DO YOU KNOW EVERYTHING THAT YOU NEED TO KNOW ABOUT CORRELATIONS?

3.1

1. a) What type of graph is used to show correlational data?

    b) Sketch one of the above that shows:

     i) a positive correlation   ii) a negative correlation   iii) no correlation

2. What are the highest and lowest correlation coefficients that are possible?

3. There is no IV and DV in a correlation – what do correlations use instead?

4. How would you judge the strength of a correlation from its graph?

5. What does it mean in terms of variables if they are described as:
    a) Negatively correlated

    b) Positively correlated

6. Suggest two variables that might be:
    a) Negatively correlated

    b) Positively correlated

7. Many hypotheses you have written will be experimental but correlations need one too. Write appropriately operationalised null and directional hypotheses for the following:
    a) The relationship between age and running speed over 100 metres.

    b) Time taken to revise for Psychology Mock exam and the score obtained.

    c) Reaction time and alcohol units consumed.

8. State two potential **advantages** of using a correlational analysis.

9. State two potential **disadvantages** of using a correlational analysis.

Analyse that personal statement!

3.2

Beth has decided to study psychology at university and wants to ensure that her personal statement is the best it can be. She decides to do her own piece of research on what are the best things to include in a personal statement and turns to Google to help her. She knows that the first few lines are vital and decides to try and find out what people generally include here. She finds some examples and selects five that she thinks are good. Then she thinks it might be fun to run a content analysis on them. Read through the examples she has selected and then answer the questions below.

1 From an early age I have been intrigued by what makes humans tick and this interest has only been enhanced through my study of psychology. Although I have only been studying it for just over a year I have really enjoyed all the topics that we have covered at A level and feel that I have a solid grounding for university level study. I recognise that we can apply what has been learned in psychology to many aspects of life and as a subject it is just so 'relevant'. That is what makes me sure I want to make this my career. What could be better than studying the human mind?

2 Sometimes it seems like I have wanted to study psychology forever! The way people behave and why they do certain things has always fascinated me. I am really glad I chose the subject at A level because it has been my favourite. I really want to learn more now and study psychology at university and even have a career in it.

3 It was the odd behavioural traits displayed by my uncle, who was diagnosed with bipolar disorder that really sparked my interest in psychology. What was it that caused him to one day be a sparkling, witty person running around and playing silly games with me and then the next time I saw him, a hunched over isolated person that could not bear to make eye contact? Psychology has the answer to this and many other everyday questions so I want to study it to find out the answers so that I can have a career that is focused on helping people like my uncle.

4 Last year I was unfortunately involved in quite a serious car accident. What amazed me was that as I sat among the wreckage, not one person stopped to help me. Didn't anyone care at all about another human being? Studying psychology meant I could answer this – I realised I knew the answer. At that moment I thought – if I know this after just studying it for a few months what could I learn from a degree in the subject. Despite my injuries, I was excited and knew that psychology was for me, not just as a degree but as a career.

5 What motivates humans to behave like we do? Is it all about external factors or perhaps it is down to internal factors? This is one of the key questions in psychology and one that I hope to answer. Psychology has already told us lots about important topics such as the causes and treatment of mental illnesses but the beauty of the subject overall is the fact that it can literally be applied to everyday life. No longer do I want psychology to be just an interest of mine, but want it to be my career.

Questions

1. Look through the opening paragraphs above and construct a list of things that people include in this paragraph, you can include the type of language or sentences as well as the actual content, e.g. experiences.

2. Working in a pair, construct a list of behavioural categories (aim for around 4–6) that you can use to compare these statements.

3. Analyse each example in terms of these categories by adding a tick each time that category is present in an example.

4. Transfer this data to a bar chart.

5. What conclusions can Beth draw about opening paragraph content?

6. Beth thinks that she now knows exactly what is required for a perfect Psychology personal statement but her Psychology teacher urges caution. Explain why this might not tell her what is the best content.

