



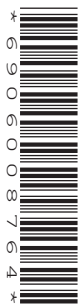
Oxford Cambridge and RSA

A Level Psychology

H567/01 Research methods

Monday 4 June 2018 – Afternoon

Time allowed: 2 hours



You must have:

- a scientific or graphical calculator



First name	StemUp Model Solution									
Last name										
Centre number						Candidate number				

INSTRUCTIONS

- Use black ink.
- Complete the boxes above with your name, centre number and candidate number.
- Answer **all** the questions.
- Write your answer to each question in the space provided. If additional space is required, use the lined page(s) at the end of this booklet. The question number(s) must be clearly shown.
- Do **not** write in the barcodes.

INFORMATION

- The total mark for this paper is **90**.
- The marks for each question are shown in brackets [].
- Quality of extended responses will be assessed in questions marked with an asterisk (*).
- This document consists of **20** pages.

SECTION A: Multiple choice

Answer **all** the questions.

1 Which of these is a weakness of a quasi experiment?

- A control of the dependent variable
- B control of the independent variable
- C control of the measurement of the dependent variable
- D control of the measurement of the independent variable

Your answer

B

[1]

2 Which does not have both an IV and DV?

- A correlation
- B field experiment
- C laboratory experiment
- D quasi experiment

Your answer

A

[1]

3 What is it best to do with extraneous variables?

- A eliminate them
- B ignore them
- C monitor them
- D record them

Your answer

A

[1]

4 What best describes the target population?

- A the people you want to study and apply the findings to
- B the people you want to study and conduct research on in a follow-up study
- C the people you want to study and obtain data from
- D the people you want to study and use in the research

Your answer

A

[1]

5 What is the name given to data before any analysis is performed?

- A interval
- B ordinal
- C quantitative
- D raw

Your answer

D

[1]

6 What is the name for the type of reliability that involves dividing a test into two parts and comparing scores on both parts of the test?

- A semi-structured
- B split-half
- C test-retest
- D two-tailed

Your answer

B

[1]

7 Which of these inferential statistical tests does not require the data to be ranked as part of the calculation?

- A Chi-square
- B Mann-Whitney U test
- C Spearman's Rho
- D Wilcoxon Signed Ranks test

Your answer

A

[1]

8 Which of these is an advantage of secondary data?

- A already exists
- B easy to analyse
- C easy to interpret
- D ecologically valid

Your answer

A

[1]

9 What is meant by the term 'significant result'?

- A it is a figure that you compare the answer from an inferential statistical test with
- B it is an answer that exceeds a certain probability level
- C it is an answer that tells us something important
- D it is an approximate answer

Your answer

B

[1]

10 In research terms, what is meant by 'social desirability'?

- A responding in a way that is approved of by society
- B responding in a way to be perceived as more friendly
- C responding in a way to please the researcher
- D responding in a way to provide the researcher with what is expected

Your answer

A

[1]

11 In which section of a practical report write-up would you find details of standardised instructions given to participants?

- A abstract
- B appendices
- C discussion
- D introduction

Your answer

B

[1]

12 What is 'the Harvard system'?

- A a way of presenting results from an inferential statistical test in a practical report
- B a way of providing details of the materials used in a practical report
- C a way of summarising how participants were obtained in a practical report
- D a way of writing academic references in a practical report

Your answer

D

[1]

13 Which best describes what a semantic differential rating scale is?

- A** selecting a point on a line to indicate your strength of opinion about something
- B** selecting a point on a line to respond to how much you agree with something
- C** selecting a point on a line with different numbers along it
- D** selecting a point on a line with words that have opposite meanings at either end

Your answer

D

[1]

14 Who conducts a peer review?

- A** a government minister
- B** a statistician
- C** fellow academics
- D** lay persons

Your answer

C

[1]

15 Which of these different types of data would not include any information in the form of words?

- A** interval
- B** nominal
- C** qualitative
- D** secondary

Your answer

A

[1]

16 Which of these could not be the answer from a Spearman's Rho inferential statistical test?

- A** -0.728
- B** 0.3
- C** 0.892
- D** 1.52

Your answer

D

[1]

17 What is the mode in this set of data?

17, 19, 12, 23, 17, 25, 19, 17

A 17

B 17.5

C 18

D 19

Your answer

A

[1]

18 The variance of a set of scores is 14.44. What is the standard deviation?

A 1.4

B 3.8

C 14.0

D 208.51

Your answer

B

[1]

19 Which variable was negatively correlated with length of time as taxi driver in the Maguire et al. study?

A volume of grey matter in the anterior hippocampus

B volume of grey matter in the central hippocampus

C volume of grey matter in the hippocampus

D volume of grey matter in the posterior hippocampus

Your answer

A

[1]

20 What type of data was collected in the Piliavin et al. study for the dependent variable of length of time that it took for help to be offered?

