ALevel Psychology Paper 2

Research Methods 1 Past Questions and Mark Scheme

Name:

Class:

Time:

Marks:

Comments:

1 Read the item and then answer the questions that follow.

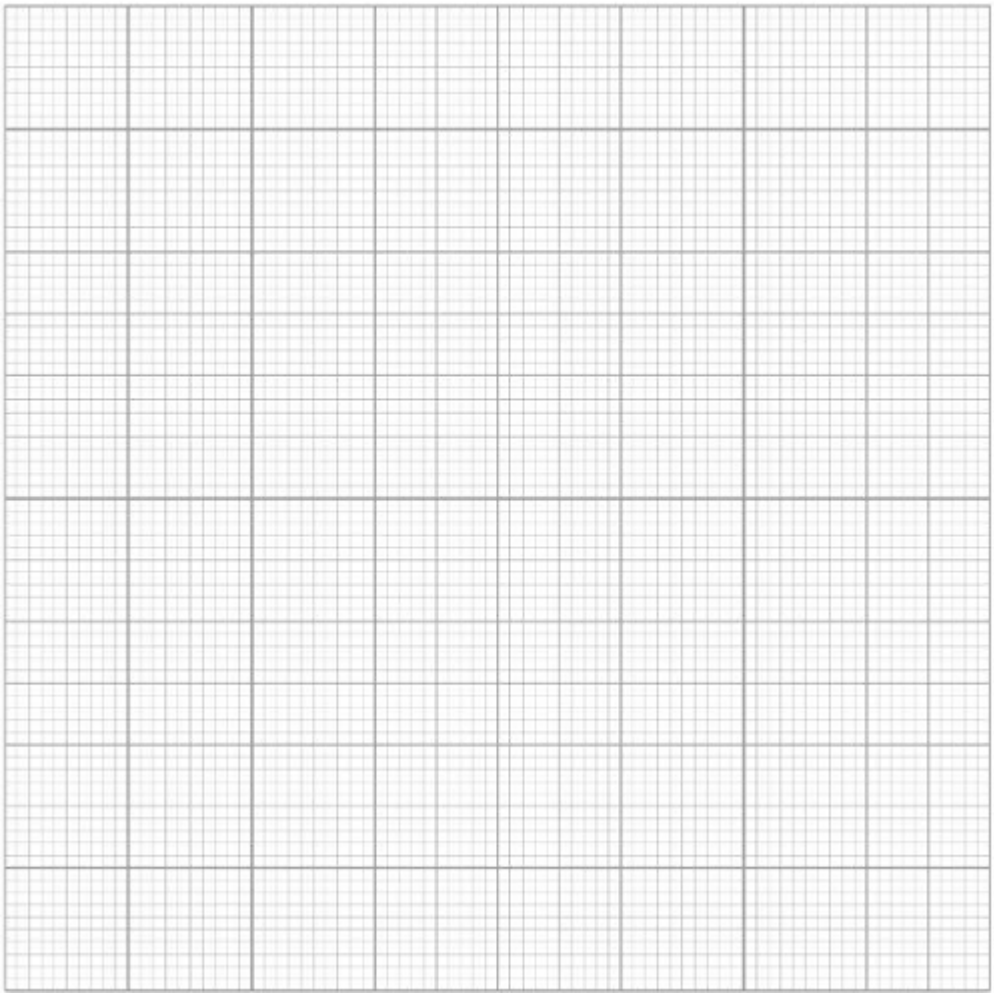
Participants in an experiment were shown a film of a robbery. The participants were then divided into two groups. One group was interviewed using a standard interview technique and the other group was interviewed using the cognitive interview technique. All participants were then given an 'accuracy score' (out of 20) based on how closely their recall matched the events in the film (20 = completely accurate, 0 = not at all accurate).

The results of the experiment are shown in the table below.

The median accuracy score for the standard interview and the cognitive interview

	Standard interview	**Cognitive interview**
Median	10	15

(a) Sketch an appropriate graphical display to show the median accuracy scores in the table above.



(6)

(b) The experiment used an independent groups design.

Explain how this study could have been modified by using a matched pairs design.

(4)
(Total 10 marks)

2 Read the item and then answer the question that follows.

> Studies of attachment often involve observation of interactions between mother and baby pairs like Tasneem and Aisha. Researchers sometimes write down everything that happens as it takes place, including their own interpretation of the events.

Explain how such observational research might be refined through the use of behavioural categories.

(Total 4 marks)

3 Read the item and then answer the questions that follow.

A psychologist wanted to see if creativity is affected by the presence of other people. To test this he arranged for 30 people to participate in a study that involved generating ideas for raising funds for a local youth club. Participants were randomly allocated to one of two conditions.

Condition A: there were 15 participants in this condition. Each participant was placed separately in a room and was given 40 minutes to think of as many ideas as possible for raising funds for a local youth club. The participant was told to write down his or her ideas and these were collected in by the psychologist at the end of the 40 minutes.

Condition B: there were 15 participants in this condition. The participants were randomly allocated to 5 groups of equal size. Each group was given 40 minutes to think of as many ideas as possible for raising funds for a local youth club. Each group was told to write down their ideas and these were collected by the psychologist at the end of the 40 minutes.

The psychologist counted the number of ideas generated by the participants in both conditions and calculated the total number of ideas for each condition.

Total number of ideas generated in Condition A (when working alone) and in Condition B (when working in a group)

	Condition A Working alone	**Condition B Working in a group**
Total number of ideas generated	110	75

(a) Identify the experimental design used in this study **and** outline **one** advantage of this experimental design.

_____

_____

_____

_____

_____

_____

(3)

(b) Describe **one other** experimental design that researchers use in psychology.

_____

_____

_____

_____

(2)

(c) Apart from using random allocation, suggest **one** way in which the psychologist might have improved this study by controlling for the effects of extraneous variables. Justify your answer.

_____

_____

_____

_____

(2)

(d) Write a suitable hypothesis for this study.

_____

_____

_____

_____

_____

_____

(3)

(e) From the information given in the description, calculate the number of participants in each group in Condition B.

_____

_____

(1)

Read the item and then answer the questions that follow.

The psychologist noticed that the number of ideas generated by each of the individual participants in **Condition A** varied enormously whereas there was little variation in performance between the 5 groups in **Condition B.** He decided to calculate a measure of dispersion for each condition.

(f) Name a measure of dispersion the psychologist could use.

_____

_____

(1)

(g) The psychologist uses the measure of dispersion you have named in your answer to **question (f)**. State how the result for each condition would differ.

_____

_____

(1)

(h) Explain how the psychologist could have used random allocation to assign the 15 participants in **Condition B** into the 5 groups.

_____

_____

_____

_____

_____

_____

(3)

(i) Using the information given in the table above, explain how the psychologist could further analyse the data using percentages.

_____

_____

_____

_____

(2)

(j) At the end of the study the psychologist debriefed each participant. Write a debriefing that the psychologist could read out to the participants in **Condition A.**

Extra space

(6)
(Total 24 marks)

4 Read the item and then answer the questions that follow.

> In a study of androgyny, a group of 100 18-year-old students completed a self-report sex-role inventory. The inventory gave two sets of scores: a femininity score and a masculinity score. Each set of scores was on a scale of 0–20, with 0 representing no masculinity or no femininity and 20 representing extreme masculinity or extreme femininity.
>
> The researchers calculated measures of central tendency for the masculinity scores. They found that the mean masculinity score was 10.3, the median masculinity score was 9.5 and the mode masculinity score was 7.

(a) Sketch a graph to show the most likely distribution curve for the masculinity scores in this study. Label the axes of your graph and mark on it the positions of the mean, median and mode.

(3)

(b) What sort of distribution does your graph show?

(1)
(Total 4 marks)

5 The following results are percentages of participants who gave the maximum shock, in variations of Milgram's experiment into obedience to authority.

Condition	% Participants obeying
Experimenter and two obedient confederates are in the same room as the participant.	92.5%
Experimenter is in the same room as the participant.	65%
Experimenter is in a different room from the participant.	20.5%
Experimenter and two disobedient confederates are in the same room as the participant.	10%

What do these results suggest about the power of the confederates in variations of Milgram's study?

Extra space _____

(Total 4 marks)

6 Outline **one** strength and **one** weakness of using correlations in stress research.

Extra space

(Total 4 marks)

7 A psychologist carried out a research study to investigate the effects of institutional care. To do this, she constructed a questionnaire to use with 100 adults who had spent some time in an institution when they were children.

She also carried out interviews with ten of the adults.

(a) For this study, explain **one** advantage of collecting information using a questionnaire.

(3)

(b) In this study, the psychologist collected some qualitative data. Explain what is meant by qualitative data.

_____

_____

_____

_____

(2)

(c) Write **one** suitable question that could be used in the interviews to produce qualitative data.

_____

_____

_____

_____

(2)

(d) Identify **two** ethical issues that the psychologist would need to consider in this research.

Explain how the psychologist could deal with **one** of these issues.

Ethical Issue 1 _____

_____

Ethical Issue 2 _____

_____

How the psychologist could deal with one of these issues _____

_____

_____

_____

_____

_____

_____

(5)
(Total 12 marks)

8 A researcher studied the effect of context on memory. He used an independent groups design. He tested participants in one of two conditions.

In **Condition 1**, a group of 20 schoolchildren learned a list of 40 words in a classroom. This group then recalled the words in the same classroom.

In **Condition 2**, a different group of 20 schoolchildren learned the same list of 40 words in a classroom. This group then recalled the words in the school hall.

The researcher recorded the results and compared the mean number of words recalled in each condition.

(a) Identify the independent variable in this study.

(1)

(b) Use your knowledge of retrieval failure to explain the likely outcome of this study.

(3)

(c) In this study, participants were randomly allocated to one of the two conditions. Explain how this might have been carried out.

(2)

(d) In this study, the researcher used an independent groups design. The researcher decided to repeat the study with different participants and to use a matched pairs design.

Explain how these participants could be matched and then allocated to the conditions.

(2)

(Total 8 marks)

9 The report was subjected to peer review before it was published in a journal.

What is meant by peer review?

(Total 2 marks)

10

Dave, a middle-aged male researcher, approached an adult in a busy street. He asked the adult for directions to the train station. He repeated this with 29 other adults.

Each of the 30 adults was then approached by a second researcher, called Sam, who showed each of them 10 photographs of different middle-aged men, including a photograph of Dave. Sam asked the 30 adults to choose the photograph of the person who had asked them for directions to the train station.

Sam estimated the age of each of the 30 adults and recorded whether each one had correctly chosen the photograph of Dave.

(a) Identify **one** aim of this experiment.

_____

_____

_____

_____

(2)

(b) Suggest **one** reason why the researchers decided to use a field experiment rather than a laboratory experiment. .

_____

_____

_____

_____

(2)

(c) Name the sampling technique used in this experiment. Evaluate the choice of this sampling technique in this experiment.

Sampling technique _____

Evaluation _____

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(4)

(d) Identify **one** possible extraneous variable in this experiment. Explain how this extraneous variable could have affected the results of this experiment.

Extraneous variable _____

How this extraneous variable could have affected the results of this experiment

_____

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(4)
(Total 12 marks)

11

A psychology student was asked to design an investigation to see whether taking exercise could increase feelings of happiness. She proposed to do an experiment. She decided to recruit a sample of volunteers who had just joined a gym, by putting up a poster in the gym. She planned to carry out a short interview with each volunteer and to give each one a happiness score. She intended to interview the volunteers again after they had attended the gym for six weeks and to reassess their happiness score to see if it had changed.

The psychology student's teacher identified a number of limitations of the proposed experiment.

Explain **one or more** limitations of the student's proposal **and** suggest how the investigation could be improved.

(Total 10 marks)

12 A group of researchers conducted a survey about helping behaviour. They asked an opportunity sample of 200 university students to complete a questionnaire. The questionnaire contained open and closed questions. The following are examples of questions used in the questionnaire:

A Do you think that you are generally a helpful person? Yes No

B What do you think most people would do if they were driving in the rain and saw a woman standing alone next to her broken-down car?

C How would you react if someone walking in front of you slipped and fell over?

(a) Identify an open question from **A**, **B** or **C** above. Give **one** advantage of using open questions.

Example of open question (write **A**, **B** or **C**) _____

(1)

Advantage _____

_____

_____

(1)

The researchers then categorised the responses given to question **C** above.
The results are shown in **Table 1**.

Table 1: The number of participants who gave the following responses to question C

Help the person	Ignore the person	Laugh at the person	Other reactions
137	23	31	9

(b) What conclusion might the researchers draw from the responses given in **Table 1** above? Justify your answer.

_____

_____

_____

_____

(2)

On the basis of the responses to question **C**, the researchers decided to conduct a further investigation. The aim was to see whether an individual's helping behaviour might be affected by the presence of other people.

The participants were an opportunity sample of 40 first-year students. The students were told that they would be interviewed about university life. Each student was met by an interviewer and asked to wait. The interviewer then went into the next room. After two minutes there was a loud noise and a cry of pain from the next room.

Twenty participants took part in **Condition 1** and the other 20 participants took part in **Condition 2**.

Condition 1 Each participant waited alone.

Condition 2 Each participant waited with another person who had previously been told by the researchers not to react to the sounds from the next room.

The researchers counted the number of participants in each condition who went to help the interviewer in the next room.

(c) Write a suitable experimental hypothesis for the further investigation.

_____

_____

_____

_____

(2)

(d) Suggest **one** extraneous variable that might be present in the further investigation. Explain why this variable should be controlled and how it could be controlled.

_____

_____

_____

_____

_____

_____

_____

_____

(3)

(e) Identify the experimental design used in the further investigation. Explain why this is a suitable experimental design for this study.

_____

_____

_____

_____

_____

_____

(3)

(f) Explain how random sampling might have been used to select the participants in the further investigation.

_____

_____

_____

_____

(2)

The results of the further investigation are given below.

Table 2: Number of participants who went to help the interviewer in Condition 1 and Condition 2

Condition 1 (Participant waiting alone)	Condition 2 (Participant waiting with another person)
20	9

(g) Suggest a suitable graphical display that could be used to represent the data in **Table 2**. Justify your choice.

_____

_____

_____

_____

(2)

(h) After the further investigation, the researchers debriefed the participants. Discuss **two** points that the researchers should have included when they debriefed the participants.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

(4)
(Total 20 marks)

13 The report was subjected to peer review before it was published in a journal.

Explain why peer review is an important aspect of the scientific process.

(Total 4 marks)

14 Some studies have suggested that there may be a relationship between intelligence and happiness. To investigate this claim, a psychologist used a standardised test to measure intelligence in a sample of 30 children aged 11 years, who were chosen from a local secondary school. He also asked the children to complete a self-report questionnaire designed to measure happiness. The score from the intelligence test was correlated with the score from the happiness questionnaire. The psychologist used a Spearman's rho test to analyse the data. He found that the correlation between intelligence and happiness at age 11 was +0.42.

(a) Write an operationalised non-directional hypothesis for this study.

(2)

(b) Identify an alternative method which could have been used to collect data about happiness in this study. Explain why this method might be better than using a questionnaire.

(4)

(c) A Spearman's rho test was used to analyse the data. Give **two** reasons why this test was used.

(2)

Extract from table of critical values from Spearman's rho(r_s) test

N (number of participants)	Level of significance for a two-tailed test	
	0.10	**0.05**
	Level of significance for a one-tailed test	
	0.05	**0.025**
29	0.312	0.368
30	0.306	0.362
31	0.301	0.356

Calculated r_s must equal or exceed the table (critical) value for significance at the level shown.

(d) The psychologist used a non-directional hypothesis. Using the table above, state whether or not the correlation between intelligence and happiness at age 11 (+0.42) was significant. Explain your answer.

(3)

(e) Five years later, the same young people were asked to complete the intelligence test and the happiness questionnaire for a second time. This time the correlation was –0.29.

With reference to **both** correlation scores, outline what these findings seem to show about the link between intelligence and happiness.

(4)
(Total 15 marks)

15 A researcher believed that there is a biological basis to aggression in males. She predicted that there would be a significant difference between the levels of the hormone testosterone in aggressive males and the levels of the hormone testosterone in non-aggressive males. In order to test her prediction, the researcher statistically analysed the levels of testosterone in saliva samples from 20 aggressive males and 20 non-aggressive males.

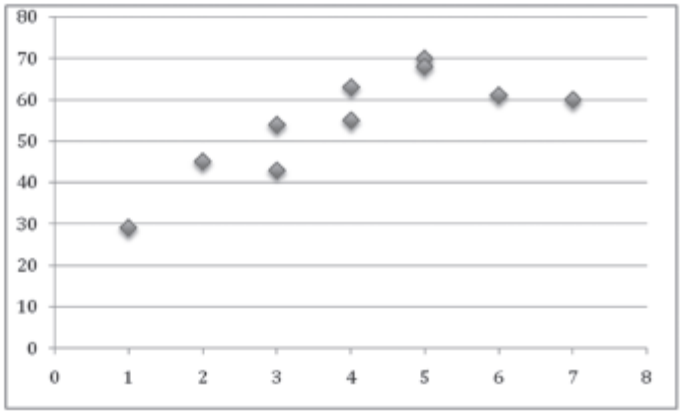
Outline **three** ways in which the study described above could be considered to be scientific.

(Total 3 marks)

16 Research has shown that there is a relationship between stress and illness.
The figure below shows the number of days off work through illness in a year and scores on a stress questionnaire, where a high score indicates more stress.

Relationship between days off work in a year through illness and stress scores

Scores on stress questionnaire



Number of days off work in a year through illness

What does the figure above tell you about the relationship between stress and illness?

_____

_____

_____

_____

Extra space _____

_____

(Total 2 marks)

17 A researcher used content analysis to investigate how the behaviour of young children changed when they started day care.

He identified a group of nine-month-old children who were about to start day care.

He asked the mother of each child to keep a diary recording her child's behaviour every day for two weeks before and for two weeks after the child started day care.

(a) Explain how the researcher could have used content analysis to analyse what the mothers had written in their diaries.