Fill in the gaps – Reliability

3.3

Use the terms below to fill in the gaps.

behavioural categories	closed	time	same person	same results
inter-observer	time	compared	pilot study	leading
psychological tests	interviewer	same results	interviews	structured
correlation coefficient	questionnaires	time	complex	standardised
correlation coefficient	ambiguous	open		

If a psychological measure can be used multiple times with the ……………… ………………… it is said to be reliable.

For example, a reliable test of intelligence would yield the ……………… ………………… on the ……………… ………………… each time the measurement is taken.

Assessing reliability

There are two key ways of assessing reliability and they are: the test-retest method and inter-rater reliability.

The test-retest method is commonly applied when assessing the reliability of …………………………………………,

……………………………………………… ……………, e.g. of personality or IQ and ……………… . The test is administered

twice and the results are ……………………………… . If scores are obtained then a ……………………………… ……………

……………………………… can be calculated. A reasonably reliable result would be one where the

……………………………… ……………………………………………… was +.80 or more. Note that one of the challenges

with this method is deciding on the ……………… lapse between tests. There must be sufficient ………………… to be

sure that the participant is not simply recalling their previous responses but not too much ………………… in case the

attitude or ability being tested actually changes.

Inter-rater or ……………………………………… reliability involves checking the consistency of ratings that two or

more independent researchers have completed. The two observers will apply the behavioural categories to a test run (also

called a ……………………… …………………………) and the consistency of the results will be assessed.

Improving reliability

It is important to note that assessing reliability is only measuring it and if the reliability level is found to be poor then

further steps will need to be taken.

In questionnaires that are found to be unreliable, ……………………………… or ………………… questions

may be removed. Sometimes this involves replacing ……………… questions (open to individual interpretation) with

………………………… ones with a more restricted range of responses.

In interviews, reliability can be improved by using the same ……………………………… with all participants or at

least training them all in the same way, e.g. to avoid ………………… questions. The more ………………………………

an interview is, the more reliable it is likely to be.

In experiments, reliability can be improved by ensuring that ……………………………… instructions and

procedures are used.

Observation reliability can be improved by careful operationalisation and explanation of the …………………………

……………………………… . They should be as independent as possible, i.e. avoiding overlap.

Working on reliability

3.4

RELIABLE

Task: Try these questions to check your understanding of assessing and improving reliability.

1. Joshvir has designed a questionnaire to measure stress that he believes is reliable.

   a) Explain fully how he should test the reliability claim.

   b) What would indicate that he needs to improve the reliability of the questionnaire?

   c) What actions could he take to improve reliability if it was not found to be reliable the first time?

2. Abi wants to interview young people about their attitudes to British foreign policy. She has recruited 20 volunteer participants and her friend Holly has agreed to run some of the interviews to save Abi some time. Abi jots down a few possible questions for each of them to ask but really wants to let the interviews flow freely to get lots of information. Abi's teacher is concerned about the reliability of this methodology and asks Abi to rethink her plans.

   Explain why the teacher is right (of course!) to be concerned.

3. Max is planning an experiment to investigate whether smells might enhance learning. He hopes to find that if people have a particular scent present when they are learning something they will find recall of what they have learned easier. He has lavender and peppermint oil, which he will spray into a classroom. Students are asked to learn a list of words selected at random from a box. They do this either in the room smelling of lavender or a room smelling of peppermint. Later he meets up with each student and asks them what words they can recall.

   Accurate methodological descriptions of studies are important to allow replication – what factors here would need to be made explicit to allow accurate replication?

4. Andy and Charlotte are heading for the sixth form common room to observe the difference in behaviour between Year 12 and 13 students.

   a) Suggest five behavioural categories that they may use – ensuring that they will produce reliable results. Justify why you believe that they can elicit reliable data.

   b) If they both observe at the same time, explain how their results may be compared to check reliability.

   c) If it was found from the process that the reliability of the results was poor – what should they do to address this problem?

Validity and reliability definition cards

3.5

Validity

The extent to which an observed effect is genuine. The extent to which the researcher has measured what they intended to and the extent to which the findings can be generalised.