A interval

B non-parametric

C qualitative

D secondary

Your answer

A

[1]

PLEASE DO NOT WRITE ON THIS PAGE

Turn over for the next question.

SECTION B: Research design and response

Answer **all** the questions.

Travel log. How do people pass the time on a long journey? Some people seem to cope better than others and manage to keep themselves occupied or simply don't mind (perhaps even enjoy) the rest. Others find it more difficult and can become bored and restless. Psychologists want to use the naturalistic observation method to investigate this and find out if some types of behaviour and actions are more common than others.

21 Write an appropriate research aim for the study.

The aim is to investigate which behaviours people engage in to pass the time on a long train journey, and whether some behaviours, such as reading, using a phone or sleeping, are more common than others.

[2]

22* Explain how you would use the naturalistic observation method to conduct this research. Justify your decisions as part of your explanation.

In your answer, the required features that you must refer to are:

- participant or non-participant observation
- behavioural categories
- time or event sampling
- how data will be recorded during the observation

You should use your own experience of practical activities to inform your response.

I would carry out a covert non-participant observation on a long distance train journey from Manchester to London, sitting in an ordinary seat at the end of a carriage with a clear view of about ten passengers but not interacting with them. Non-participant observation is justified because I can concentrate fully on recording accurately rather than having to join in conversations, and because the passengers will not know they are being watched, so their behaviour will be natural and free from demand characteristics. A train is a public place, so observing without formal consent is acceptable.

Before the journey I would draw up a list of behavioural categories, operationalised so that each is clear, observable and mutually exclusive: reading a book or newspaper, using a phone or laptop, listening to music with headphones, sleeping, eating or drinking, talking to another passenger, and looking out of the window. Clear categories are justified because

they turn the aim into measurable behaviours and let a second observer code the same behaviour in the same way, which raises inter-rater reliability.

I would use time sampling: every five minutes for the whole two hour journey I would look at each of the ten passengers in turn and record which category they are showing at that exact moment. Time sampling is justified because behaviours such as sleeping or reading last a long time and have no clear start or end, so event sampling, where every single occurrence of each behaviour is tallied across the whole journey, would be impractical for one observer and would over count brief behaviours. Time sampling keeps the data manageable and gives a frequency for each category that can be compared.

Data would be recorded on a pre-printed tally chart on a clipboard, with a row for each passenger and a column for each five minute interval, writing the code letter of the category observed. Recording would be covert, as someone writing in a notebook looks ordinary on a train, and I would not film because recording members of the public on video without consent would be unethical. The tallies are quantitative and can be totalled to show which categories are most common.

In my own practical activity I carried out a non-participant observation of mobile phone use in the sixth form common room using time sampling every two minutes and a tally chart. I found that two of my categories overlapped, so here I would pilot the categories first, and I found a tally chart quick enough to keep up with ten people, so it would work well on the train.

23 Describe **two** things that may influence the inter-rater reliability of this study.

The first thing is how clearly the behavioural categories are operationalised. If a category such as 'using technology' is vague, one observer on the train might count a passenger listening to music through their phone as using technology while the other counts it as listening to music, so their tallies would differ and inter-rater reliability would be low. Precise definitions agreed and piloted beforehand would make it high.

The second thing is where the observers are positioned in the carriage. If one observer sits at the front and the other at the back, each will see different passengers' faces and screens, so one might record a passenger as sleeping while the other, who can see the phone in their hand, records them as using a phone. Sitting side by side with the same view of the same passengers would increase agreement between the two sets of records.

[6]

24 Outline **one** strength and **one** weakness of using the naturalistic observation method in this study.

One strength is that passengers on the train are in their normal environment and do not know they are being observed, so the behaviours recorded, such as sleeping, reading or looking out of the window, are natural rather than altered by demand characteristics. This means the findings about which behaviours are most common on a long journey are high in ecological validity and reflect what people really do to pass the time.

One weakness is that the researcher has no control over extraneous variables on the journey. For example, if the train is crowded, the wifi stops working or there is a delay, passengers may talk, look out of the window or become restless far more than usual, so the frequencies recorded would reflect that particular journey rather than long journeys in general, lowering the validity of the conclusions about common behaviours.