Extra space

(4)

(b) Explain **one or more** possible limitations of this investigation.

Extra space

(4)
(Total 8 marks)

18

A psychologist used an independent groups design to investigate whether or not a cognitive interview was more effective than a standard interview, in recalling information. For this experiment, participants were recruited from an advertisement placed in a local paper. The advertisement informed the participants that they would be watching a film of a violent crime and that they would be interviewed about the content by a male police officer.

The psychologist compared the mean number of items recalled in the cognitive interview with the mean number recalled in the standard interview.

(a) Name the sampling technique used in this experiment.

(1)

(b) Suggest **one** limitation of using this sampling technique.

(2)

(c) Identify the independent variable **and** the dependent variable in this experiment.

Independent variable _____

_____

Dependent variable _____

_____

(2)

(d) Explain **one** advantage of using an independent groups design for this experiment.

_____

_____

_____

_____

(2)

(e) Discuss whether or not the psychologist showed an awareness of the British Psychological Society (BPS) Code of Ethics when recruiting participants for this experiment.

_____

_____

_____

_____

_____

_____

(3)
(Total 10 marks)

19 Briefly outline **two** problems that might arise when making generalisations on the basis of psychological research findings.

(Total 4 marks)

20

(a) One technique used in cognitive interviews is 'report everything'. When using this technique, the police officer in this investigation read the following instructions to the participants:

"Please tell me everything you can remember about what you saw in the film. Do not leave anything out, even the small details you think may be unimportant."

Identify **one other** technique which could have been used by the police officer in this cognitive interview. Write down the instructions that he could have read out to the participants.

Technique _____

_____

Instructions to participants _____

_____

_____

_____

_____

_____

(3)

(b) The psychologist also recorded the number of correct items recalled and the number of incorrect items recalled in each type of interview. The following results were obtained:

	Cognitive Interview	**Standard Interview**
Mean number of correct items recalled	45	32
Mean number of incorrect items recalled	8	8

From these results, what might the psychologist conclude about the effectiveness of cognitive interviews?

_____

_____

_____

_____

(2)
(Total 5 marks)

21

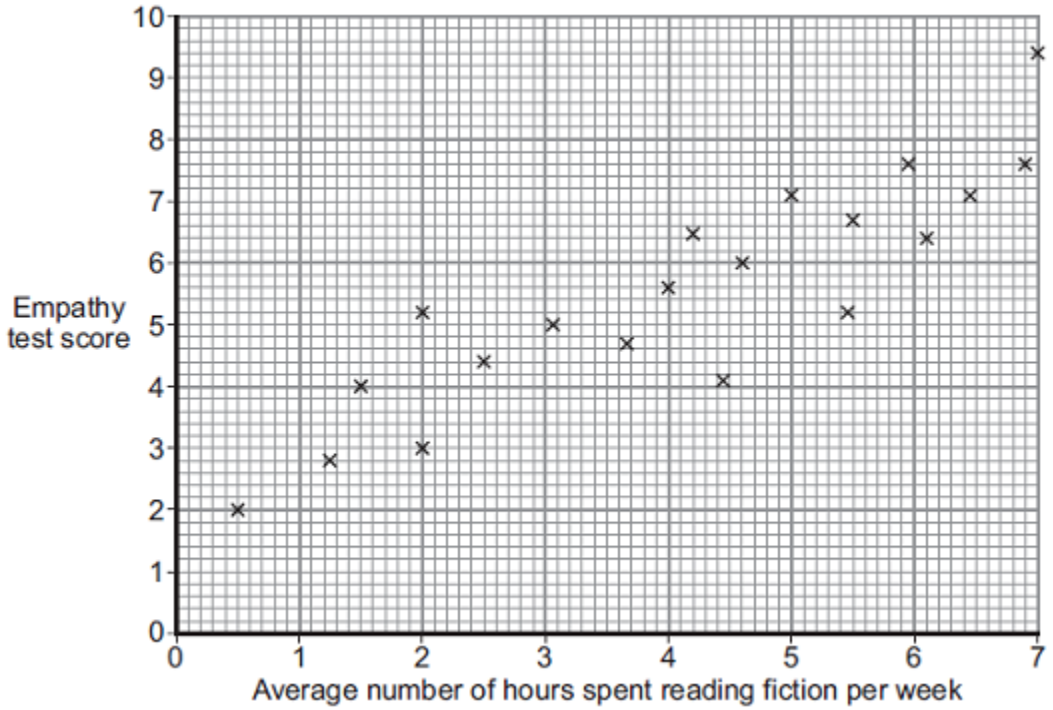
A student teacher was interested in the relationship between empathy (consideration and feelings for others) and the time spent reading fiction. She decided to investigate whether or not such a relationship was present in children.

The student teacher designed her own questionnaire to measure empathy in 8-year-old children. The higher the score achieved, the greater the empathy. Twenty children, all from one school, took part. Each child completed the questionnaire individually.

The student teacher designed another questionnaire to measure 'time spent reading fiction'. Each child was given this questionnaire to take home and complete with his or her parents over a four-week period. 'Time spent reading fiction' included the time spent by parents reading to the child as well as the time the child spent reading independently. Using the responses to this questionnaire, the student teacher calculated how much time per week, on average, each child spent reading fiction.

The data obtained are shown in the graph below.

Scattergram of children's scores on a test of empathy and the average number of hours spent reading fiction per week.



(a) Outline the relationship between empathy and the average number of hours spent reading fiction per week shown in the graph above.

(1)

(b) Name an appropriate test to determine whether or not there is a significant relationship between the two variables in the graph above. Justify your answer with reference to levels of measurement.

(2)

The student teacher decided to use a two-tailed test.

(c) Outline **one** way in which the student teacher could have assessed the validity of the empathy questionnaire.

(2)

(d) Apart from the issue of validity, identify and briefly explain **one** methodological limitation of the study.

(2)

(e) Explain why it was appropriate for the student teacher to use a correlation study rather than an experiment.

(3)

(f) The student teacher noticed that some students on her course commented that they were better able to recall information if they could read the information rather than listen to it in lectures.

Design an experiment to test the following hypothesis:

'People who are given written information will recall more than people who hear information in spoken form.'

In your answer, you should refer to the following and justify your design decisions:

- the variables to be considered
- the experimental design to be used
- the sample
- relevant materials
- an outline of the proposed procedurr.

(8)
(Total 18 marks)

22

A psychodynamic psychologist wished to investigate the function of dreams. He asked five friends to keep a 'dream diary' for a week by writing a descriptive account of their dreams as soon as they woke up in the morning. He interpreted the content of their dreams as an expression of their repressed wishes.

Referring to the study above, explain why psychodynamic psychologists have often been criticised for neglecting the rules of the scientific approach.

(Total 3 marks)

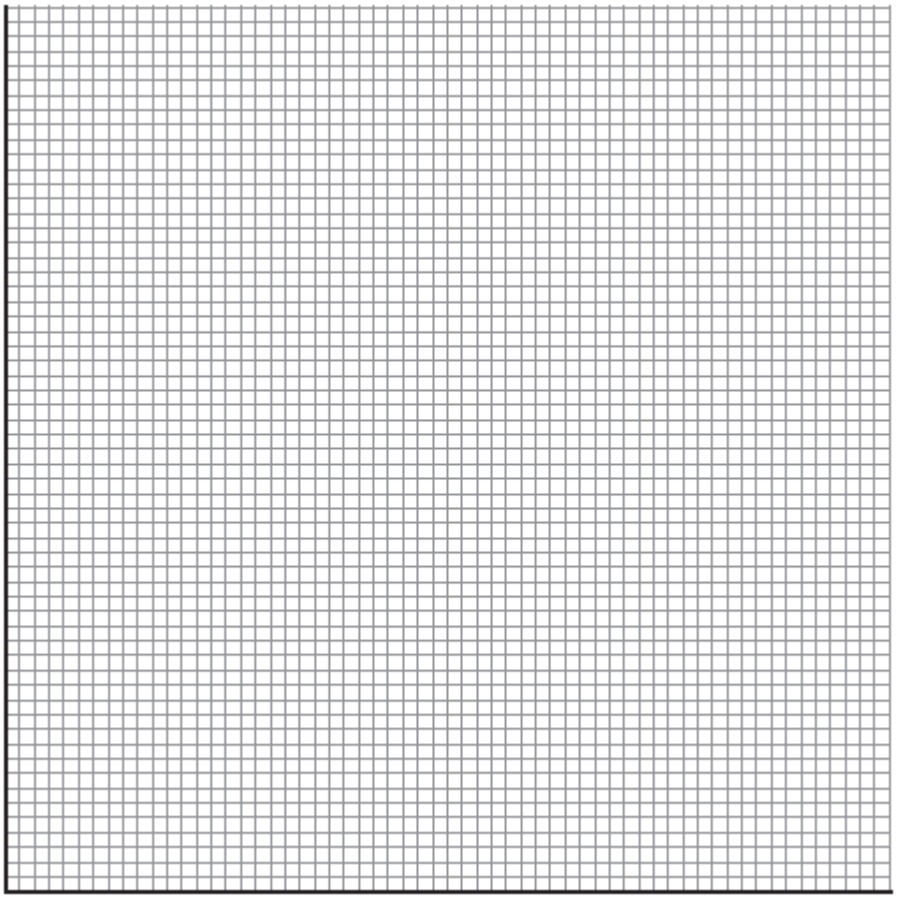
23

Two groups of patients took part in a trial to compare the effectiveness of two different drug therapies. One of the groups was given **Drug A** and the other group was given **Drug B**. All patients completed a rating scale at the start of a ten-week course of treatment and again at the end of the course. This scale measured the severity of symptoms.

The **Drug A** group had an average score of 9 before the therapy and an average score of 4 at the end of the course.

The **Drug B** group had an average score of 7 before the therapy and an average score of 5 at the end of the course.

Sketch and label a bar chart to illustrate the data.



(Total 4 marks)

24

(a) Describe **one** way in which psychologists have investigated caregiver-infant interaction in humans. Refer to a specific study in your answer.

(3)

(b) Evaluate the way of investigating caregiver-infant interaction that you have described in your answer to part (a). Do **not** refer to ethical issues in your answer.

(3)
(Total 6 marks)

25

A psychologist wanted to see whether or not there is a difference in the expectations that men and women have of their own numeracy skills. She obtained a sample of 15 men and 15 women from a factory. She conducted her study in two parts.

In the **first part** of the study, the psychologist said to each participant: "I want you to estimate how many marks you think you will get on a maths test that is suitable for 14-year-old children. If the test has a maximum score of 50, what mark do you think you will get?"

The psychologist recorded the estimate given by each participant and calculated the median estimates for the men and for the women.

The results of the study are given in **Table 1**.

Table 1: Median estimated maths test scores for men and women

	Median estimated maths test score
Men	31
Women	19

(a) Explain how a median score is calculated.

_____

_____

(1)

(b) Identify the dependent variable in this study.

_____

(1)

(c) Write a suitable hypothesis for this study.

_____

_____

_____

_____

(2)

(d) Identify and explain the experimental design used in this study.

_____

_____

_____

_____

(2)

(e) Explain how the psychologist could have obtained a random sample of 15 men and a random sample of 15 women for this study.

_____

_____

_____

_____

_____

_____

_____

_____

(3)

(f) What conclusion could the psychologist draw from the median estimated scores in **Table 1**? Justify your answer.

_____

_____

_____

_____

_____

(2)

In the **second part** of the study, each participant took a 30-minute maths test suitable for 14-year-old children. The test took place under examination conditions. The psychologist marked the test. The maximum mark was 50.

The results of the maths test are given in **Table 2**.

Table 2: Median maths test scores for men and women

	Median maths test score
Men	25
Women	25

(g) Taking the results from **both** parts of the study (Table 1 and Table 2), what can the psychologist now conclude?

_____

_____

_____

_____

_____

_____

(3)

(h) After both parts of the study had been completed, the psychologist needed to debrief the participants.

Write a debrief that the psychologist could read out to the participants.

(4)

(i) This psychologist did not conduct a pilot study. Explain **one** reason why psychologists sometimes conduct pilot studies.

(2)
(Total 20 marks)

26

A researcher investigated whether people with obsessive-compulsive disorder (OCD) are more aware of their own heartbeat than people who do not have OCD. A matched pairs design was used. This involved 10 people with OCD and 10 people without OCD. The researcher asked each participant to estimate how fast his or her heart was beating (in beats per minute) and this was compared to his or her actual heartbeat. It was found that people with OCD were more accurate at estimating their own heartbeat than people without OCD.

(a) Identify the independent variable in this study.

(1)

(b) This study is a quasi-experiment. Explain why this study is a quasi-experiment.

(2)

(c) The researcher used a matched pairs design. Identify **one** relevant variable that could have been used to match participants in this study.

(1)

(d) Outline **one** advantage of using a matched pairs design in this study.

(2)

(Total 6 marks)

27

In an observational study, 100 cars were fitted with video cameras to record the driver's behaviour. Two psychologists used content analysis to analyse the data from the films. They found that 75% of accidents involved a lack of attention by the driver. The most common distractions were using a hands-free phone or talking to a passenger. Other distractions included looking at the scenery, smoking, eating, personal grooming and trying to reach something within the car.

(a) What is content analysis?

(2)

(b) Explain how the psychologists might have carried out content analysis to analyse the film clips of driver behaviour.

(4)

(c) Explain how the two psychologists might have assessed the reliability of their content analysis.

The psychologists then designed an experiment to test the effects of using a hands-free phone on drivers' attention. They recruited a sample of 30 experienced police drivers and asked them to take part in two computer-simulated driving tests. Both tests involved watching a three-minute film of a road. Participants were instructed to click the mouse as quickly as possible, when a potential hazard (such as a car pulling out ahead) was spotted.

Each participant completed two computer-simulated driving tests:

- Test A, whilst chatting with one of the psychologists on a hands-free phone
- Test B, in silence, with no distractions.

The order in which they completed the computer tests was counterbalanced.

(3)

(d) Explain why the psychologists chose to use a repeated measures design in this experiment.

(3)

(e) Identify **one** possible extraneous variable in this experiment. Explain how this variable may have influenced the results of this experiment.

(3)

(f) Explain **one or more** ethical issues that the psychologists should have considered in this experiment.

(4)

(g) Write a set of standardised instructions that would be suitable to read out to participants, before they carry out Test A, chatting on a hands-free phone.

The computer simulator measured two aspects of driver behaviour:

- the number of hazards detected by each driver
- the time taken to respond to each hazard, in seconds.

The mean scores for each of these measures is shown in the table below.

Table to show the mean number of hazards detected and mean reaction times in seconds for Test A and Test B

Mean scores	**Test A: with hands-free phone**	**Test B: in silence**
Number of hazards detected	26.0	23.0
Reaction time in seconds	0.45	0.27

The psychologists then used an inferential statistical test to assess whether there was a difference in the two conditions.

(5)

(h) Identify an appropriate statistical test to analyse the difference in the number of hazards detected in the two conditions of this experiment. Explain why this test of difference would be appropriate.

They found no significant difference in the number of hazards detected ($p > 0.05$), but there was a significant difference in reaction times (p . 0.01).

(3)

(i) Explain why the psychologists did not think that they had made a Type 1 error in relation to the difference in reaction times.

(2)

(j) Replication is one feature of the scientific method. The psychologists decided to replicate this experiment using a larger sample of 250 inexperienced drivers.

Explain why replication of this study would be useful.

(3)
(Total 32 marks)

28 Type A personality can be measured by using a questionnaire. Explain **two** strengths of using questionnaires.

Strength One _____

_____

_____

_____

Extra space _____

_____

Strength Two _____

_____

_____

_____

Extra space _____

_____

(Total 4 marks)

29 A psychologist was interested in the role of sensation-seeking in the development of addictive behaviour. She tested ten participants addicted to smoking (Group A) and ten participants who had no addictive behaviours (Group B). Each participant was given a questionnaire that measured sensation-seeking. Scores on the questionnaire are given in the table below:

Sensation seeking scores for those with addictive behaviours and for those with no addictions

Group A (Addicted to smoking)	Score on sensation-seeking questionnaire	Group B (No addictive behaviours)	Score on sensation-seeking questionnaire
1	65	1	16
2	32	2	25
3	25	3	27
4	29	4	24
5	28	5	59
6	30	6	26
7	18	7	33
8	30	8	21
9	35	9	18
10	28	10	23
Median		**Median**	

Complete the table by calculating the median and range for the two groups. Why did the psychologist use the median rather than the mode?

(Total 4 marks)

30 A researcher investigated obedience. The table shows the percentages of people who obeyed a simple request from a confederate who was either smartly dressed or casually dressed.

Request	**Smartly dressed confederate**	**Casually dressed confederate**
Pick up some litter	80%	61%
Post a letter lying near a post box	61%	40%
Carry a heavy box up some stairs	30%	30%

What do these results suggest about obedience?

Extra space

(Total 4 marks)

31

Social influence research helps us to understand how it is possible to change people's behaviour: for example, understanding how to persuade people to eat more healthily.

With reference to this example of social change, explain how psychology might affect the economy.

_____

_____

_____

_____

_____

_____

_____

_____

(Total 4 marks)

32

A psychologist wanted to investigate the effects of age of adoption on aggressive behaviour. He compared children who had been adopted before the age of two with children who had been adopted after the age of two. The children were observed in their school playground when they were six years old.

(a) Suggest **two** operationalised behavioural categories the psychologist could use in his observation of aggressive behaviour. Explain how the psychologist could have carried out this observation.

Behavioural Category 1 _____

_____

Behavioural Category 2 _____

_____

Explanation of how the observation could have been carried out _____

_____

_____

_____

_____

(4)

(b) Explain **one** ethical issue the psychologist would have needed to consider when carrying out this research. How could the psychologist have dealt with this issue?

Extra space

(4)

The psychologist wanted to investigate how aggressive the children were when they were at home. He interviewed a sample of their parents to investigate this.

(c) Explain why using interviews might be better than using questionnaires in this situation.

