Click on *Phi and Cramer's V* under *Statistics*. Phi (φ) is used for variables with 2 levels and Cramer's V for variables with more than 2 levels. A value of .1 is considered a small effect, .3 a medium effect, and .5 a large effect.
8. In other statistics programs it is sometimes possible to visualize the results of a chi-square analysis in a barplot. In JASP this is not quite possible (yet), but it is possible to make two histograms under *Descriptive Statistics*, one for *social class* and one for *reply*. Try to make these histograms. Which values in the contingency table do the bars in these histograms correspond to?
9. A template for reporting the results of a chi-square would be (please choose an option or fill in the correct numbers between the {accolades}):

> 'A chi-square analysis revealed that the association between {variable 1} and {variable 2} was {significant/not significant}, χ^2({fill in the *df*}, *N*={fill in the total number of participants}) = {fill in χ^2-value}, p = {fill in the exact p-value or, if it is less than .001, $p < .001$}.'

In the case of significant results, do not forget to report on the direction of the association (e.g. men having a stronger preference for the colour blue). Also remember that, especially in the case of significant results, you would want to report Phi (φ) or Cramer's V.

Report on the results of this study in the way that it is conventionally done in APA research papers.

Part C: Gender and intelligence

A researcher wants to find out whether boys or girls are more intelligent. Eleven girls and eight boys (randomly selected) participated in an experiment in which the range of scores was 1–20 (interval).

Data:

Girls	Boys
17	16
16	15
14	13
19	19
18	15
17	14
16	13
15	12
16	
15	
19	