[6]

- 25 Describe **two** ways you would address the ethical consideration of 'responsibility' in relation to this study.

One way is to protect the passengers from harm. Responsibility means ensuring participants are not caused physical or psychological distress, so I would observe discreetly from my seat without staring, filming or approaching anyone, and if a passenger noticed and seemed uncomfortable at being watched I would stop recording them straight away. I would also record no names or faces, only coded behaviours, so no individual could be identified or embarrassed.

A second way is to debrief. Because the observation on the train is covert, at the end of the journey I would, where practical, explain to the passengers observed that I had been recording how people pass the time on a long journey, explain why they were not told beforehand, offer to share the findings and give them the right to withdraw their data from the tally chart if they were unhappy about it.

[6]

SECTION C: Data analysis and interpretation

Answer **all** the questions.

Love is in the (question) **air**. We can fall 'in to it' and 'out of it', it can make us feel both happy and sad or even angry and mad, but what is love? Psychologists decided to investigate this further by using the self-report method in which they posted out a questionnaire to members of the public to complete. Some of the data from the first twenty males and first twenty females to respond are presented in the tables below.

Table 1: Ratings of the importance of physical appearance for being in love given by male and female participants

(1 = 'not very important' to 10 = 'extremely important')

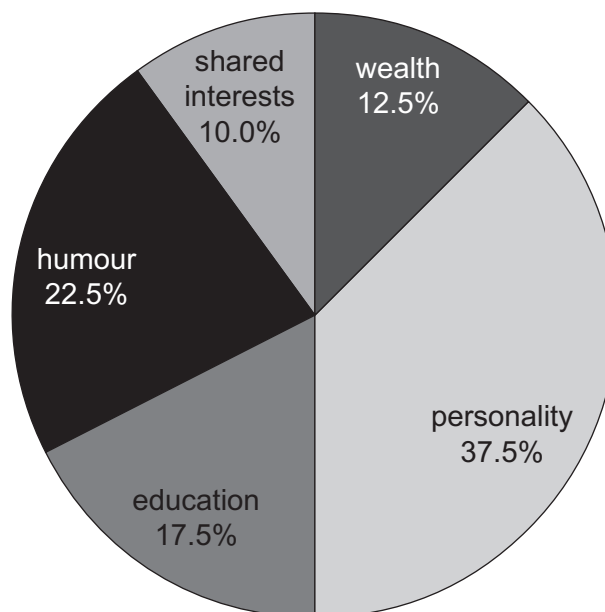
males	females
7	3
5	7
9	1
10	6
8	4
7	9
7	6
9	3
8	5
10	3
9	5
10	3
9	5
9	5
7	3
6	7
5	5
6	1
7	3
9	2

Table 2: Responses to the question ... 'do you believe in love at first sight?'

	yes	no
males	5	15
females	12	8

Figure 1

Pie chart showing the percentage of people who thought which attributes were the most important for being in love with someone



- 26 A partly completed table of measures of central tendency for the data collected in Table 1 is presented below.

Measures of central tendency for the ratings of the importance of physical appearance for being in love given by male and female participants		
	males	females
mean		4.3
median	8.0	
mode	9.0	

Answer the following questions so that the rest of the table can be completed. Show all your workings.

- (a) Calculate the mean rating of the importance of physical appearance for being in love given by males. Write your answer to two significant figures.

Mean = total of all the male ratings divided by the number of males

$$7+5+9+10+8+7+7+9+8+10+9+10+9+9+7+6+5+6+7+9 = 157$$

$$157 / 20 = 7.85$$

To two significant figures the mean male rating = 7.9

[3]

- (b) Calculate the median rating of the importance of physical appearance for being in love given by females.

Female ratings in order: 1,1,2,3,3,3,3,3,4,5,5,5,5,5,6,6,7,7,9

With 20 values the median is halfway between the 10th and 11th values.

$$(4 + 5) / 2 = 4.5, \text{ so the median female rating} = 4.5$$

[2]

- (c) What is the mode for the importance of physical appearance for being in love given by females?

The mode is 3, as it is the most frequent female rating (it appears six times).

[1]

- 27 (a) What is the range for the rating of the importance of physical appearance for being in love given by males **and** females? Show your workings.