Extra space

(4)
(Total 12 marks)

33 Prison staff compared two methods of managing anger in offenders. One group of offenders took part in cognitive therapy. Another group of offenders took part in discussion therapy.

After one month following the training, levels of anger for each individual were rated by prison staff on a scale of 0 – 100. The results are given in the table below:

Ratings of anger in offenders given either systematic CBT anger management training or general advice

Cognitive Group	Anger rating	Discussion group	Anger rating
1	37	1	44
2	45	2	22
3	23	3	74
4	17	4	36
5	41	5	66
6	32	6	63
7	27	7	44
8	26	8	81
9	38	9	56
10	52	10	45
Median		**Median**	

Complete the table by calculating the median for the two groups. Show your working.
Why did the psychologist use the median as a measure of central tendency rather than the mean?

(Total 4 marks)

34 A researcher investigated whether memory for words presented with pictures was better than memory for words presented without pictures. The researcher used an independent groups design.

In **Condition 1**, participants were given a limited time to learn a list of 20 words. They were then asked to recall the 20 words in any order.

In **Condition 2**, participants were given the same time to learn the same 20 words, but this time each word was presented with a picture. For example, the word ‘apple’ was presented alongside a picture of an apple. They were then asked to recall the 20 words in any order.

(a) A pilot study is a small-scale investigation carried out before the main study.

Explain why it would be appropriate for this researcher to use a pilot study. In your answer you must refer to details of the experiment given above.

Extra space

(4)

(b) State a non-directional hypothesis for this experiment.

(2)

(c) Explain **two** reasons why it was more appropriate to use an independent groups design than a repeated measures design.

Reason 1 _____

_____

_____

_____

Reason 2 _____

_____

_____

_____

(4)

The range and median number of words correctly recalled for participants shown words without pictures and for participants shown words with pictures

	Condition 1 Words without pictures	**Condition 2 Words with pictures**
Median number of words correctly recalled	13	16
Range	11	13

(d) What do the scores in the table above show?

_____

_____

_____

_____

(2)

After he had carried out the experiment, the researcher noticed that one participant in **Condition 1** had recalled all 20 words. The researcher thought that this participant might have used a strategy for memory improvement, even though he had not been told to do so.

(Total 12 marks)

35 A psychologist was interested in the effects of violent computer games on aggression in young boys. Following appropriate ethical procedures she set up a study in which she identified ten boys who played violent computer games for at least two hours a day (Group A), and another group of ten boys who did not play violent computer games (Group B). The boys were systematically observed in their school playground on five separate occasions and the total number of aggressive behaviours they demonstrated was recorded. The data are given in the table below:

The effects of playing violent computer games on aggressive behaviour in boys

Group A	Number of aggressive acts	Group B	Number of aggressive acts
1	19	1	8
2	9	2	7
3	3	3	11
4	18	4	7
5	13	5	6
6	16	6	24
7	5	7	9
8	3	8	10
9	7	9	5
10	35	10	10
Median		**Median**	

Complete the table by calculating the median for the two groups. Why did the psychologist use the median as a measure of central tendency rather than the mean?

(Total 4 marks)

36 Distinguish between a Type I error and a Type II error.

(Total 4 marks)

37 A researcher investigated the effectiveness of typical and atypical psychotics in schizophrenia patients with either negative or positive symptoms.

Percentages of patients with either negative or positive symptoms, responding well to typical or atypical antipsychotics.

	Number of patients responding well to atypical antipsychotics	Number of patients responding well to typical antipsychotics
Patients with negative symptoms	30	16
Patients with positive symptoms	60	60

What does the data in the table seem to show about the effectiveness of typical and atypical antipsychotics in the treatment of schizophrenia?

(Total 4 marks)

38 Read the item and then answer the questions that follow.

In a study of antisocial activity and social background, researchers interviewed 100 children aged 14 years. They then classified each child according to their level of antisocial activity. They concluded that 26 were 'very antisocial', 40 were 'mildly antisocial' and 34 were 'not antisocial'. The researchers found that the majority of the 'very antisocial' children attended Crayford secondary school, whereas most of the other two groups of children attended another local school.

(a) The study on the opposite page is an example of socially sensitive research.

Briefly explain how the researchers could have dealt with the issue of social sensitivity in this study.

(4)

(b) What level of measurement is being used in this study?

(1)

(c) Explain **one** limitation of the level of measurement you have identified in your answer to (b).

(2)

(Total 7 marks)

39

Briefly explain **one** reason why it is important for research to be replicated.

_____

_____

_____

_____

(Total 2 marks)

40

Imagine you have been asked to design a study to investigate possible gender differences in card sorting behaviours. You decide you will ask participants to sort a shuffled pack of playing cards into their suits of hearts, clubs, diamonds and spades. You decide you will time the participants as they do this using a stop watch.

Discuss the following aspects of this investigation:

- with reference to the card sorting task, explain how you would ensure that this is made the same task for all participants
- **one** methodological issue you should take into account when obtaining suitable participants for this study and explain how you would deal with this issue
- how you would ensure that the experience of your participants is ethical.

(Total 9 marks)

41

Researchers studying male and female map reading ability calculated the following statistics

	Map reading scores	
	Males	Females
Mean	15.4	5.25
Sd	2.70	2.22

(a) What do the mean and standard deviation values suggest about the male and female performances in the investigation?

(4)

(b) In a replication of the part of the study in which map reading skills were investigated, 20 men and 20 women completed the original map reading task and the researchers obtained the following data:

Male map reading scores	17, 20, 13, 12, 13, 11, 8, 17, 12, 15, 14, 18, 20, 17, 17, 15, 13, 10, 5, 9.
Female map reading scores	12, 8, 10, 11, 4, 2, 11, 18, 17, 12, 13, 10, 3, 15, 11, 9, 10, 11, 16, 10.

The mean map reading score for both groups together was 12.23.

What percentage of the male group scored above the mean score and what percentage of the female group scored above the mean score? Show your calculations.

(4)

(c) Using your answers to both **question (a)** and **question (b)**, comment on the performances of the male and the female participants in this study.

_____

_____

_____

_____

(2)
(Total 10 marks)

42

A psychologist wanted to test the effects of biological rhythms on the ability to solve maths problems. She used random sampling to form two groups each of 20 students.

She tested one group on one set of maths problems at 3 am in the morning. The other group were tested on another set of maths problems at 3 pm in the afternoon. She found that performance of the group tested at 3 pm was significantly better than the group tested at 3 am.

When submitted for peer review the paper was rejected because of serious design problems.

Explain **one** problem with the design of this study and suggest ways of dealing with this problem.

_____

_____

_____

_____

_____

_____

_____

_____

(Total 4 marks)

43

Read the item and then answer the questions that follow.

Researchers were interested in the spatial awareness skills of motorists. They decided to investigate a possible relationship between different aspects of spatial awareness. Motorists who had between ten and twelve years of driving experience and held a clean driving licence with no penalty points were asked to complete two sets of tasks.

Set 1: To follow a series of instructions and using a map, to identify various locations correctly. This provided a map reading score for each motorist with a maximum score of 20.

Set 2: To complete a series of practical driving tasks accurately. This involved tasks such as driving between cones, driving within lines and parking inside designated spaces. Each motorist was observed completing the **Set 2** tasks by a single trained observer who rated each performance by giving the driver a rating out of 10.

The following results were obtained.

The map reading scores and driver ratings of motorists

Participant driver	Map reading score	Driver rating
1	17	9
2	8	4
3	15	7
4	12	6
5	3	2
6	4	4
7	6	8
8	14	6
9	19	10

(a) Should the hypothesis be directional? Explain your answer.

_____

_____

_____

_____

_____

(2)

(b) Write a suitable hypothesis for this investigation.

(3)

(c) Identify a suitable graphical display for the data in the table and briefly explain why this display would be appropriate.

(2)

(d) Using the data in the table, comment on the relationship between the map reading scores and the driver rating scores of the participants.

(3)

(e) Briefly outline **one** problem of using a single trained observer to rate the participants' driving skills in the practical task. Briefly discuss how this data collection method could be modified to improve the reliability of the data collected.

(6)

(f) The researchers decided to analyse the data using a Spearman's rho test. Explain why this is a suitable choice of test for this investigation.

(3)

Table of critical values for a Spearman's rho test

		Level of significance for a two-tailed test			
		0.10	0.05	0.02	0.01
		Level of significance for a two-tailed test			
		0.05	0.025	0.01	0.005
N=	8	0.643	0.738	0.833	0.881
	9	0.600	0.700	0.783	0.833
	10	0.564	0.648	0.745	0.794

Calculated r_S must EQUAL or EXCEED the critical value for significance at the level shown.

(g) After analysis of the data the researchers obtained a calculated value of **r_S= 0.808.**

Using the information in the table above, what conclusion can the researchers draw about the relationship between the map reading and driving skills of the motorists? Explain your answer.

(4)
(Total 23 marks)

44 Explain **one** limitation of asking parents to rate their own children.

(Total 2 marks)

45

A researcher wanted to see whether cognitive behaviour therapy was an effective treatment for depression. Twenty depressed patients who had all recently completed a course of cognitive behaviour therapy were involved in the investigation. From their employment records, the researcher kept a record of the number of absences from work each patient had in the year following their treatment. This was compared with the number of absences from work each patient had in the year prior to their treatment.

Those patients who had fewer absences from work in the year following their treatment than in the year prior to their treatment were classified as ‘improved’ (+). Those patients who had more absences were classified as ‘deteriorated’ (-). Those patients who had the same number of absences were classified as ‘neither’ (0).

The results of the investigation are included in **Table 1** below.

Table 1

Patient	Improved	Deteriorated	Neither
1	+		
2			0
3		–	
4	+		
5	+		
6	+		
7		–	
8		–	
9			0
10	+		
11		–	
12	+		
13	+		
14	+		
15	+		
16		–	
17	+		
18	+		
19	+		
20			0

The researcher decided to use the sign test to analyse the data.

(a) Explain **two** factors that the researcher had to take into account when deciding to use the sign test. Refer to the investigation above in your answer.

(4)

(b) Calculate the sign test value of s for the data in **Table 1**. Explain how you reached your answer.

(2)

Table 2: Critical values for the sign test

n	0.005 (one tailed) 0.01 (two tailed)	0.01 (one tailed) 0.02 (two tailed)	0.025 (one tailed) 0.05 (two tailed)	0.05 (one tailed) 0.10 (two tailed)
16	2	2	3	4
17	2	3	4	4
18	3	3	4	5

For significance, the value of the less frequent sign is equal to, or less than, the value of the table.

(c) With reference to the critical values in **Table 2**, explain whether or not the value of s that you calculated in response to **question (b)** is significant at the 0.05 level for a two tailed test.

(2)

(d) The investigation above is based on secondary data.

In what ways would the use of primary data have improved this investigation?

(3)

(e) Outline the implications of psychological research for the economy. Refer to the investigation above in your answer.

(5)
(Total 16 marks)

46 Briefly explain **one** reason why it is important for research to undergo a peer review process.

(Total 2 marks)

47 Read the item and then answer the questions that follow.

> A psychologist investigating the investment model of relationships, devised a self-report Investment Scale for use with a group of 100 female participants. The scale gave an investment score for each participant on a scale of 0–20, with 0 representing no investment in relationships and 20 representing extreme investment in relationships.
>
> The psychologist calculated measures of central tendency for the investment scores. He found that the mean investment score was 8.6, the median investment score was 9.5 and the mode investment score was 13.

(a) Sketch a graph to show the most likely distribution curve for the investment scores in this study. Label the axes of your graph and mark on it the positions of the mean, median and mode

(3)

(b) What sort of distribution does your graph show?

(1)
(Total 4 marks)

48 Explain **one** limitation of a self-report technique.

(Total 2 marks)

49 Read the item and then answer the question that follows.

> The psychologist focused on fluency in spoken communication in her study. Other research has investigated sex differences in non-verbal behaviours such as body language and gestures.

Design an observation study to investigate sex differences in non-verbal behaviour of males and females when they are giving a presentation to an audience.

In your answer you should provide details of:

- the task for the participants
- the behavioural categories to be used and how the data will be recorded
- how reliability of the data collection might be established
- ethical issues to be considered.

(Total 12 marks)

50 Read the item and then answer the questions that follow.

Researchers used a test to measure the mathematical reasoning ability of pairs of identical and non-identical twins. If both members of a pair had a similar score on the test, they were said to be 'concordant'. This type of study is known as a concordance study.

Outcome of the research with the concordance rates expressed as a percentage

Genetic relationship group	Concordance rate for mathematical reasoning ability
Identical twins (100% shared genes)	58%
Non-identical twins (50% shared genes)	14%

(a) Briefly explain the outcome of the study in relation to the nature-nurture debate.

(2)

(b) Some ways of establishing validity involve the use of a statistical test.

Outline how these researchers could have used a statistical test to establish **concurrent** validity of the mathematical reasoning ability test.

(4)

(Total 6 marks)

51 Read the text below and then answer the questions that follow.

Two researchers obtained a sample of ten people whose ages ranged from 20-years-old to 60-years-old.

Each participant was asked to take part in a discussion of social care issues. This included discussion about who should pay for social care for elderly people and how to deal with people struggling with mental health problems. A confederate of the researchers was given a script to follow in which a series of discussion points was written for the confederate to introduce.

Each participant then came into a room individually and the discussion with the confederate took place. The maximum time allowed for a discussion was 30 minutes.

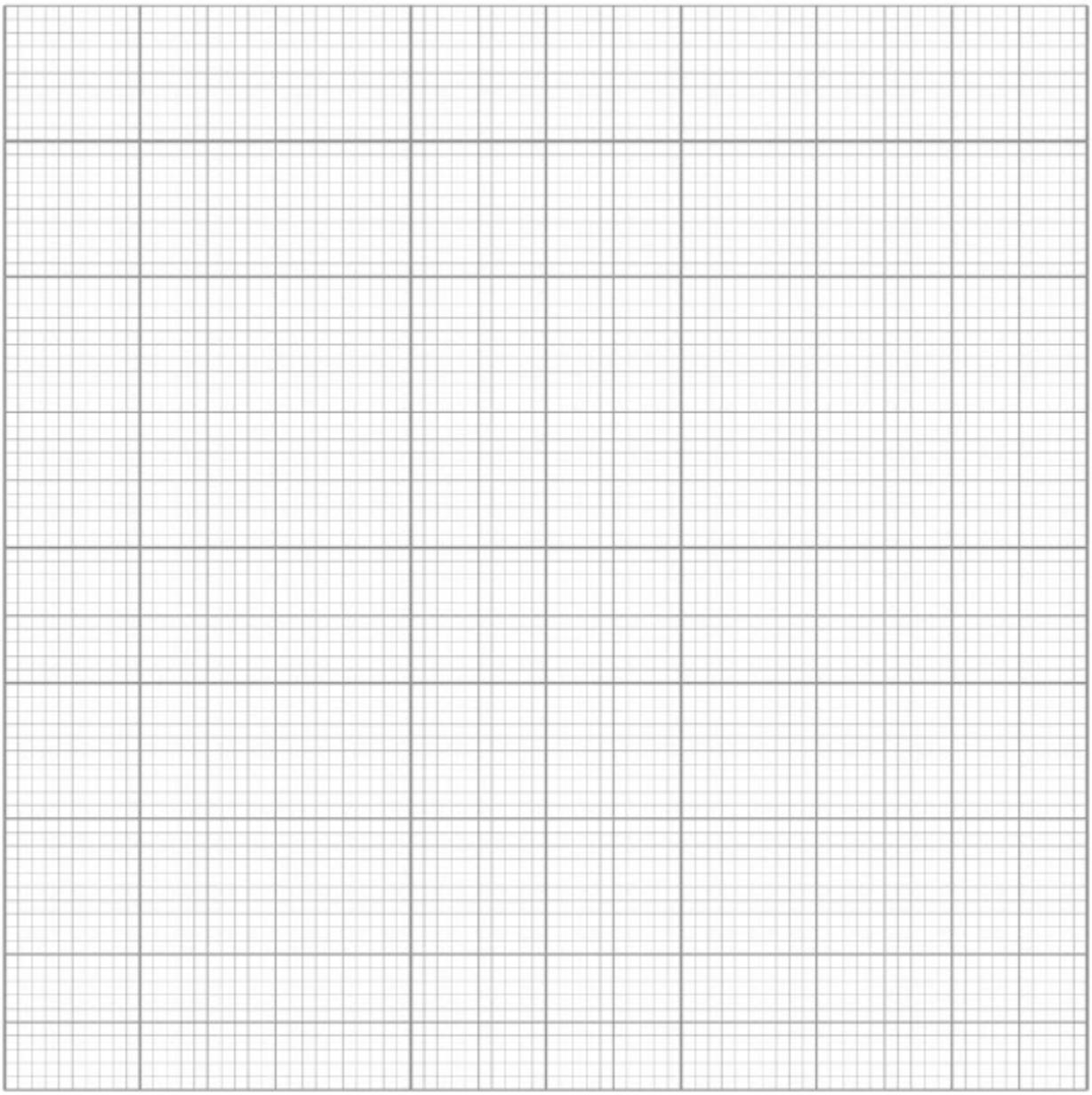
The researchers observed the discussions between the confederate and participants and rated the active engagement of the participants in the discussion. The ratings were between 1, (not at all interested) and 20, (extremely interested.) The researchers believed that the rating provided a measurement of the participants' attitudes towards social care issues.

The following data were obtained in the study:

The relationship between age and attitude to social care.

Age of participant	Attitude to social care issues rating
21	5
23	3
34	8
36	12
40	10
47	13
52	17
53	15
58	18
60	20

(a) Use the graph paper below to sketch a display of the data given in the table above. You do not need to give your display a title.



(3)

(b) What does the display you have drawn in your answer in part (a) suggest about the relationship between age and attitude to social care issues? Explain your answer.

_____

_____

_____

_____

(2)

(c) The researchers rated the active engagement of the participants in the discussion on social care. They used this rating as a measure of each participant's attitude to social care issues.