External validity

The extent to which an observed effect can be generalised, for example to other times, people and settings

Internal validity

The extent to which the effects of a study are due to the manipulation of the independent variable. The extent to which the researcher has measured what they intended to measure.

Ecological validity

A form of external validity focusing on the extent to which the research effect is likely to be demonstrated in other settings.

Temporal validity

A form of external validity focusing on the ability to generalise findings beyond the particular historical context of the study.

Concurrent validity

A form of internal validity. The extent to which a psychological measure relates to an existing similar measure.

Face validity

A form of internal validity focusing on the extent to which a questionnaire, etc., appears (on the face of it) to be measuring what it claims to.

Reliability

The consistency of measurements of procedures in a research study.

Inter-rater reliability

An assessment of the extent to which two or more ratings/observations of the same behaviour are consistent.

Test—retest reliability

An assessment of the consistency of a psychological test or questionnaire to see how similar scores/responses are when a person is tested twice.

WORKING ON VALIDITY   3.6

Task: Try these questions to check your understanding of assessing and improving validity and reliability

VALID

1. Joshvir has designed a questionnaire to measure stress that he believes is valid.

   a) Explain fully how he should assess the concurrent validity of the test.

   b) Suggest one way he might improve the validity of the questionnaire.

2. Abi wants to interview young people about their attitudes to British foreign policy. She has recruited 20 volunteer participants and her friend Holly has agreed to run some of the interviews to save Abi some time. Abi jots down a few possible questions for each of them to ask but really wants to let the interviews flow freely to get lots of information. Abi's teacher is concerned about both the reliability and validity of this methodology and asks Abi to rethink her plans.

   What might Abi do to ensure that her attitude measurements are as valid as possible?

3. Max is planning an experiment to investigate whether smells might enhance learning. He hopes to find that if people have a particular scent present when they are learning something they will find recall of what they have learned easier. He has lavender and peppermint oil, which he will spray into a classroom. Students are asked to learn a list of words selected at random from a box. They do this either in the room smelling of lavender or a room smelling of peppermint. Later he meets up with each student and asks them what words they can recall.

   What are the concerns in terms of validity of this research plan and how might they be addressed?

4. Andy and Charlotte are heading for the sixth form common room to observe the difference in behaviour between Year 12 and 13 students.

   Suggest two issues that they might consider to ensure the best possible validity of their data.

Chapter 3: Research methods
Choosing a statistical test

70–71

Data Analysis

Types of data practice

3.8

Revising qualitative and quantitative data

In each of the following cases, imagine you are writing a questionnaire to collect data and state one question that would enable you to collect **quantitative** data and one that would yield **qualitative** data.

1.   A researcher wants to measure stress levels in teachers.

2.   A researcher wants to measure differences in smoking behaviour and attitudes towards smoking between teenage boys and girls.

3.   A researcher wants to investigate sleep problems in new mothers.

4.   A researcher wants to investigate levels of anger in rail passengers whose train has been delayed.

Types of data

Quantitative data can be classified as **Nominal**, **Ordinal** or **Interval**.
In each of the cases below identify the data produced.

1. Heart rate in newborn babies.	
2. How many people in a class would vote for each of the main political parties.	
3. Ratings of how happy each student feels about their psychology test score.	
4. The time taken for participants to complete a questionnaire.	
5. Patients' ratings of satisfaction with the service of their dentist.	
6. The number of calories consumed by a participant on each day.	
7. How many boys and girls each choose to do either an apprenticeship or a degree.	
8. Favourite foods of a group of nurses.	
9. A study of the main mode of transport to school of Year 7.	
10. The weight of participants beginning a diet study.	

Extension activity

Thinking of studies that you have covered so far, state **two** that have collected quantitative data and **two** that have collected qualitative data and state what data you are referring to.

In each of the following cases, write a suitably operationalised:
a) **Null hypothesis** H_0
b) **Directional alternative hypothesis** H_1

1 An investigation to see if people who regularly watch football are better at remembering a set of football scores than people who never watch football.

H_0

H_1

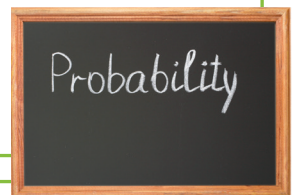**2** An investigation to see if people farming in Wales or England suffer more symptoms of stress.