Range = highest value minus lowest value

Males: highest rating = 10, lowest rating = 5

Male range = $10 - 5 = 5$

Females: highest rating = 9, lowest rating = 1

Female range = $9 - 1 = 8$

So the range for males is 5 and the range for females is 8.

[4]

- (b) Outline **one** conclusion from the calculation of the range for the rating of the importance of physical appearance for being in love given by males and females.

The female range of 8 is larger than the male range of 5, so females vary much more in how important they think physical appearance is for being in love, with some seeing it as barely important and others as very important. Males are more consistent, suggesting men broadly share a similar opinion that appearance does matter when falling in love, whereas women disagree with each other more about it.

[3]

- 28 (a) Using the formula provided calculate the value of Chi-square for the data in the table below. The E values (expected frequencies) have already been provided (in the table in italics). Show your workings.

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Responses to the question ... 'do you believe in love at first sight?'		
	yes	no
males	5 (8.50)	15 (11.50)
females	12 (8.50)	8 (11.50)

Chi-square = the sum of (O - E) squared divided by E for each of the four cells

Males yes: 5 - 8.5 = -3.5, squared = 12.25, 12.25 / 8.5 = 1.441

Males no: 15 - 11.5 = 3.5, squared = 12.25, 12.25 / 11.5 = 1.065

Females yes: 12 - 8.5 = 3.5, squared = 12.25, 12.25 / 8.5 = 1.441

Females no: 8 - 11.5 = -3.5, squared = 12.25, 12.25 / 11.5 = 1.065

Sum of the four values = 1.441 + 1.065 + 1.441 + 1.065 = 5.012

Chi-square = 5.01 (to two decimal places)

[5]

- (b) Using the extract of the tables of critical values for the Chi-square test presented below, what is the critical value at the 5% probability level for data collected in this study?

	Probability level		
Degrees of freedom (df)	0.5	0.05	0.01
1	0.455	3.841	6.635
2	1.386	5.991	9.210
3	2.366	7.815	11.345
4	3.357	9.488	13.277
5	4.351	11.070	15.086

Degrees of freedom = (rows - 1) x (columns - 1) = (2 - 1) x (2 - 1) = 1

At df = 1 and the 0.05 (5%) probability level the critical value is 3.841.

..... [2]

- (c) Write the significance statement for the analysis performed on this data using the Chi-square test.

Chi-square = 5.01, df = 1, $p < 0.05$. The result is significant at the 5% level because

the calculated value of 5.01 is greater than the critical value of 3.841, so the null

hypothesis is rejected.

..... [2]

- 29 Using the data presented in the pie chart in Figure 1 on page 12, calculate the ratio of how many people said that personality was the most important thing for love compared to those that said wealth was. Show your workings.

Total number of people = 20 males + 20 females = 40

Personality: 37.5% of 40 = $37.5 / 100 \times 40 = 15$ people

Wealth: 12.5% of 40 = $12.5 / 100 \times 40 = 5$ people

Ratio of personality to wealth = 15:5

Dividing both sides by 5 gives the simplified ratio 3:1

[4]

- 30 Evaluate the population validity of the data collected in this study.

Population validity is the extent to which findings about love can be generalised to people who did not take part. A strength is that the sample is balanced, with 20 males and 20 females, so the findings about physical appearance and love at first sight are not biased towards one gender and can be applied to both men and women more confidently. However, the sample is small, with only 40 people, so the ratings of physical appearance and the 37.5% who chose personality may not represent the views of the wider population about love. The sample is also biased because the questionnaire was posted out and only the first twenty of each gender to reply were used, which is self-selecting. People who are currently in love or recently heartbroken are more likely to bother returning a questionnaire about love, so their ratings may be more extreme than the population's, lowering population validity.

[6]

- 31 The discussion section of the write-up of a practical report includes a conclusion made from the analysis of the data collected. Outline **one** conclusion from the discussion section of any of your own practical activities.

In my practical replicating Loftus and Palmer, participants who were asked how fast the cars were going when they 'smashed' gave higher speed estimates than those asked when they 'hit'. I concluded that the wording of a leading question alters how an eyewitness reconstructs a memory, because the verb 'smashed' implies a more violent crash, and that police should therefore avoid leading language when interviewing witnesses.

[3]

END OF QUESTION PAPER

This image shows a blank sheet of white paper designed for handwriting practice. It features a solid black vertical line on the left side, creating a narrow margin. The rest of the page is filled with evenly spaced, horizontal dashed lines for writing. There are no other markings, text, or illustrations on the page.

stemup.co.uk