Briefly explain how investigator effects might have occurred in this study.

(2)

(d) Outline how the researchers could have avoided investigator effects having an impact on the study.

(2)

The researchers thought it might be interesting to investigate further the attitudes of the participants in the study. They decided to interview each participant. The researchers devised a questionnaire in order to collect the data they required. The questionnaire included both open and closed questions.

(e) Briefly discuss the benefits for the researchers of using **both** closed **and** open questions on their questionnaire about attitudes to social care.

(4)

(f) Write **one** question that you think the researchers might have put on their questionnaire. Explain which type of question you have written and why you think this would be a suitable question for this study.

_____

_____

_____

_____

_____

_____

(3)

The researchers have obtained both qualitative and quantitative data in the observations and interviews they have conducted.

(g) Identify the qualitative and quantitative data collected in this study. Explain your answer.

_____

_____

_____

_____

_____

_____

_____

_____

(4)

(h) Explain how the researchers should have addressed **two** ethical issues in the investigation.

_____

_____

_____

_____

_____

_____

_____

_____

(4)
(Total 24 marks)

52 Read the item and then answer the questions that follow.

A psychologist wanted to see if verbal fluency is affected by whether people think they are presenting information to a small group of people or to a large group of people.

The psychologist needed a stratified sample of 20 people. She obtained the sample from a company employing 60 men and 40 women.

The participants were told that they would be placed in a booth where they would read out an article about the life of a famous author to an audience. Participants were also told that the audience would not be present, but would only be able to hear them and would not be able to interact with them.

There were two conditions in the study, **Condition A** and **Condition B**.

Condition A: 10 participants were told the audience consisted of 5 listeners.

Condition B: the other 10 participants were told the audience consisted of 100 listeners.

Each participant completed the study individually. The psychologist recorded each presentation and then counted the number of verbal errors made by each participant.

(a) Identify the dependent variable in this study.

(2)

(b) Write a suitable hypothesis for this study.

(3)

(c) Identify **one** extraneous variable that the psychologist should have controlled in the study **and** explain why it should have been controlled.

(3)

(d) Explain **one** advantage of using a stratified sample of participants in this study.

(2)

(e) Explain how the psychologist would have obtained the male participants for her stratified sample. Show your calculations.

(3)

(f) The psychologist wanted to randomly allocate the 20 people in her stratified sample to the two conditions. She needed an equal number of males in each condition and an equal number of females in each condition. Explain how she would have done this.

(4)

(Total 17 marks)

53

Read the item and then answer the questions that follow.

A child psychologist carried out an overt observation of caregiver-infant interaction. She observed a baby boy interacting separately with each of his parents. Using a time sampling technique, she observed the baby with each parent for 10 minutes. Her findings are shown in the table below

Frequency of each behaviour displayed by the infant when interacting with his mother and when interacting with his father

	Gazing at parent	**Looking away from parent**	**Eyes closed**	**Total**
Mother	12	2	6	20
Father	6	10	4	20
Total	18	12	10	40

(a) Using the data in the table, explain the procedure used for the time sampling technique in this study.

_____

_____

_____

_____

_____

_____

(3)

(b) In what percentage of the total observations was the baby gazing at his mother? Show your calculations.

_____

_____

_____

_____

(2)

(c) Which **one** of the following types of data best describes the data collected in this study? Shade **one** box only.

A Primary data ⬭

B Qualitative data ⬭

C Secondary data ⬭

D Continuous data ⬭

(1)
(Total 6 marks)

54 Read the item and then answer the questions that follow.

A psychologist wanted to see if verbal fluency is affected by whether people think they are presenting information to a small group of people or to a large group of people.

The psychologist needed a stratified sample of 20 people. She obtained the sample from a company employing 60 men and 40 women.

The participants were told that they would be placed in a booth where they would read out an article about the life of a famous author to an audience. Participants were also told that the audience would not be present, but would only be able to hear them and would not be able to interact with them.

There were two conditions in the study, **Condition A** and **Condition B**.

Condition A: 10 participants were told the audience consisted of 5 listeners.

Condition B: the other 10 participants were told the audience consisted of 100 listeners.

Each participant completed the study individually. The psychologist recorded each presentation and then counted the number of verbal errors made by each participant.

The results of the study are given in the table.

Mean number of verbal errors and standard deviations for both conditions

	Condition A (believed audience of 5 listeners)	Condition B (believed audience of 100 listeners)
Mean	11.1	17.2
Standard deviation	1.30	3.54

(a) What conclusions might the psychologist draw from the data in the table? Refer to the means **and** standard deviations in your answer.

(6)

(b) Read the item and then answer the question that follows.

> The psychologist had initially intended to use the range as a measure of dispersion in this study but found that one person in **Condition A** had made an exceptionally low number of verbal errors.

Explain how using the standard deviation rather than the range in this situation, would improve the study.

(3)

(c) Name an appropriate statistical test that could be used to analyse the number of verbal errors in the table. Explain why the test you have chosen would be a suitable test in this case.

(4)

(d) The psychologist found the results were significant at $p<0.05$. What is meant by 'the results were significant at $p<0.05$'?

(2)

(e) Briefly explain one method the psychologist could use to check the validity of the data she collected in this study.

(2)

(Total 17 marks)

55 Briefly discuss how observational research might be improved by conducting observations in a controlled environment.

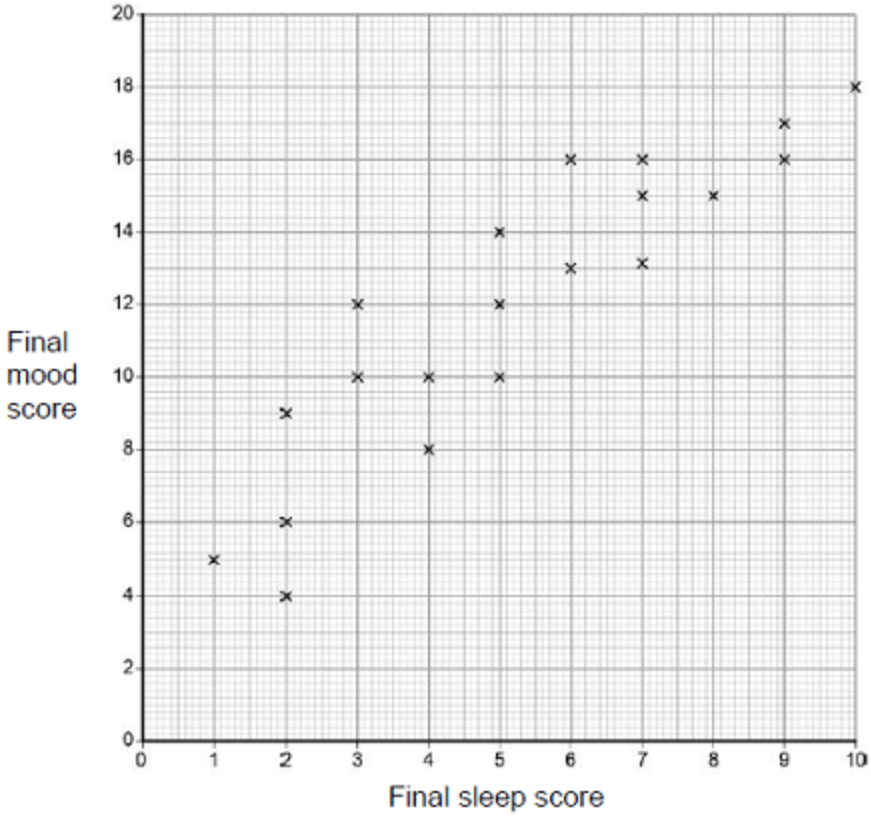
(Total 4 marks)

56 Read the item and then answer the questions that follow.

Twenty depressed patients were treated using cognitive behavioural therapy. Over the course of the six-week treatment, each patient's mood was monitored every week using a self-report mood scale (where a score of 20 = extremely positive mood and a score of 0 = extremely negative mood). Each week they also completed a quality of sleep questionnaire which was scored from 10 = excellent sleep to 0 = very poor sleep.

At the end of the study the researchers correlated each patient's final mood score with his or her final sleep score. The results are shown in the graph below.

Scattergram to show the relationship between final mood scores and final sleep scores for 20 patients at the end of therapy



(a) Outline the type of relationship shown in the graph above and suggest why it would not be appropriate for the researchers to conclude that better sleep improves mood.

_____

_____

_____

_____

(2)

(b) Outline **one** way in which the researchers should have dealt with ethical issues in this study.

_____

_____

_____

_____

(2)

(c) The sleep questionnaire used by the researchers had not been checked to see whether or not it was a reliable measure of sleep quality.

Explain how this study could be modified by checking the sleep questionnaire for test-retest reliability.

(4)
(Total 8 marks)

57 Read the item and then answer the question that follows.

> A group of researchers used 'event sampling' to observe children's friendships over a period of three weeks at break times and lunchtimes during the school day.

Explain what is meant by 'event sampling'.

(Total 2 marks)

58 Explain what is meant by 'overt observation'.

(Total 2 marks)

59 Read the item and then answer the questions that follow.

An experiment was carried out to test the effects of learning similar and dissimilar information on participants' ability to remember.

In **Stage 1** of the experiment, 10 participants in **Group A**, the 'similar' condition, were given a list of 20 place names in the UK. They were given two minutes to learn the list. 10 different participants in **Group B**, the 'dissimilar' condition, were given the same list of 20 place names in the UK. They were also given two minutes to learn the list.

In **Stage 2** of the experiment, participants in **Group A** were given a different list of 20 more place names in the UK, and were given a further two minutes to learn it. Participants in **Group B** were given a list of 20 boys' names, and were given a further two minutes to learn it.

In **Stage 3** of the experiment, all participants were given five minutes to recall as many of the 20 place names in the UK, from the list in **Stage 1**, as they could. The raw data from the two groups is below.

Number of place names recalled from the list in Stage 1

Group A	Group B
5	11
6	10
4	11
7	13
8	12
4	14
5	15
4	11
6	14
7	14

(a) What is the most appropriate measure of central tendency for calculating the average of the scores, from the table, in each of the **two** groups? Justify your answer.

_____

_____

_____

_____

(2)

(b) Calculate the measure of central tendency you have identified in your answer to **part (a)** for **Group A** and **Group B.** Show your calculations for each group.

_____

_____

_____

_____

_____

_____

_____

_____

(4)

(c) In **Stage 3** of the experiment, several participants in **Group A**, the 'similar' condition, recalled words from the **Stage 2** list rather than the **Stage 1** list.

Use your knowledge of forgetting to explain why this may have occurred.

_____

_____

_____

_____

(2)
(Total 8 marks)

60 A cognitive psychologist investigating how memory works gave participants the same word list to recall in one of two conditions. All the words were of equal difficulty.

Condition 1: Ten participants recalled the words in the same room in which they had learned the words.

Condition 2: Ten different participants recalled the words in a room that was not the same room as that in which they had learned the words.

The following results were obtained:

Mean values and standard deviations for Condition 1 and Condition 2 in a memory experiment.

	Condition 1	**Condition 2**
Mean	15.9	10.6
Standard deviation	3.78	1.04

(a) Why are the standard deviation values found in the study above useful descriptive statistics for the cognitive psychologist?

_____

_____

_____

_____

(2)

(b) Outline **one** problem of studying internal mental processes like memory ability by conducting experiments such as that described in part (a) above.

_____

_____

_____

_____

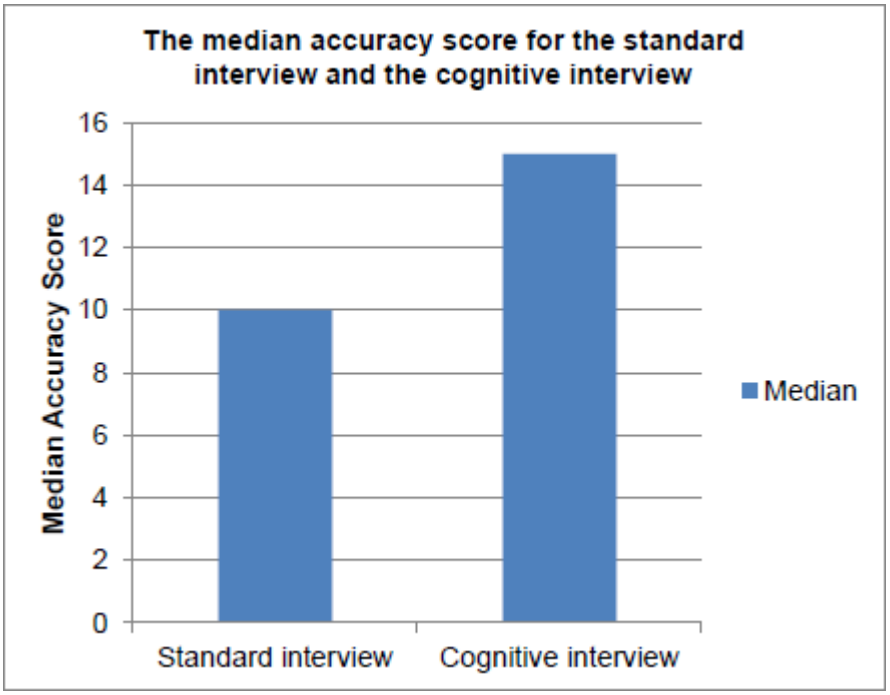
(2)
(Total 4 marks)

1

(a) **[AO2 = 6]**

1 mark for each of the following:

- display as a bar chart
- both axes labelled correctly
- an informative title with reference to the IV and DV
- y axis has appropriate scaling
- bars are separate
- bars are plotted reasonably correctly.



(b) **[AO3 = 4]**

Award **one mark** for each of the following points:

- the researcher needs to ensure that the two groups are matched for key variables
- example of at least one key variable – any that might reasonably be expected to affect memory in this situation, eg eyesight, age, intelligence
- all participants should be pre-tested / assessed for the key variable / variables
- for each person in one condition, the researcher should assign a 'matched' person in the other condition.

Credit other relevant points or this information embedded in the example.

2 **[AO3 = 4]**

Award **1 mark** for any four points explained from the following points, to a maximum of 4 marks:

- behavioural categories allow observers to tally observations into pre-arranged groupings
- examples of behavioural categories appropriate in this situation might be ……
- using categories provides clear focus for the researcher
- categorisation enables proposal of a testable hypothesis
- categories allow for more objective / scientific data recording
- use of categories should result in greater reliability
- categories provide data that is easier to quantify / analyse
- contrast with method described in the stem (own interpretation is too subjective / opinion-based).

Credit other valid points.

3 (a) **[AO1 = 1 and AO3 = 2]**

1 mark for identification of the correct experimental design – independent groups / independent measures.

Plus

2 marks for a clear and coherent outline of an advantage using appropriate terminology.

OR

1 mark for a brief / vague / muddled outline of an advantage.

Possible advantages:

- performances not affected by order effects as people only do one condition
- demand characteristics less likely as participants only aware of own condition
- same task / materials can be used in both conditions as participants are always naïve to the task.

Credit other relevant advantages.

(b) **[AO1 = 2]**

2 marks for a clear and coherent outline of how participants are used in either a repeated measure or a matched pairs design.

1 mark for a vague, muddled or incomplete outline of a repeated measure or a matched pairs design.

If the answer to **(a)** is incorrect, credit a different design to that given.

(c) **[AO3 = 2]**

1 mark for an appropriate and plausible suggestion.

Plus

1 mark for an appropriate justification.

Likely suggestions:

- testing all participants in the same room
- making sure that all participants hear the same instructions
- ensuring that all participants are tested by the same researcher.

Credit other relevant suggestions.

(d) **[AO2 = 3]**

3 marks for an appropriate non-directional (or directional) operationalised hypothesis: 'There is a difference in the number of ideas generated when participants work alone and when they work in groups.'

2 marks for a statement with both conditions of the IV and DV that lacks the clarity or has only one variable operationalised.

1 mark for a muddled statement with both conditions of the IV and DV where neither variable is operationalised.

0 marks for expressions of aim / questions / correlational hypotheses or statements with only one condition.

Full credit can be awarded for a hypothesis expressed in a null form.

(e) **[AO2 = 1]**

1 mark: 3 (in each group)

(f) **[AO1 = 1]**

1 mark for naming a suitable measure of dispersion (range or standard deviation).

(g) **[AO2 = 1]**

1 mark for stating that the statistic calculated (either the range or the SD) would be greater in **Condition A** than in **Condition B**.

or written as

1 mark for stating that the statistic calculated (either the range or the SD) would be less in C**ondition B** than in **Condition A**.

(h) **[AO2 = 3]**

Marks for a clear description of a practical way as follows:

1 mark – all the participants allocated a number from 1 to 15.
1 mark – the 15 numbers are put in a hat.
1 mark – assign first three numbers drawn to a group and repeat process for other 4 groups.

Accept other valid descriptions that would be practical and produce the same outcome.

(i) **[AO3 = 2]**

1 mark : for each condition, the overall number of ideas generated should be divided by the overall total of 185.

Plus

1 mark : the result for each condition should then be multiplied by 100 to give the percentage.

(j) **[AO2 = 6]**

Level	Marks	Description
3	5 – 6	Both elements of required content are clear and mostly well detailed. The debrief is all in verbatim format.
2	3 – 4	Both elements of required content are present. The answer lacks detail and / or clarity in places. Some of the answer is in verbatim format.
1	1 – 2	There is some information about at least one element of required content. The answer lacks clarity. Verbatim format is lacking. For one mark there must be some relevant content, eg an optional point about ethics.
	0	No relevant content.