H_0

H_1

3 An investigation into the relationship between alcohol consumed per month and work absence.

H_0

H_1

4 What is another term for a one-tailed hypothesis?

5 When would you use a one-tailed hypothesis?

6 What is the usual level of significance employed in psychology? Write as decimal and percentage.

7 Justify why psychologists use the level stated above.

8 Under what circumstances might a more stringent level of significance be used and what would the level be?

9 What is a Type II error?

10 When is a Type II error most likely to occur?

WHICH TEST OF DIFFERENCE?

3.10

Below you will find four studies – for each you have to decide and justify which test of difference you would use to test significance of the results. You must state the design and the type of data being collected.

A psychologist, who is also a football fan, believes that her club's fans (United) sing more passionately than their rival's fans (City). She collects data by getting two observers to sit with one or other of the two sets of fans and rate the passion of the singing. In total they attend ten games at each club.
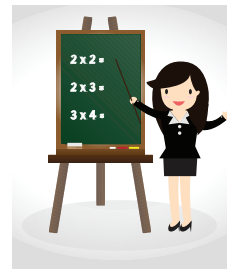
Level of measurement:

Design:

Test to be used:

A primary school teacher wants to see whether her class will enjoy Maths lessons better before lunch or after. She gives them a rating scale each time she holds a Maths lesson and wants to see if there is a significant difference in their rating of how much they enjoy Maths. The scale is 1 to 10 and is accompanied by smiling or unhappy faces according to the rating.

Level of measurement:

Design:

Test to be used:

A psychologist wants to investigate the issue of parking and gender. He visits a multi-storey car park for 2 hours a day every day for a week and gives marks out of 10 for how well people park, taking into account number of attempts, time taken and the straightness of the vehicle. He adds these marks and then ranks the parkers on how well they completed the manoeuvre. He now wants to see if there is a significant difference in how well males and females parked.

Level of measurement:

Design:

Test to be used:

A researcher believes that aggression in television programmes varies in summer and winter months. They decide to conduct a study on whether programmes in winter months can be considered more aggressive than those in summer months. They ask 10 observers to rate the aggression of 25 popular programmes – one episode from each series in January and June. They ask for ratings on a scale of 0–50 based on incidents, intensity of violence, etc. (where 0 is not at all aggressive and 50 is extremely). The researcher will produce a mean rating (from the 10 observers) for each of the 25 programmes and compare those in January with those in June.

Level of measurement:

Design:

Test to be used:

Extension activity
Write null and suitable alternative hypotheses for each of these studies.

Essay Masterclass

3.11

A Psychology teacher is keen to find out whether the masterclass he gave on essay writing has actually improved his students' ability to write A level Psychology essays. He sets an essay on social influence, then runs the masterclass and sets another on attachment.

He marks both sets and sorts them into who they were written by and whether it was before or after the masterclass. The teacher then wants to see whether the difference in marks is significant. The marks for each student are given on the right.

Student	Essay mark before masterclass	Essay mark after masterclass	Difference	Rank of difference
1	12	17		
2	14	20		
3	15	16		
4	19	17		
5	13	17		
6	14	15		
7	15	18		
8	16	16		
9	15	17		
10	19	18		
11	16	14		
12	15	17		
13	18	18		
14	16	18		
15	16	15		

1. Write a suitable **null** and **alternative** hypothesis for this study. State whether the alternative hypothesis you have chosen is **directional** or **non-directional**.

2. Calculate the mean and modal results for before and after the masterclass. Give the result to one decimal place where appropriate.

3. Justify why Wilcoxon would be the correct test to use in this case.

4. Complete the columns in the table above marked *Difference* and *Rank of Difference*.

5. Once the teacher had gathered all the results he used Wilcoxon to find out whether or not the difference was significant. He ended up with a calculated value of $T = 19$

   Were the results significant at the 5% level? Explain your answer.