Required content:

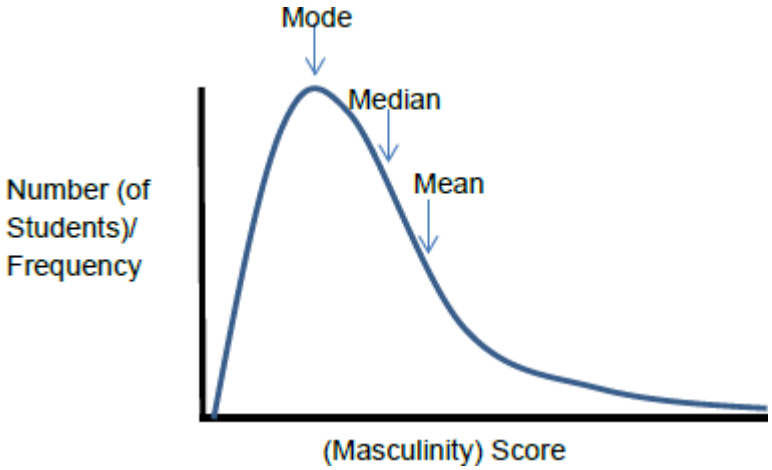
- explanation of the aim: to see if creativity is affected by the presence or absence of others
- information about the other condition – in an independent design people need to know about the condition in which they did not take part.

Optional content:

- specific ethical issues, eg right to withdraw data / be informed of results / check of welfare
- general ethical considerations, eg respect for participants.

4

(a) **[AO2 = 3]**



Credit a rough sketch of a positively skewed distribution as follows:

1 mark for shape of curve with tail to the right.

1 mark for axis labels – '(Masculinity) Score' on horizontal axis, 'Number (of Students)' / 'Frequency' on vertical axis.

1 mark for positioning the mean, median and mode appropriately in relation to one another.

(b) **[AO2 = 1]**

1 mark for stating a positive skew.

If the graph sketched in (a) does not show a positive skew, credit answers that match the sketch given.

5

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

AO3 = 4

The data suggest that the confederates have a considerable influence on whether or not the participant obeys; candidates could consider the implications of the difference between 92.5% and 10%. They might consider whether the confederates are acting as role models, informing the participant how to behave. Credit could also include comparison of power of confederates with power of having the experimenter in the same room.

The question is not just asking candidates to describe the data in the table, but to consider the effect that the confederates have, to access the top bands answers need to be shaped to fit the question.

AO3 Interpretation of data
4 marks Accurate and reasonably detailed Accurate and reasonably detailed answer that demonstrates sound knowledge and understanding of what the data suggest about obedience. There is appropriate selection of material to address the question.
3 marks Less detailed but generally accurate Less detailed but generally accurate answer that demonstrates relevant knowledge and understanding. There is some evidence of selection of material to address the question.
2 marks Basic Basic answer that demonstrates some relevant knowledge and understanding but lacks detail and may be muddled. There is little evidence of selection of material to address the question.
1 mark Very brief/flawed or inappropriate Very brief or flawed answer demonstrating very little knowledge. Selection and presentation of information is largely or wholly inappropriate.
0 marks No creditworthy material.

6

AO3 = 6

Strength: can study relationship between variables that occur naturally. Can measure things that cannot be manipulated experimentally. Can suggest trends that can lead to experiments.
Weakness: It is not possible to say that one thing causes another. Just because there is a correlation between stress scores and days off it does not mean that stress caused people to take days off work, or there may be another variable connecting them. Elaboration through the use of an appropriate example can also receive credit.
Any other appropriate answer can get credit.

One mark for a brief outline of strength / limitation and a further mark for elaboration. For example, cannot say one thing causes another (1 mark) there may be a third variable that connects the two (2nd mark for elaboration.)

7

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

(a) **AO3 = 3**

Advantages of using a questionnaire in this study could include that data from the hundred adults could be collected relatively quickly because the researcher would not need to be present when the questionnaires were completed; participants might be more willing to answer honestly because they would feel more anonymous; there might be a reduction in investigator effects because the researcher's reactions would not be visible. The advantage must be one that could be applied to this study.
1 mark for a slightly muddled or very brief outline of an advantage. Further marks for accurate elaboration.

(b) **AO3 = 2**

Qualitative is non-numerical and uses words to give a full description of what people think or feel.
1 mark for a very brief or slightly muddled answer eg qualitative data uses words.
2nd mark for accurate elaboration eg by comparison or by using an example.

(c) **AO3 = 2**

One mark for a question which would produce qualitative data but is not appropriate eg "How are you feeling?"
Two marks for an appropriate question eg "Tell me what it was like in the institution"
(Full marks can be awarded if it is not in the form of a question)
0 marks for a question that would not produce qualitative data.

(d) **AO3 = 1 + 1 + 3**

There are no ethical issues named in the specification, so any potentially relevant issues should be credited.
Likely ethical issues include informed consent, right to withdraw, protection from harm, confidentiality, respect or the need for debriefing in this particular case.
Other issues such as deception (deliberate or by omission) can be credited as they could apply in this research.
One mark each for identification of a relevant ethical issue.
One mark for a brief mention of how the issue could be dealt with.
Two further marks for elaboration appropriate to this research.
There is a depth / breadth trade-off. Candidates may explain one way of dealing with the issue in some depth, or mention several ways (of dealing with one issue) more briefly.
Ethical issue one eg, right to withdraw (1 mark); ethical issue two eg confidentiality (1 mark); Don't identify the participants (1 mark). Don't use photographs or names in published research. Names of people and / or places should be changed (2 further marks).

8

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

(a) **[AO3 = 1]**

One mark for the independent variable.
Likely answers: the context of recall / whether participants recalled the words in the same room or a different room / the classroom or the school hall.
Reference to both conditions might be implicit rather than clearly stated.

(b) **[AO3 = 1, AO2 = 2]**

AO3

Award one mark for stating the likely outcome.
Likely answers: Participants who learned and recalled in the same context are likely to recall more words than those who learned and recalled in different contexts / there will be a higher mean number of words recalled in Condition 1 than Condition 2.
Accept alternative wording.

AO2

Award up to two marks for explanation of the likely outcome based on knowledge of retrieval failure as an explanation for forgetting. Credit reference to environmental cues / context triggering recall; the absence of cues / context in Condition 2.

For two AO2 marks there must be some reference to condition two's participants failing to retrieve / recall information.
Credit use of evidence and / or use of an example as part of the discussion.

(c) **[AO3 = 2]**

Award up to two marks for an explanation of how random allocation to one of the two conditions might have been carried out. Two marks for a full explanation, one mark for a brief / vague answer.
Possible answer: All participants' names / numbers are placed into a hat / lottery system / computer (1) the first name drawn is assigned to condition one, the next to condition two / the first twenty are allocated to condition one, the second twenty to condition two (1).

(d) **[AO3 = 2]**

Award up to two marks for an explanation of how participants could be matched and then allocated to the two conditions for a matched pairs design.
Possible answer: Participants are paired on some relevant variable (eg memory ability, IQ, age, etc.), (1) and then one from each pair is allocated to each condition (1).
Answers based on the use of identical twins can get full marks as long as there is some reference to the idea that twins are likely to have a similar level of recall.

9

AO1 = 2

Peer review is the process of subjecting a piece of research to independent scrutiny by other psychologists working in a similar field who consider the research in terms of its validity, significance and originality.

0 marks for 'other psychologists look at the research'.

1 mark for a very brief outline eg 'other psychologists look at the research report before it is published'. Award one further mark for elaboration.

10

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

(a) **AO3 = 2**

One aim of the investigation is to see if the age of participants affects their ability to identify a person.
(Credit relevant alternatives)
1 mark for a very brief or muddled aim eg to investigate whether participants can identify a man in a photograph or to investigate EWT or to investigate memory. For 2 marks the aim must be more detailed eg to investigate the effect on EWT or to investigate EWT in a natural setting.

(b) **AO3 = 2**

Participants are less likely to show demand characteristics because in the first part of the experiment they are unaware they are taking part and so are likely to respond more genuinely. In real life settings research has high validity because the findings can be generalised to other similar situations. It is therefore more likely to be relevant to eyewitness testimony in court cases.
1 mark for a very brief or muddled answer eg high ecological validity.
2 marks for accurate elaboration.

(c) **AO3 = 4**

Opportunity sample 1 mark. Volunteer or random = 0 marks.
One limitation is the lack of a target population. This means that the sample is not representative of any population so there are problems in generalising the findings. However, selecting participants for availability is an appropriate way to select a sample when no names are available. Comparison with alternative sampling methods is creditworthy. 1 mark for identifying a limitation or advantage eg biased sampling. Further marks for accurate elaboration or identification of further limitations / advantages. Candidates may refer to one or more limitations, advantages or both. Candidates who identify the sample incorrectly can still gain marks for correct evaluation of opportunity sampling.

(d) **AO3 = 4**

Extraneous variables are anything other than the independent variable that could affect the dependent variable. In this study they could include situational variables such as how the researcher asked for directions or time of day, and participant variables such as gender or eyesight.
1 mark for identification of any possible extraneous variable in this study. Eg One possible extraneous variable is the length of time the researcher spends with each participant.
3 marks for accurate explanation of how this variable could have affected this study. This might have affected the results of this study because old people are more likely to have time to stop and chat than younger participants. They therefore spend longer giving directions and would therefore find it easier to identify the researcher.
1 mark for very brief or slightly muddled explanation.
Further marks for accurate elaboration.

11

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

There are a number limitations of the proposal included in the stem. Some of the most obvious are as follows:

- The independent variable (exercise) is not operationalised. There is no attempt to specify the amount of exercise taken, frequency or intensity. These could vary substantially.
- The DV (happiness) is measured through an interview. Interviews are prone to demand characteristics and social desirability effects. Both of these could affect the validity of the measurement. Students may suggest using a happiness questionnaire to measure the DV.
- As the student proposes to carry out the interviews herself, there is a likelihood of investigator effects. An independent interviewer could be used to reduce this.
- There is a lack of a control group for comparison purposes. The experiment could be modified to use an independent group design, with a control group who do not undertake an exercise programme.
- The use of a volunteer sample means that the study is unlikely to be representative.
- Ethical issues – although a volunteer sample has been recruited, there is no mention of informed consent, confidentiality, debriefing etc.
- Competence, the student is unlikely to have received training to carry out interviews of this nature.

In order to gain credit, students are required to identify one or more of these limitations and suggest appropriate modifications. There is a depth – breadth trade off here: students can cover one limitation in detail or consider several limitations in less detail.

AO2 / AO3 = 10

AO3 Mark bands
10 – 9 Effective The answer demonstrates sound knowledge and understanding of research methods. The student clearly identifies and explains one limitation and provides a detailed discussion of ways to overcome it OR the student identifies and explains several limitations and suggests appropriate modifications for these in less detail.
8 – 6 Reasonable The answer demonstrates reasonable knowledge and understanding of research methods. One or more limitations are identified and there are some reasonable suggestions for modifications.
5 – 3 Basic The answer demonstrates limited knowledge and understanding of research methods. There is some awareness of one or more limitations of the study with weak suggestions for improvement.
2 – 1 Rudimentary The method demonstrates rudimentary knowledge of research methods. The explanation lacks clarity is muddled and incomplete.

0 marks
No creditworthy material presented.

12

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

(a) **[AO3 = 2]**

One mark for either B or C.

One mark for an appropriate advantage of using open questions.
Likely points: open questions provide depth / detail / greater diversity of responses / more meaningful information in the response; they avoid participant frustration associated with fixed choice responses.

(b) **[AO3 = 2]**

One mark for an appropriate conclusion that might be drawn, eg: the majority of people **regard themselves** as kind and helpful people.(Accept alternatives such as 'see themselves, believe or think they are / say they would')

One mark for justification of the answer with reference to the data given, eg: the number of people who reported they would help the person is much higher than any other response given (about 75% said they would help the person).

Accept other valid conclusions with an appropriate matching justification.

(c) **[AO3 = 2]**

Up to 2 marks for an appropriate experimental hypothesis. For full credit the hypothesis must be a testable statement and contain both the IV and DV.

Possible answers for 2 marks:
Non-directional: There is a difference in the number of participants who go to help / help someone when the participant waits alone and when the participant waits with another person.

Directional: More participants who wait alone go to help / help someone than participants who wait with another person. (Accept 'Fewer'.)

Accept null version of the hypothesis.

Possible answers for 1 mark:

There will be a difference in the number of participants who go to help / help in Condition 1 and Condition 2

People who wait alone are more likely to go to help / help than people than people who wait with someone else.

(d) **[AO3 = 3]**

One mark for identification of a possible extraneous variable.
Likely answers: the behaviour of the interviewer who 'falls'; the behaviour of the confederate in the waiting room. Accept EVs based on participant variables eg gender and appropriate condition variables such as 'noise.'

One mark for explaining why the EV should be controlled.

One mark for explaining how it could be controlled.

Possible answers:
The behaviour of the interviewer who falls must be the same – the same sounds and cries so that each participant has the same incident to react to. This could be controlled by using a taped recording of the falling and crying out.
The behaviour of the confederate must be the same so that each participant has the same environment in the waiting room. This could be controlled by using the same person as a confederate who has a script he / she follows for each participant.

(e) **[AO3 = 3]**

One mark for identification of the experimental design as independent groups / measures.

Up to 2 marks for explanation of why this is a suitable design for this study.
Likely points: the participants can only be exposed to the person 'falling' once (1) as they will then have some understanding of what the study is trying to find out and their behaviour will be affected by this knowledge (lack of naivety) (1).

Maximum of 1 mark for generic explanations not linked explicitly to the study

(f) **[AO3 = 2]**

Up to 2 marks for an outline of the procedure of random sampling:

Possible answer:
Put the name of every first year student at the university into a hat (number every first year student)(1).
Draw out 40 names or numbers for the sample (use a random number table / computer program to generate a set of 40 numbers – this represents the sample) (1).

(g) **[AO3 = 2]**

One mark for an appropriate suggestion.
Likely answer: Bar chart / bar graph, frequency graph. Accept pie chart.

One mark for justification of the suggestion.
Likely point: the display clearly demonstrates the numerical difference between the two conditions. Credit discrete data / categorical data.

If more than one graphical display is listed – mark the first answer.

(h) **[AO3 = 4]**

For each of the TWO points, allow one mark for identification of the point and one further mark for discussion of why that point should be raised when the participants are debriefed. Max 2 marks for each point.

For full marks at least one of these points must focus on imparting the aim / purpose of the study or detail of the two conditions.

One further mark for discussion of the chosen point.

Maximum 2 marks if only ethical issue(s) discussed. These 2 marks can only be given for **one** ethical issue (1) that is appropriately discussed (1).

Likely points: explanation of the aim of the study; explanation of the use of independent groups; ethical issues, (these include deception, protection from harm / treating participants with respect; right to withdraw data from the study.)

Verbatim answers are likely to be credited with a maximum of two marks as there would be no discussion / explanation.

13

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

AO2 / AO3 = 4

Peer reviewed research may be accepted, sent back for revisions or rejected. Peer review is an important part of the scientific process because:

- It is difficult for authors and researchers to spot every mistake in a piece of work. Showing the work to others increases the probability that weaknesses will be identified and addressed.
- It helps to prevent the dissemination of irrelevant findings, unwarranted claims, unacceptable interpretations, personal views and deliberate fraud.
- Peer reviewers also judge the quality and the significance of the research in a wider context.
- This process ensures that published research can be taken seriously because it has been independently scrutinised by fellow researchers.

No credit for merely re–stating what is meant by peer review.

AO2 / AO3 Mark Band
4 marks Effective Effective analysis and understanding. The answer contains a coherent explanation of the importance of peer review. Ideas are well structured and expressed clearly and fluently.
3 marks Reasonable Reasonable analysis and understanding. The answer includes one or more of the above points about the importance of peer review. Most ideas appropriately structured and expressed clearly.
2 marks Basic Basic, superficial understanding. The answer shows a basic understanding of the importance of peer review.
1 mark Rudimentary Rudimentary with very limited understanding. The answer is weak, muddled and may be mainly irrelevant.
0 marks No creditworthy material is presented.

14

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

(a) **AO2 / AO3 = 2**

Award 2 marks for an appropriate non-directional hypothesis which is operationalised. 'There is a relationship between happiness scores on a questionnaire and intelligence test scores'.

Award 1 mark for a non-directional hypothesis which is not fully operationalised or lacks clarity ('there is a relationship between happiness and intelligence').
Award no marks for a null or directional hypothesis, or one that predicts a difference / link / association / connection.

(b) **AO2/AO3 = 4**

An interview is the most likely answer. An interview would be a more appropriate method than a questionnaire as it enables questions to be clarified and responses to be probed, thus overcoming the main disadvantages of questionnaires.

Students could also make a case for the analysis of diaries/written materials as a way of collecting data about happiness. These would generally overcome the problems of social desirability and demand characteristics inherent in questionnaires. Students could also make a case for the use of observation.

Award one mark for identifying an appropriate method. Award up to three further marks for an explanation of why this method would be better than a questionnaire.

(c) **AO2/AO3 = 2**

Award 1 mark each for any two of the following reasons:

- Study is looking for a correlation (relationship)
- Suitable for pairs of scores
- The data type obtained is ordinal, at least ordinal or interval level
- Linear relationship between scores.

(d) **AO2/AO3 = 3**

Students should state that the obtained value of + 0.42 exceeds the critical value for a twotailed test (.362) for N = 30. The results are therefore statistically significant ($p \leq 0.05$) Award 2 marks for a student who supplies two pieces of information. Award 1 mark for a student who states that the results are significant but does not provide an explanation OR the student who states results are significant but uses incorrect values from the table. Award 0 marks for students who argue that results are not significant.

(e) **AO2/AO3 = 4**

This question requires students to interpret a further correlation co-efficient (this time demonstrating a non-significant negative correlation) and put both findings together. For full marks, answers should cover the two key bullet points below:

- At age 11, there is a significant positive correlation between happiness and intelligence, demonstrating that more intelligent children tend to be happier
- At age 16, the correlation is not statistically significant.

Students may also make the point that there may be a weak tendency for more intelligent teenagers to be less happy at 16 years of age, although this is not statistically significant. Students may also refer to the contradiction in the results or provide an overall conclusion.

AO2 / AO3 Mark bands
4 marks Effective Effective analysis and understanding. The answer includes the findings of the two studies which are expressed clearly and fluently with appropriate reference to intelligence and happiness. Effective use of statistical terminology.
3 marks Reasonable Reasonable analysis and understanding. The answer is generally focussed and includes reference to both of the key findings which are reasonably clear. There is reasonable use of statistical terminology.
2 marks Basic Basic, superficial understanding. The answer is sometimes focussed OR covers only one of the key conclusions. Expression of ideas lacks clarity. Limited use of statistical terminology.
1 mark Rudimentary Rudimentary with very limited understanding. The answer is weak, muddled and may be mainly irrelevant. Deficiency in expression of ideas results in confusion and ambiguity. The answer lacks structure, often merely a series of unconnected assertions.
0 marks No creditworthy material is presented.

15

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

[AO3 = 3]

One mark per outline of a way: each identified feature of the scientific approach explained in relation to the study.
Possible features from the study: measuring levels of hormones; use of saliva samples; sample of 40 participants; prediction based on theory; statistical testing.
Accept other features that can be inferred eg replication.
Explanations might refer to: empirical method; factual, verifiable, objective measures; precision / measuring on interval / ratio scale; operational prediction / testable hypothesis derived from theory; theory amenable to scientific testing; possible to replicate the procedure; theory capable of refutation; sample size.

Markers should be aware that some of the above scientific principles may overlap.
1 mark for two or more features and / or scientific principles named but not explained.

16

AO3 = 2

The graph indicates a fairly strong, positive correlation between scores on a stress questionnaire and days off through illness. The following can all receive a mark: direction, strength and a description of their relationship. Credit can also be given for mentioning the flattening of the graph at higher stress levels.

17

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

(a) **AO3 = 4**

Content analysis is a way of analysing data such as text using coding units such as themes. In this case mothers were asked to write down how their child behaved, so students might suggest.

Create a checklist / categories
Relevant example(s) of behaviours eg aggression, crying
Read through the diaries / mothers' writing / reports
Counting behaviours or tallying
Compare before and after day care

Any 1 of these equals 1 mark
Any 2 of these equals 2 marks
For 3 marks any 3 components but must refer to reading diaries / mothers' writing / reports.
For 4 marks any 4 components but must refer to reading diaries / mothers' writing / reports.
Ie Max 2 marks if there is no reference to reading diaries.

AO3 Knowledge and understanding of content analysis
4 marks Effective explanation Explanation is accurate, reasonably detailed and demonstrates sound knowledge and understanding of how content analysis could be used. Includes reference to both coding / categorizing and counting.
3 marks Reasonable explanation Explanation is generally accurate but less detailed and demonstrates reasonable knowledge and understanding of how content analysis could be used.
2 marks Basic explanation Explanation demonstrates basic knowledge of how content analysis could be used.
1 mark Rudimentary explanation Explanation demonstrates rudimentary knowledge of how content analysis could be used.
0 marks No creditworthy material.

(b) **AO3 = 4**

Credit all possible limitations of this investigation such as mothers not having time to write much, or to problems in the analysis such as difficulties deciding on appropriate categories. Other limitations could be demand characteristics, mothers dropping out of the study, bias in recording, lack of control of time spent in day care, nine-month-olds not representative of all young children etc. Also ethical issues such as maintaining confidentiality could be made relevant.

Students may explain one limitation in detail, or more than one in less detail.

AO3 Knowledge and understanding of limitations of this investigation
4 marks Effective explanation Explanation is accurate, reasonably detailed and demonstrates sound knowledge and understanding of one or more limitations of this investigation.
3 marks Reasonable explanation Explanation is generally accurate but less detailed and demonstrates reasonable knowledge and understanding of one or more limitations of this investigation.
2 marks Basic explanation Explanation demonstrates basic knowledge of one or more limitations of this investigation.
1 mark Rudimentary explanation Explanation demonstrates rudimentary knowledge of one or more limitations of this Investigation.
0 marks No creditworthy material.

18

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

(a) **AO3 = 1**

Volunteer / volunteering or self-selected / self-selecting sample. 1 mark
Voluntary 0 marks.

(b) **AO3 = 2**

A limitation of a volunteer sample is that it is biased / not representative (1 mark) because some people are more likely to volunteer than others (1 mark) or the findings cannot be generalised to a population (1 mark).

1 mark for a very brief or slightly muddled explanation eg it is biased. 2nd mark for accurate elaboration. For 2 marks the answer must relate explicitly to volunteer sampling.

(c) **AO3 = 2**

IV The interview, type of interview, method of interview, 'standard interview or / and cognitive interview', whether or not cognitive interview.
DV Number of items recalled, recall, what they remembered.

(d) **AO3 = 2**

There is better control because the same film can be used in both conditions. The participants are less likely to show demand characteristics because they take part in only one condition. There are no order effects such as practice or fatigue, because participants take part in one condition. 1 mark for very brief or slightly muddled advantage. 2nd mark for accurate elaboration.
0 marks for simply stating there are different participants in each condition or takes less time.

(e) **AO3 = 3**

The question asks about recruiting participants, so answers referring to debriefing are not relevant.
There was no deception. Participants knew they would be watching a film of a violent crime and that they would be interviewed about the content by a male police officer before they volunteered. This gave them the opportunity to give informed consent. Students may argue that the psychologist did not follow BPS guidelines eg because they were not told of their right to withdraw.

1 mark for a very brief or slightly muddled answer, linking a relevant ethical issue to whether or not awareness was shown. Further marks for accurate elaboration / discussion.

Eg He told them what he was going to do. (1 mark) They could give informed consent because he told them what he was going to do. (2 marks) The participants were told that they would be watching a violent crime so they were able to give informed consent. (3 marks)

19 **[AO3 = 4]**

Up to two marks for outlining each problem. One mark for a brief point, 2nd mark for elaboration / explanation.
Possible content: problem of small sample not being representative; individual differences affecting generalisation; problem of sample generalisation including animals to humans; often difficult to represent the many different factors that characterise a population in the sample; problem of generalisability of findings from one culture to another / different cultures; general issue of subject matter being humans, thus varied and less predictable than subject matter in other sciences; generalisability across time; generalisability relating to task, context and location; relating findings from an experiment to life in the real world / beyond the immediate setting (ecological validity).
Credit use of evidence as elaboration.

20 (a) **AO2 = 3**

The answer must clearly relate to one or more of the main techniques used in a cognitive interview (other than report everything):-

Context reinstatement
Recall from a changed perspective
Recall in reverse order

Some of the main additional features of the enhanced cognitive interview could be relevant, as long as it could be explained to the participant: – eg Encourage to relax

1 mark for identification of a relevant cognitive technique.
1 mark for very brief statement eg “tell me what you saw in reverse order”.
Second mark for appropriate elaboration eg “Tell me what you saw on the film in a different order to how it actually happened.” If instructions are not suitable to be read out maximum 1 mark for this part.
For 3 marks technique and instructions must match.

(b) **AO3 = 2**

The researcher might conclude that the cognitive interview was effective because more correct items were recalled, but it did not affect the number of incorrect items recalled.

0 mark - the cognitive interview was effective with no explanation.
1 mark - it was effective because there were more correct items recalled or it was not effective because the number of incorrect items stayed the same.
2 marks - it was effective because there were more correct items recalled **and** the number of incorrect items stayed the same / didn’t increase.

1 mark for stating there were more correct items recalled with the cognitive interview than with the standard interview **and** the number of incorrect items recalled was the same. (There is no reference to effectiveness).

21

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

(a) **[AO3 = 1]**

One mark for answers either:

- referring to the strength and the direction of the relationship – a positive correlation between the number of hours spent reading fiction and the empathy test score.
  or:
- describing the relationship – the more hours spent reading fiction, the greater the empathy test score.

No credit for just stating type of correlation eg strong positive.

(b) **[AO3 = 2]**

One mark for naming a test: Spearman's rank order correlation / rho or Pearson's product moment correlation.
One mark for justification. For Spearman's rank order correlation accept: not all data is interval – data collected for empathy test score most likely treated at ordinal level of measurement due to self-report.
For Pearson accept: Pearson's product moment correlation is a robust test, even if not all data can be treated as truly interval.

Just stating ordinal / interval no credit. Accept ordinal or interval providing this is justified with reference to at least one variable.

Unlikely but allow for an informed argument made for treating both sets of data at interval level.

(c) **[AO3 = 2]**

1 mark for a knowledge of a way (not just naming a type of validity) and 2nd mark for explaining how this would be implemented in this case. Most likely answers will address face validity or concurrent validity, but accept any other way such as construct validity, content validity, criterion validity and predictive validity.

For full marks, the answer must refer to either the empathy questionnaire or empathy test items. The 'way' need not be named or defined.

(d) **[AO3 = 2]**

One mark for the identifying a methodological limitation of the study.
Likely answers: size / composition of sample / one school only; for test of empathy – no evidence of testing reliability; parental involvement in 'time spent reading questionnaire'; self-report measures; correlation study.
One mark for a brief explanation.
Suggested explanations might cover: limits to generalisation; confidence in a test and its findings rests on it being deemed reliable; social desirability of parental responses and consequent bias; honesty of reporting / memory recall; cause and effect issues in correlation studies.
Accept any other plausible answers.

(e) **[AO3=3]**

Up to three marks for a discussion of reasons for correlation studies rather than experiments when investigating behaviour.
Likely answers: unethical / impossible to manipulate these variables (reading and empathy in children) to investigate cause and effect; impractical to sometimes do an experiment; may discover a link between two existing variables which might suggest future research ideas; interested in relationships **rather than** a causal explanation.
Accept comparison with the experimental approach.
For full marks, the answer must be coherent and applied to this study.
Maximum of two marks for general answers not applied to this study.

(f) **[AO3 = 8]**

Up to 8 marks for answers demonstrating an ability to design an experiment effectively. Answers should refer to:

- clearly identified independent and dependent variables and at least one extraneous variable identified and control suggested;
- the experimental design – independent groups, repeated measures or matched pairs;
- detail of sample;
- materials required for carrying out the research, eg task for assessing levels of recall, timing device if needed;
- sufficient procedural details to carry out a replication (might include standard instructions, ethics, etc.)

Note: standardised instructions and ethical issues are not required for full marks.

Mark bands

8 – 7 marks	**Very good answers** All 5 points well addressed and some sound justification. Answer shows sound knowledge and understanding and an ability to design an appropriate experiment. The proposal is coherent and feasible, and includes details of all the essential elements of the chosen design. Information allows for clear understanding of the proposed design. There may be some minor omission(s) at the bottom of the band.
6 – 5 marks	**Good answers** 3 or 4 points well addressed and some justification. The design shows knowledge and understanding and some ability to design an appropriate experiment. The proposal is feasible but may lack the clarity and coherence of the top band. There may be some inaccuracies and omissions.
4 – 3 marks	**Average to weak answers** At least 3 points are addressed and attempt at justification. The answer shows some knowledge and understanding but detail of the proposal may lack clarity. There are inaccuracies and omissions.
2 – 1 marks	**Poor answers** 1-2 points are addressed. There must be some relevant material. The experimental method may not be obvious. There may be substantial confusion, inaccuracy and / or irrelevance.
0 marks	**No relevant content**

22 **[AO3 = 3]**

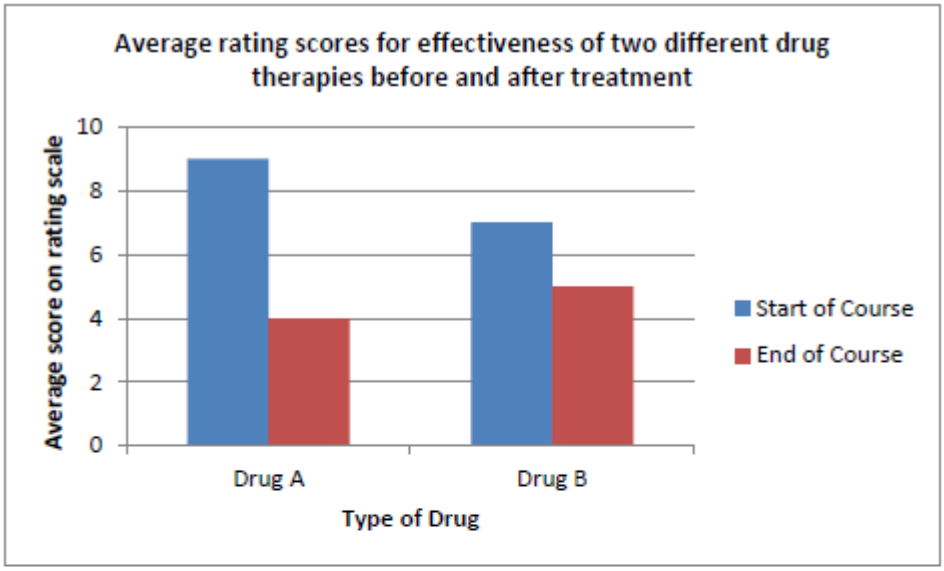
Up to three marks for explaining how the psychodynamic approach as depicted in the stem neglects the rules of science. Students may offer a brief elaboration on two or more rules of science identified in the study as ‘neglected’ or may choose to elaborate on a single one.
Likely answers: interpretation of content of dreams open to bias and subjectivity; no verifiable evidence; small sample; opportunity sample of friends and implications for generalizability; qualitative data collected and implications for statistical analysis; retrospective data / memory distortions – reports written on waking; dreams are private experience and covert; problem of replicability. Credit other possible answers if made relevant to the scenario, eg no reference to a testable hypothesis.
Markers should be aware that some of the above points may overlap and should look for a coherent answer for full marks.

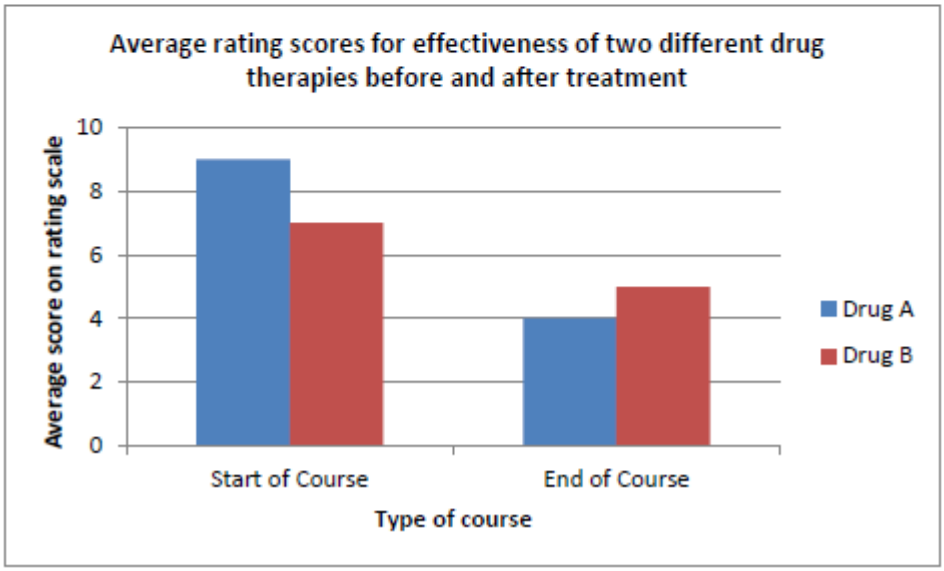
23

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

AO3 = 4





1 mark for correctly labelled x-axis: either with over-arching label, e.g. Type of Drug, or by clearly labelling the 2 conditions e.g. Start and End of course.

1 mark for correctly labelled y-axis: average score on (rating) scale.

Up to 2 marks for clearly sketching a bar chart. For full marks, there needs to be an appropriate use of graph paper and bars labelled correctly.

24

(a) **[AO1 = 3]**

Up to 3 marks for description of a valid way, one mark for each relevant detail. Full mark answers should refer to the method and DV / what was being measured (do not credit aims / conclusion). Likely answers include: studies of imitation, eg Melzoff and Moore (1977); studies of interactional synchrony, eg Condon and Sander, Murray and Trevarthen (1985); studies of skin-to-skin contact, eg Klaus and Kennell (1976); studies of sensitive responsiveness and the Strange Situation, eg Ainsworth et al (1978), De Wolff and van Ijzendoorn (1997).
More generic methodological answers which cannot be identified as a specific study (either by name or description) may gain a maximum of two marks.
No credit for animal studies.

(b) **[AO3 = 3]**

Up to 3 marks for evaluation of the way described in (a). Students who present an inappropriate study or no study in (a) may still gain marks for (b) where it becomes clear that a specific study / way of investigating caregiver-infant interaction is being evaluated. Students may choose to elaborate on one issue or may mention more than one issue in less detail. Evaluative points will vary according to the method described but likely issues, include: usefulness of controlled experimentation in researching social relationships eg artificiality v cause and effect; usefulness of combining data from several studies as in meta-analysis; inferences based on findings, eg studies of imitation and the issue of intentionality; short-term v long-term effects.

For full marks evaluative point(s) must be fully applied to the study of caregiver-infant interaction. One mark only for a totally generic yet valid response.

25

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

(a) **[AO3 = 1]**

A median score is calculated by putting all the scores in order from lowest to highest (or vice versa) and finding the middle score in the set.

Credit explanations that refer to sets with even numbers of values, ie by finding the numerical mid-point between the two middle scores.

(b) **[AO3 = 1]**

One mark for identification of the dependent variable:

The estimate of how many marks would be scored on the maths test.

Accept alternative wording.

(c) **[AO3 = 2]**

2 marks for a clear, testable statement containing both conditions of the IV and an operationalised DV.

There is a difference in the estimates men will give of their scores on a maths test and the estimates women will give of their scores on a maths test. (Accept a null version)

Men will give higher (or lower) estimates of their scores on a maths test than the estimates women will give of their scores on a maths test.

For 1 mark – a statement with both conditions of the IV and a DV which may not be operational or testable.

No marks for expressions of aim / questions / correlational hypotheses or statements with only one condition.

(d) **[AO3 = 2]**

One mark for identification of the experimental design as independent groups / measures / samples. No credit if incorrect design is chosen.