   Use the critical value tables for Wilcoxon as given on page 75 of the textbook and refer to whether or not your hypothesis was directional or non-directional.

6. The method here has some flaws – suggest **two** improvements that could have been made to the design of the study to ensure that the results were more reliable and/or valid.

Using tables of critical values

3.12

How confident are you about this topic as you start? **Please tick**	Not at all	A little confident	Neutral	Quite confident	Very confident

Using tables of critical values 1

For a one-tailed test at 5%

N_1	5	6	7	8	9	10	11	12	13	14	15
N_2											
5	4	5	5	8	9	11	12	13	15	16	18
6	5	7	8	10	12	14	16	17	19	21	23
7	6	8	11	13	15	17	19	21	24	26	28
8	8	10	13	15	18	20	23	26	28	31	33
9	9	12	15	18	21	24	27	30	33	36	39
10	11	14	17	20	24	27	31	34	37	41	44
11	12	16	19	23	27	31	34	38	42	46	50
12	13	17	21	26	30	34	38	42	47	51	55
13	15	19	24	28	33	37	42	47	51	56	61
14	16	21	26	31	36	41	46	51	58	61	68
15	18	23	28	33	39	44	50	55	61	66	72

For a two-tailed test at 5%

N_1	5	6	7	8	9	10	11	12	13	14	15
N_2											
5	2	3	5	6	7	8	9	11	12	13	14
6	3	5	6	8	10	11	13	14	16	17	19
7	5	6	8	10	12	14	16	18	20	22	24
8	6	8	10	13	15	17	19	22	24	26	29
9	7	10	12	15	17	20	23	26	28	31	34
10	8	11	14	17	20	23	26	30	33	36	39
11	9	13	16	19	23	26	30	33	37	40	44
12	11	14	18	22	26	29	33	37	41	45	49
13	12	16	20	24	28	33	37	41	45	50	54
14	13	17	22	26	31	36	40	45	50	55	59
15	14	19	24	29	34	39	44	49	54	59	64

Observed value of U must be **equal to** or **less than** the critical value in this table for significance at the level shown.

Reminder – in order to use a critical value table you need to know:

a) If the hypothesis is one- or two-tailed
b) Number of participants (degrees of freedom in Chi^2)
c) Level of significance

It is good practice to write out any significance decisions in full, e.g.

The calculated value of T is less than the critical value so the result is significant for a one-tailed hypothesis where $N_1 = 15$ and $N_2 = 15$ at the 5% significance level.

Practise doing this with the examples below, using the appropriate table above.

1. Observed Value U = 22 $N_1 = 10$ $N_2 = 10$ One-tailed test   5%

2. Observed Value U = 18 $N_1 = 8$ $N_2 = 8$ Two-tailed test   5%

3. Observed Value U = 37 $N_1 = 12$ $N_2 = 13$ Two-tailed test   5%

4. Observed Value U = 74 $N_1 = 15$ $N_2 = 15$ One-tailed test   5%

5. Observed Value U = 16 $N_1 = 7$ $N_2 = 10$ Two-tailed test   5%

How confident are you? **Please tick**	Not at all	A little confident	Neutral	Quite confident	Very confident

Chapter 3: Research methods
Parametric tests of difference:
Unrelated and related t-tests

📖 76-77

Related and unrelated t-tests

3.13

Complete the following table as a reminder:

Test	Experimental design(s)	Type of data	Distribution of data	Degrees of freedom
Related				
Unrelated				

Related and unrelated *t*-tests

First establish which test would have been used and then use the table of critical values for *t* in the text book (page 76) to find out whether or not results are significant in the following cases (show your working!).

1
Calculated value of *t* was 3.68
Hypothesis directional
Design matched pairs
Participants 15 in each group
Significance level 0.05.