One mark for an explanation of this design.
The participants in the male group are different people from those in the female group or reference to only participating in one condition.

(e) **[AO3 = 3]**

Up to 3 marks for a clear explanation of the procedure for obtaining a random sample in this study:

- put all the names / numbers of the men (from the factory) in a hat / computer (1)
- draw out 15 names for the sample or get computer to randomly generate 15 numbers (1).
- repeat for all the women (of the factory) (1) or vice versa.

Accept other ways that would clearly generate a random sample.

(f) **[AO3 = 2]**

Up to 2 marks for a suitable conclusion drawn from the data in **Table 2.**

One mark for the conclusion:
There is a difference in the expectations men and women have of their own numeracy skills.
Accept a directional statement such as:
Men think they are likely to achieve better scores on a maths test than women think they will achieve.

One mark for the justification:
The median estimate for men was much higher than that for the women.

(g) **[AO3 = 3]**

Up to 3 marks for what comparison of the estimated and actual maths scores of the men and women indicates.

Men overestimated their numeracy skills / numerical ability / score they would get (or similar) (1 mark)

Women underestimated their numeracy skills / numerical ability / score they would get (or similar) (1 mark)

Overall conclusions:

- People / Men / Women are not very good at estimating our ability
- Expectations were wrong
- (Although estimates are different) ability was the same
- Men are over confident re ability
- Women under confident re ability

(1 mark for any of these)

(h) **[AO3 = 4]**

Information must be written in verbatim form for more than 1 mark.

Essential points	Purpose / aim of the study Ethical point
Optional points	Background information / Elaboration of the aim and conditions Any questions? Thanking for participation Interest in the results?

4 marks	**Verbatim** Both essential and at least one optional point are addressed clearly such that an understanding of the study is achieved. Information should be clear, relevant, sensible and logically structured.
3 marks	**Verbatim** Both essential points are addressed such that there is reasonable understanding of the study. There may be deficiencies in clarity, some irrelevance, illogical sequencing or inappropriate content.
2 marks	**Verbatim** Any 2 points are addressed. There may be omissions / irrelevances / muddle such that understanding of the study might be limited.
1 mark	There must be at least one relevant point. Information may be unclear / inappropriate / irrelevant such that understanding of the study would be very limited or most points addressed but not in verbatim form.
0 marks	No relevant information. Understanding of the study would not be possible.

(i) **[AO3 = 2]**

One mark for a valid reason and a further mark for elaboration of the reason given.

Reason: to identify any possible flaws in (the design of) the study (1 mark).
Elaboration: to provide an opportunity to improve the study / or so that the researcher does not waste time collecting data that will have to be discarded (1 mark).

Accept any other valid answer.

26

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

(a) **[AO3 = 1]**

One mark for identifying the independent variable in this study.
Likely answer: whether the participants in the study have OCD or not.

(b) **[AO3 = 2]**

Award two marks for an explanation of why this study is a quasi-experiment.

Possible points: This study is a quasi-experiment because the IV (whether the participants have OCD or not) is pre-existing / naturally occurring (1) the IV has not been manipulated / could not have been controlled by the researcher (1) random allocation is not possible (1).

(c) **[AO3 = 1]**

One mark for identifying one relevant variable that could have been used to match participants in this study.
Likely answers: gender; age; health; IQ; ethnicity; weight.

(d) **[AO3 = 2]**

Award one mark for an outline of an advantage of matched pairs and one further mark for an explanation of why this is an advantage. For two marks there must be some application to the study described.

One mark only for an advantage of matched pairs not linked to the study described.

Possible answer: One advantage of matched pairs is that participant variables / individual differences are controlled / reduced (1) so the researcher can be more confident that the results are due to OCD, rather than other variables (1).

Do not credit participant variables are eliminated / removed.

27

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

(a) **AO1 = 2**

Content analysis is a technique for analysing qualitative data of various kinds. Data can be placed into categories and counted (quantitative) or can be analysed in themes (qualitative).

Award 1 mark for a brief statement and a further mark for elaboration.

(b) **AO3 = 4**

- The psychologist could have begun by watching some of the film clips of driver behaviour.
- This would enable the psychologist to identify potential categories which emerged from the data of the different types of distractions seen in the film.
- Such categories / themes might include: passenger distractions, gadget distractions, etc.
- The psychologists would then have watched the films again and counted the number of examples which fell into each category to provide quantitative data.

Credit variations in so far as they explain the process.

Note: maximum 1 mark if no engagement with the stem.

AO3 Mark bands
4 marks Effective Effective explanation of the processes involved in content analysis referring to some or all of the above points.
2 – 3 marks Reasonable Reasonable accurate coverage of the processes involved.
1 mark Basic Basic identification of the processes involved in content analysis (‘watching the films and counting’).
0 marks No creditworthy material.

(c) **AO3 = 3**

1 mark for identification of an appropriate way of assessing reliability in this investigation. By far the most likely answers here are inter-rater reliability or test-retest reliability.

2 marks for some explanation / elaboration: ‘the two psychologists could carry out content analysis of the films separately and compare their answers’ or ‘they could re-code the films at a later date and compare the two sets of data’.

3 marks for an accurate and clear explanation which refers to deriving the categories and checking the data. ‘The two psychologists could watch the films separately and devise a set of categories. They could compare these and use categories they both agreed on. They could carry out content analysis of the films separately and compare their answers looking for agreement’.

(d) **AO3 = 3**

Candidates can cover one reason explained in detail here or several reasons in less detail.

A repeated measures design was chosen in this experiment:

- to remove the effects of individual differences in reaction times which would occur if an independent groups design was used
- to avoid the potential difficulties involved in matching participants
- to reduce the number of participants required for the experiment.

(e) **AO3 = 3**

This is a repeated measures design and is counter-balanced hence points about order effects and individual differences will not gain credit.

There are a range of potential extraneous variables here including:

- the nature and content of the conversation with the psychologist on the hands-free phone
- interaction between the sex of the psychologist and sex of participant which could influence the type of conversation
- the number of hazards in the computer-based test, hence difficulty of the tests
- the presence of the hands-free headset could have produced distraction.

Award 1 mark for basic identification of a confounding variable and a further 2 marks for elaboration of how this could have affected the dependent variable.

Example: The chat with the psychologist was not controlled (1 mark) so the difficulty or number of questions could have varied (2 marks). This would influence the DV as more or less attention would be required (3 marks).

(f) **AO3 = 4**

There are several potential ethical issues here. Candidates can focus on one in detail or several in less detail.

- Protection of participants from harm whilst studying the effects of a hands-free phone on driving. Two key issues here are the use of a computer-based test with no risk attached and of an experienced sample of police drivers.
- Informed consent: Participants should be given full information about the nature of both tasks before deciding whether or not to participate.
- Debriefing: A full debriefing should take place at the end of the experiment. This should provide feedback on performance and allow participants to ask questions if they wish to.
- Freedom to withdraw: Participants should be made aware of their freedom to withdraw before and during the experiment. They should be made aware of their right to withdraw their data after the experiment.
- Confidentiality: Individuals should not be identified, but should retain anonymity (use of numbers or initials instead of names).

Lists of ethical issues with no elaboration 1 mark.

AO3 Mark bands
4 marks Sound An appropriate ethical issue is identified and explained in detail. Material is accurate – or several issues are identified and discussed accurately in less detail.
2 – 3 marks Reasonable One or more appropriate ethical issues are identified and discussed. The answer is generally accurate.
1 mark Basic Basic identification of an ethical issue (e.g. ‘right to withdraw’) or very brief answers which lack detail.
0 marks No creditworthy material.

(g) **AO3 = 5**

The standardised instructions should include the following information:

a. You will take part in a simulated driving test which will last for three minutes.

b. Your task will be to identify potential hazards on the road ahead.

c. When you see a hazard, you should press the mouse button as quickly as possible.

d. Whilst you are doing the test, I will chat to you on a mobile phone and I would like you to reply using the hands-free mobile phone headset.

e. Do you have any questions?

For full marks, the instructions should adopt an appropriate formal tone. Instructions which are not suitable to be read out should be awarded a maximum mark of 2.

AO3 Marks bands Standardised instructions
5 marks Effective The standardised instructions provide accurate detail of the procedure in a clear and concise form and participants' understanding is checked.
4 – 3 marks Reasonable The standardised instructions provide sufficient detail of the procedure in a reasonably clear form.
2 marks Basic The standardised instructions provide some details of the procedure though these may not be clear.
1 mark Rudimentary The standardised instructions provide few details of the procedure and may be muddled and or inaccurate. Omissions in the instructions compromise the procedure.
0 marks No creditworthy material is presented.

(h) **AO3 = 3**

Students are required to identify an appropriate test and are asked to justify their choice.

Award 1 mark for identification of the Wilcoxon (signed ranks) test. Candidates could receive credit for Sign test or related t test. Note that reasons / justification must be correct for the test supplied.

If an incorrect test is identified **no marks** can be awarded.

Award 1 mark for basic statement of a reason, and a further mark for elaboration, within the context of the experiment or a further reason.

e.g. for Wilcoxon test:

- A repeated measures design was used (1 mark) as drivers take part in both the hands-free phone and non-phone (silent) conditions (1 mark).
- A repeated measures design was used (1 mark) and the data can be treated as ordinal (1 mark).

Test of difference cannot gain credit.

(i) **AO3 = 2**

Students are told that the difference in reaction times was significant at the $p \leq 0.01$ level.

Award 1 mark for a basic understanding of this ('the result is highly significant') and a further mark for elaboration e.g. identifying that the probability of a Type 1 error here is less than 1 / 100.

(j) **AO3 = 3**

Replication is an important tool in the scientific method. It allows scientists to check findings and ensure that they are robust. In this study, replication is important, as the original sample is small (30 people) and specific (experienced police drivers). For this reason, replication on a larger sample will be used to check if findings apply outside this specific group.

Award 1 mark for a general answer on the importance of replication to check findings.

28

AO3 = 4

Two strengths of using questionnaires could include:

- Compared to interview they are easy to use (1 mark). The researcher doesn't need any special training to hand out the questionnaires (2nd mark for elaboration).
- People may be happier to disclose personal information on a questionnaire (1 mark) compared to a face-to-face situation (2nd mark for elaboration).
- Participants can answer the questions without the need for the researcher to be present (1 mark) so reducing experimenter bias (2nd mark for elaboration).
- If the questionnaire used closed questions which generate quantitative data, this is easier to analyse (1 mark) than open questions which generate qualitative data which is difficult to analyse (2nd mark for elaboration).
- Can be given to a large group of people (1 mark).

For each strength, 1 mark for identifying the strength and a further mark for explaining why it is a strength. The final bullet point is an example of a 1-mark answer as there is no explanation of why it is a strength.

Examiners should be aware that this question asks about the strengths of the method, not of the type of data collected. Answers that refer to data should not receive credit unless they are explicitly related to the type of question used (as illustrated in the bullet point above).

29

AO2 = 4

Possible content:

- Median is 29.5 (29 + 30/2) for Group A and 24.5 (24 + 25/2) for Group B

1 mark for each accurately calculated median
2 further marks for explaining the median is the more appropriate measure because of the outlying extreme scores in each group which could have distorted the mean.

Accept answers based on unsafe level of measurement.

30

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

AO3 = 4

- A smartly dressed confederate elicits more obedience (1 mark) than a casually-dressed confederate (second mark)
- The type of task (request) also influences rate of obedience (1 mark)
- If told to do something that requires effort (e.g. heavy task) obedience levels are not affected by what the person is wearing (2 marks)

AO3 Interpretation of data
4 marks Effective interpretation of data Effective interpretation that demonstrates sound knowledge of what the data shows, with reference to both what the confederate is wearing and type of task.
3 marks Reasonable interpretation of data Reasonable interpretation of what the data shows, with reference to what the confederate is wearing and the type of task, but one in more detail.
2 marks Basic interpretation of data Basic interpretation of what the data shows; in terms of, for example “more” or “less”.
1 mark Rudimentary interpretation of data Rudimentary, muddled interpretation of the data, demonstrating very limited knowledge.
0 marks No creditworthy material.

31

AO2 = 4

Level	Marks	Description
2	3 – 4	Explanation of how psychology / social influence research might affect the economy is clear. There is effective application to the example of eating healthily. The answer is generally coherent with effective use of terminology.
1	1 – 2	There is limited / partial explanation of how psychology / social influence research might affect the economy. There is limited application to the example of eating healthily. The answer lacks coherence. Use of terminology is either absent or inappropriate.
	0	No relevant content.

Possible content:

- Social influence research tells us how behaviour and attitudes can be changed: eg how minority influence can be exerted or how people tend to conform to perceived norms (or reference to any other relevant social influence process).
- In this case, the resulting change of eating more healthily means that people should be more healthy.
- Economic implication: eg saves health service / care resources; means less time off work sick.

Credit other relevant information.

32

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

(a) **AO3 = 4**

Suitable behavioural categories for investigating children's aggressive behaviour could be:–
pushing, hitting, biting, punching, swearing, etc.
Maximum 2 marks – 1 for each suitable behaviour category.
Candidates may suggest recording playground behaviour on CCTV for later analysis by ticking a box when a relevant behaviour is shown by the child. Alternatively the researcher could watch each child's behaviour in the playground and tick the box when each behaviour is shown. In this case where the researcher stands and whether the children know they are being observed would be relevant.
1 mark for a very brief or slightly muddled explanation e.g. use a tally chart
2nd mark for accurate elaboration

(b) **AO3 = 4**

There are no ethical issues named in the specification, so any potentially relevant issues in this research should be credited.
Although the psychologist would not be responsible for the behaviour of the children in the playground he might consider his responsibility if he saw that one of the children was being harmed.
Likely ethical issues include informed consent, right to withdraw, confidentiality or respect. Ways of dealing will depend on the issue selected.
There are different routes to achieving 4 marks depending on the ethical issue selected, but for full marks both the ethical issue and how the psychologist could have dealt with it should be clear.

AO3 Knowledge of research methods
4 marks Accurate and reasonably detailed Accurate and reasonably detailed answer that demonstrates sound understanding of one relevant ethical issue and how the psychologist could have dealt with this issue.
3 marks **Less detailed but generally accurate** answer that demonstrates relevant understanding of one relevant ethical issue and how the psychologist could have dealt with this issue. or **Accurate and reasonably detailed answer** that demonstrates sound understanding of one relevant ethical issue or how the psychologist could have dealt with an issue.
2 marks Basic Basic answer that demonstrates some relevant understanding of one relevant ethical issue and / or how the psychologist could have dealt with an ethical issue, but lacks detail and may be muddled.
1 mark Very brief / flawed Very brief or flawed answer demonstrating very little understanding of a relevant ethical issue and / or how the psychologist could have dealt with an issue.
0 marks No creditworthy material.

(c) **AO3 = 4**

There are different routes to full marks in this question. Candidates explain one advantage in reasonable detail or more advantages in less detail.
Advantages of using an interview rather than a questionnaire could include it would allow the interviewer to clarify questions and answers; it might be easier to see if participants were answering honestly because their reactions could be observed; it is easier to collect detailed qualitative data.

AO3 **Knowledge of research methods**
4 marks Accurate and reasonably detailed Accurate and reasonably detailed answer that demonstrates sound understanding of one or more advantages of using interviews rather than questionnaires in this situation.
3 marks Less detailed but generally accurate Less detailed but generally accurate answer that demonstrates relevant understanding of one or more advantages of using interviews rather than questionnaires in this situation.
2 marks Basic Basic answer that demonstrates some relevant understanding of one or more advantages of using interviews rather than questionnaires in this situation, but lacks detail and may be muddled.
1 mark Very brief / flawed Very brief or flawed answer demonstrating very little understanding of one or more advantages of using interviews rather than questionnaires in this situation.
0 marks No creditworthy material.

33 **AO2 = 4**

Content:

- Median is 34.5 for Group A (32 + 37/2) and 50.5 for Group B (45 + 56/2)

1 mark for each accurately calculated median

Plus

2 further marks for explaining that the median is used because the level of measurement is not interval – ratings data with units of variable size.

34

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have changed. Under the new Specification the following system of AOs applies:

- AO1 knowledge and understanding
- AO2 application (of psychological knowledge)
- AO3 evaluation, analysis, interpretation.

Although the essential content for this mark scheme remains the same, mark schemes for the new AQA Specification (Sept 2015 onwards) take a different format as follows:

- A single set of numbered levels (formerly bands) to cover all skills
- Content appears as a bulleted list
- No IDA expectation in A Level essays, however, credit for references to issues, debates and approaches where relevant.

(a) **AO3 = 4**

In this experiment a pilot study could be used to:-

- check how long the participant should be given to look at the stimulus material
- check whether the pictures were appropriate and clear
- check whether 20 is an appropriate number of words to use
- check whether the words were appropriate
- check the participants understand the instructions and what they are required to do
- ask a few participants about their experience of taking part

Changes can then be made to the procedure if necessary, to avoid wasting time / money.
There is a depth / breadth trade off. Candidates may cover one point in detail or more than one in less detail.

Vague or general statements which simply state “to save time / money” , “to see of it works” ,
“to see if there is a difference” = 0
To test / change the hypothesis = 0

AO3 Application of knowledge of research methods
4 marks Accurate and reasonably detailed Accurate and reasonably detailed explanation that demonstrates sound knowledge and understanding of why a pilot study would be appropriate, including at least one detail from the experiment.
3 marks Less detailed but generally accurate Less detailed but generally accurate answer that demonstrates sound knowledge and understanding of why a pilot study would be appropriate, including at least one detail from the experiment.
2 marks Basic Basic answer that demonstrates some understanding of why a pilot study would be appropriate in this study, but lacks detail and may be muddled.
1 mark Very brief / flawed Very brief or flawed answer demonstrating very little understanding of why a pilot study would be appropriate in this study.
0 marks No creditworthy material.