2
Calculated value of *t* was 1.37
Hypothesis non-directional
Design independent groups
Participants 15 and 12
Significance level 0.05

3
Calculated value of *t* was 2.26
Hypothesis non-directional
Design repeated measures
Participants 8
Significance level 0.05

4
Calculated value of *t* was 2.89
Hypothesis non-directional
Design independent groups
Participants 10 and 10
Significance level 0.10

Chapter 3: Research methods
Tests of correlation:
Spearman's and Pearson's

📖 78–79

Spearman and Pearson

3.14

Use the following table as a reminder:

Test	Method	Type of data	Number used to check critical value
Spearman's Rho	Correlation	Ordinal	N = number of participants
Pearson's r	Correlation	Interval	df = number of participants – 2

Task 1   In the following examples, which statistical test of correlation would be used?

1) An investigation into the relationship between the number of hours of day care a child experiences per week and their score on an IQ test.

2) An investigation into the relationship between the number of Christmas cards received and the number of Christmas party invitations.

3) An investigation into the relationship between the number of hours study completed per week and stress level measured by increases in pulse rate.

4) Student performance was compared on a psychology test and a maths test to see if students who did well on one test also did well on the other. Test scores were in terms of ranked position on the test.

Task 2   For each example below:

a) **Use the tables of critical values in the textbook on pages 78 and 79 to establish whether or not results are significant (show your working!).**

b) **Comment on the strength of the correlation.**

1
Calculated value of rho +.68
Hypothesis directional
Participants 15 in each group
Significance level 0.05.

2
Calculated value of rho –.18
Hypothesis non-directional
Participants 12
Significance level 0.05.

3
Calculated value of r +.52
Hypothesis non-directional
Participants 8
Significance level 0.05

4
Calculated value of r –.95
Hypothesis directional
Participants 10
Significance level 0.05

Complete the following as a reminder

The Chi-Squared test is used where the data is _____ ; for example, frequencies of a particular colour.

The table of results is known as a _____ _____.

In order to establish the critical value for a particular experiment we need to know the degrees of

_____. This is calculated by _____

Questions:

1. Which of the following examples would require the use of Chi-Squared? If it is not Chi-Squared, which test would you use? Justify your response in each case.

   a) An investigation where the short-term memory of a group of students was tested in the morning and afternoon to find out which was better.

   b) An investigation into the relationship between participants' scores on a stress questionnaire and their height.

   c) An investigation into whether or not one parent went to university and whether or not their children plan to.

   d) An investigation into the efficacy of a new antidepressant, which involved half of the participants taking the drug for three months and half taking a placebo. Effectiveness was measured using the Beck Depression Inventory and the change in symptoms for the two groups was compared.

   e) An investigation into whether or not male and female 17-year-olds were learning to drive.

2. Using the table on page 80 of the textbook find the critical values for the following:

   a) For a directional hypothesis, significance level of 0.05 and df = 2

   b) For a directional hypothesis, significance level of 0.05 and df = 3

   c) For a directional hypothesis, significance level of 1% and df = 1

   d) For a directional hypothesis, significance level of 10% and df = 1

A reporting summary

3.16

Use your textbook to write condensed notes about each section of a journal article.

Section of report	Description
Abstract	Summary of: - -
Introduction	
Method	Subsections: - - - - - A detailed method section will allow accurate replication.
Results	Summaries of the data in graphical, tabular and written formats include: - - -
Discussion	Key elements are: - - - -
Referencing	All sources must be included, for example Example of a journal reference: Example of a book reference:
Appendices	Will include raw data

Extension activity

Look through the references for the textbook. These can be found online at www.illuminatepublishing.com/psychreferences

Search Google to find a copy of one of the articles listed.

To what extent does the article follows the pattern described in your table?

What feature of science am I?

3.17

Consider the following features that are being described and simply insert the word beneath each description. Identify who, or which approach, might have made it.

1
I am the principle that a theory cannot be considered scientific unless it admits the possibility of being found untrue.

2
I involve minimising all sources of personal bias to prevent distortion or influence on the research process.

3
I am an agreed set of assumptions and methods within a scientific discipline.

4
I am the process of deriving new hypotheses from an existing theory.

5
I am the scientific process of gathering evidence through direct observation and experience.

6
I am the extent to which scientific procedures and findings can be repeated by other researchers.

7
I am a significant change in the dominant unifying theory within a scientific discipline, brought about by new findings and understanding.