(b) **AO3 = 2**

0 marks for a directional / correlational / null hypothesis.
1 mark for an appropriate non directional hypothesis where either or both variables are not operationalised e.g. memory will be different in the two conditions and / or when the hypothesis is not written as a statement e.g. " To see if ..." or " Is there.....?"
2 marks for an appropriate non directional hypothesis where both variables are operationalised e.g. there will be a difference in the number of words correctly recalled when words are presented with pictures and without pictures.

(c) **AO3 = 4**

Reasons for using an independent groups design rather than repeated measures include:-
There are no order effects because participants only do the task once.
The same words can be used in both conditions so one set of words is no easier to recall than the other set of words.
Demand characteristics are less likely because participants will be unaware of the other condition.
Credit other appropriate reasons.
Simply stating IGD is quicker / saves time = 0.
In each case 1 mark for a very brief / slightly muddled potentially relevant reason that could explain the use of IGD.
2nd mark for some elaboration of a reason that is relevant / appropriate to this study.

(d) **AO3 = 2**

The focus of this question is on understanding the outcome of this experiment.
Simply re-stating the data in table 1 = 0
e.g. The range for Condition 1 is 11 and for Condition 2 is 13.
Or The range is higher for Condition 1 than for Condition 2.
Or The median for Condition 1 is 13 and Condition 2 is 16.
Or The median for Condition 2 is higher than Condition 1.
1 mark for accurate reference to either median or range
e.g. more words were correctly recalled with pictures than without pictures.
Or The spread / dispersion of scores is larger with pictures than without pictures.
Or There is more individual variation with pictures than without.
2 marks for accurate reference to both difference and dispersion (spread) as above.

35 **AO2 = 4**

Content

- Median is 11 for Group A (9 + 13/2) and 8.5 for Group B (8 + 9/2)

1 mark for <u>each</u> accurately calculated median

Plus

2 further marks for explaining that the median is used because of the outlying/extreme scores (one in each group) which would have distorted the mean.

Also accept answers based on unsafe level of measurement.

36

AO1 = 4

1 mark each for a correct definition of both a Type I and a Type II error
Plus
Up to 2 marks for a clear distinction between these two errors.

Possible content:

- A Type I error occurs when a researcher claims support for the research hypothesis with a significant result when the results were caused by random variables
- A Type II error occurs when the effect the researcher was attempting to demonstrate does exist but the researcher claims there was no significance in the results/erroneously accepts the null hypothesis
- The difference is that in a Type I error the null hypothesis is rejected when it is true and in a Type II error it is retained when it is false.

37

AO2 = 4

Level	Marks	Description
2	3 – 4	Knowledge of the effectiveness of atypical and typical antipsychotics on positive and negative symptoms is clear and mostly accurate. The findings in the table are used appropriately. The answer is generally coherent with effective use of terminology.
1	1 – 2	Some knowledge of the effectiveness of atypical and typical antipsychotics and positive and negative symptoms is evident. Use of findings from the table is not always effective. The answer lacks accuracy and detail. Use of terminology is either absent or inappropriate.
	0	No relevant content.

- Atypical and typical antipsychotics are equally effective against positive symptoms with more than half of patients responding well
- The main difference is that negative symptoms respond better to atypical antipsychotics, 30% improve compared with typical antipsychotics 16%
- Atypical antipsychotics are more effective against negative symptoms
- These findings support the view that they act on different neurotransmitters

38

(a) **AO3 = 4**

Level	Marks	Description
2	3 – 4	Clear understanding of the notion of social sensitivity is demonstrated through effective application to the stem. Explanation of how the researchers could have dealt with the issue of social sensitivity in this case is clear. The answer is generally coherent with effective use of terminology.
1	1 – 2	Some understanding of the notion of social sensitivity is demonstrated through limited application to the stem. There is limited/partial explanation of how the researchers could deal with the issue of social sensitivity in this case. The answer lacks accuracy and detail. Use of terminology is either absent or inappropriate.
	0	No relevant content.

Content:

- Awareness of issue: Researchers should be aware of the implications of their research: possible negative impact for the children in the sample; possible negative implications of the research for the reputation of Crayford school and the wider community; possible self-fulfilling prophecy
- Dealing with the issue: Researchers should take adequate steps to counter the above: sensitive briefing/debriefing of participants, parents, teachers etc; care in relation to publication, disclosure of results and confidentiality/anonymity.

(b) **AO2 = 1**

1 mark for nominal level/categorical level

(c) **AO3 = 2**

1 mark – categorical data is crude/unsophisticated/does not enable very sensitive analysis

Plus

1 mark – because it does not yield a numerical result for each participant

39

AO1 = 2

Possible content:

- The likelihood of the same differences occurring twice (or more), by chance alone are much smaller than when they occur the first time.
- Effects that occur in a study are more likely to be reliable if they occur in a repeat of the study – replication therefore increases (external) reliability.

40 **AO2 = 9**

Level	Marks	Description
3	7 – 9	Suggestions are generally well detailed and practical, showing sound understanding of design of an experiment. **All three** elements are present. There is sufficient information for most aspects of the study as required to be implemented with success. The answer is clear and coherent. Specialist terminology is used effectively. Minor detail and/or explanation sometimes lacking.
2	4 – 6	Some suggestions are appropriate but there may be a lack of detail. At least **two** elements are addressed. Implementation may be difficult given the lack of information. The answer is mostly clear and organised. There is some appropriate use of specialist terminology.
1	1 – 3	**At least** one element is addressed but knowledge of task design or dealing with participants is limited. Successful implementation would be difficult given the information provided. There is substantial inaccuracy/muddle. Specialist terminology is either absent or inappropriately used.
	0	No relevant content.

Possible content:

- **The task:** the answer must show an appreciation of the fact that the usual way of merely sorting a shuffled pack of cards into suits will have to be modified in order to ensure that each participant has exactly the same task. [Initial shuffle, record the order, reinstate that order for each participant.]
- **Suitability of participants:** the answer must include information about how familiarity with cards could become a confounding variable if not controlled and how this could be controlled practically.
- **Ethical issues:** specific or more general ethical considerations as applied to this study – protection of welfare, confidentiality, respect or integrity.

41 (a) **AO2 = 2 and AO3 = 2**

2 marks for an accurate comment about the means for both males and females
Plus
2 marks for an accurate comment about the standard deviations for both sets of data

Means: the mean score for males is almost 3 times larger than that of the females which suggests they are very much better at map reading than the females

Standard deviations: sds are quite similar to each other suggesting the spread of performances of the male participants and the female participants is similar within each group.

(b) **AO2 = 4**

Award 2 marks for a correct calculation of the percentage for the male participants and 2 marks for a correct calculation of the percentage for the female participants.

If the calculation for one or both of the groups is incorrect but the procedure used is correct award 1 mark for each time this occurs to a maximum of 2 marks.

Males – 13/20 = 65%

Females 5/20 = 25%

(c) **AO2 = 2**

Up to 2 marks for a clear comment on the data

Possible content: the difference in the percentages confirms the earlier suggestion that men are much better at map reading than women.

42 **AO3 = 4**

Level	Marks	Description
2	3 – 4	Explanation of problem and way of dealing with it is clear and mostly appropriate. The answer is generally coherent with effective use of specialist terminology.
1	1 – 2	Some explanation of problem and/or appropriate way of dealing with it. The answer lacks accuracy and detail. Use of specialist terminology absent or inappropriate.
	0	No relevant content.

Possible content:

- Problem – random sampling; the 3 pm group might simply have been better at maths than the 3 am group. The solution would be a matched pairs (matched on maths ability) or repeated measures design.
- Problem – use of different maths tests, with no evidence that they were matched for difficulty. The solution would be to use the same set of maths problems if a matched pairs design was used.
- Individual differences due to independent groups design so use repeated measures but would need different but equivalent tests and counterbalancing.
- Other issues, such as individual differences in biological rhythms (‘owls’ versus ‘larks’) confounding results. Such answers should be marked on their merits – is the problem plausible and is the solution sensible?

43

(a) **AO2 = 2**

2 marks for explanation that a non-directional hypothesis is suitable or 'it should not be directional,' (1) as there is no reference to evidence that allows the researchers to predict the direction of the results (1).
1 mark for a muddled/limited explanation of why the hypothesis should be non-directional or **1 mark** for stating non-directional.

(b) **AO2 = 3**

3 marks for an appropriate non-directional operationalised hypothesis:
'There is a relationship between the map reading scores and the driving error ratings of motorists'.
2 marks for a non-directional statement with both key variables that lacks clarity or has only one variable operationalised.
1 mark for a muddled statement with some reference to variables.
0 marks for expressions of aim/questions/causal statements or statements with only one condition.

Full credit can be awarded for a hypothesis expressed in a null form.

(c) **AO2 = 2**

1 mark for stating scattergraph or scattergram.

Plus

1 mark for explanation – because it shows a relationship between two variables.

(d) **AO2 = 3**

Possible content

- General pattern - if a participant scored highly on the map reading task then they are also rated highly on the practical driving task, (or vice versa)
- This suggests a person who has good map reading ability also has good driving skills so these spatial abilities are (positively) related/correlated

Accept other relevant comments

(e) **AO2 = 2 and AO3 = 4**

Level	Marks	Description
3	5 – 6	Outline of the problem is clear and coherent. Discussion of how the method could be modified is appropriate and effective. The answer is clear and coherent. Specialist terminology is used effectively. One modification in detail can access this level.
2	3 – 4	Outline of the problem is clear. Discussion of how the method could be modified is mostly appropriate and effective. There is some appropriate use of specialist terminology.
1	1 – 2	Outline of the problem is vague/muddled. Discussion of how the method could be modified either lacks detail or is muddled. Specialist terminology is either absent or inappropriately used.
	0	No relevant content.

Possible problems:

- Researcher bias – using one observer means objectivity/reliability/validity cannot be checked

Possible modifications:

- Increasing the number of observers of the driving task because then the data is less subject to individual bias – the observations could then be correlated
- Recording the driver performance so that the data is not lost but can be reviewed as often as required.

Credit other relevant information.

(f) **AO2 = 3**

Possible content

- The test determines the strength of a relationship between two variables which is what the researchers were looking for in their initial aim
- The data are in related pairs
- The variables under test are both ratings measured at the ordinal level.

Credit other relevant information

(g) **AO2 = 2 and AO3 = 2**

Level	Marks	Description
2	3 – 4	Explanation of an appropriate conclusion for this study is clear and mostly accurate. There is appropriate justification of the conclusion with reference to the critical values table. The answer is generally coherent with effective use of specialist terminology.
1	1 – 2	Some explanation of an appropriate conclusion is evident. There may be some justification of this with reference to the critical values table. The answer lacks accuracy and detail. Use of specialist terminology is either absent or inappropriate.
	0	No relevant content.

Possible content:

Conclusion

- The null hypothesis should be rejected and the alternative hypothesis accepted
- There is a significant (positive) relationship between the map reading ability and the driving ability of the participants
- Drivers who are skilled at map reading are also skilled at driving

Justification

- This relationship is a strong positive one as the calculated value of r_s of 0.808 exceeds the critical value for a two tailed test at p=0.05 where n=9 of 0.700.

44 **[AO3 = 2]**

1 mark for a brief explanation (must be explained rather than stated).

Plus

1 mark for elaboration in relation to consequences for the research / implications.

Possible limitations: questionable validity; lack objectivity (questions about own child).

Credit other relevant limitations.

45

(a) **AO2 = 4**

2 marks for identifying two factors that are relevant for use of the sign test: nominal/categorical data; test of difference; related design/repeated measures.

Plus

Up to 2 marks for application of these to the investigation described:

- Nominal data as patients are assigned to one of three categories – 'improved', 'deteriorated' or 'neither'.
- Testing for difference in the number of absences in the year following and prior to treatment.
- Repeated measures as the same patients' work records are compared before and after treatment.

(b) **AO2 = 2**

1 mark for identifying the correct value of s as 5

Plus

1 mark for explanation/calculation of how this was arrived at:

- The most commonly occurring sign is + (12) and the least frequently occurring sign is – (5). The 0s are disregarded.
- The total for the least frequently occurring sign is the value of s = 5

(c) **AO2 = 2**

1 mark for stating that the value of s (5) is not significant at the 0.05 level.

Plus

1 mark for explanation:

- The critical value is 4. As the calculated value is higher than/exceeds the critical value, the result is significant not at the 0.05 level.

Accept alternative wording

(d) **AO3 = 3**

Marks may be awarded for a single point that is expanded/elaborated or more than one point briefly stated.

1 mark only if there is no reference to the investigation described.

Possible points:

- Primary data are obtained ‘first-hand’ from the participants themselves so are likely to lead to greater insight: e.g. into the patients' experience of treatment, whether they found it beneficial, negative, etc.
- Secondary data, such as time off work, may not be a valid measure of improvement in symptoms of depression. Primary data are more authentic and provide more than a surface understanding: e.g. participants may have taken time off work for reasons not related to their depression.
- The content of the data is more likely to match the researcher’s needs and objectives because questions, assessment tools, etc. can be specifically tailored: e.g. an interview may produce more valid data than a list of absences.

(e) **AO1 = 3 and AO2 = 2**

Level	Marks	Description
3	4 – 5	Knowledge of the implications of psychological research for the economy is clear. Application to the investigation described is effective. The answer is coherent with effective use of terminology.
2	2 – 3	Some knowledge of the implications of psychological research for the economy is present but there is a lack of detail/clarity. Application to the investigation described is limited or absent. Terminology is used appropriately on occasion.
1	1	An implication of psychological research for the economy is briefly stated.
	0	No relevant content.

AO1 – possible content:

- Psychological research may lead to improvements in psychological health/treatment programmes which may mean that people manage their health better and take less time off work.
- Absence from work costs the economy an estimated 15 billion a year annually and much of this absence is due to 'mild' mental illness: e.g. stress, anxiety.
- Psychological research may lead to better ways of managing people whilst they are at work to improve productivity: e.g. research into motivation and workplace stress.
- 'Cutting-edge' scientific research may encourage investment from overseas companies into this country.

Credit other relevant points/implications, including examples not linked to psychopathology.

AO2 – application

- If research (such as the investigation described) suggests that depressives are better able to manage their condition following CBT and return to work, then it may benefit the economy to make treatment more widely available, improve funding, etc.
- Psychological research such as this plays an important role in sustaining a healthy workforce and reducing absenteeism.

Credit other relevant application points.

46 **[AO3 = 2]**

2 marks for a clear and coherent explanation of one reason.

1 mark for a partial or muddled explanation of one reason.

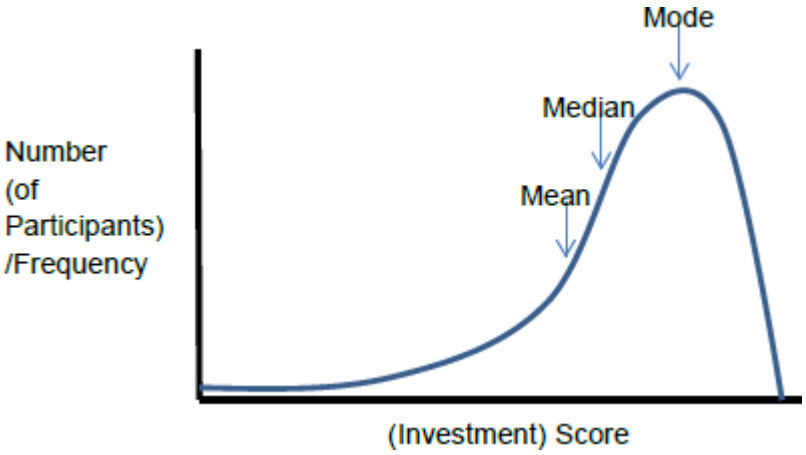
Possible content:

- prevents dissemination of irrelevant findings / unwarranted claims / unacceptable interpretations / personal views and deliberate fraud – improves quality of research
- ensures published research is taken seriously because it has been independently scrutinised
- increases probability of weaknesses / errors being identified – authors and researchers are less objective about their own work.

Accept other valid answers.

47

(a) **[AO2 = 3]**



Credit a rough sketch of a negatively skewed distribution as follows:

1 mark for shape of curve with tail to the left.

1 mark for axis labels – '(Investment) Score' on horizontal axis, 'Number (of Participants)' / 'Frequency' on vertical axis.

1 mark for positioning the mean, median and mode appropriately in relation to one another.

(b) **[AO2 = 1]**

1 mark for stating negative skew.

If graph sketched in (a) does not show a negative skew, credit answers that match the sketch given.

48

[AO3 = 2]

1 mark for a brief explanation (must be explained rather than stated).

Plus

1 mark for elaboration in relation to consequences for the research / implications.

Possible limitations: questionable validity; lack objectivity (questions about self).

Credit any other relevant limitation.

49 **[AO2 = 12]**

Level	Marks	Description
4	10 – 12	Suggestions are generally well detailed and practical, showing sound understanding of observational techniques. **All four** elements are present. There is sufficient information for most aspects of the study to be implemented with success. The answer is clear and coherent. Specialist terminology is used effectively. Minor detail and / or explanation sometimes lacking.
3	7 – 9	Suggestions are mostly sensible and practical, showing some understanding of observational techniques. **At least three** elements are present. Implementation of some aspects is possible. The answer is mostly clear and well organised. Specialist terminology is mostly used effectively.
2	4 – 6	Some suggestions are appropriate but others are impractical or inadequately explained. **At least two** elements are addressed. Implementation would be difficult based on the information given. The answer lacks clarity, accuracy and organisation on occasions.
1	1 – 3	**At least one** element is addressed but knowledge of observational techniques is limited. Implementation would be very difficult. The whole answer lacks clarity, has many inaccuracies and is poorly organised.
	0	No relevant content.

Four elements of design to be credited:

- **The task for the participants** – detail of what the men and women in the study will have to do. This must go beyond ‘give a presentation to an audience’.
- **The behavioural categories to be used and how the data will be recorded** – detail of specific and observable behaviours to be recorded. This must go beyond the idea of global constructs such as ‘body language’ or ‘gesture’. Also detail of recording method to be used, eg record sheet.
- **How reliability of the data collection might be established**, eg using two observers / raters and comparing separate recordings; statistical comparison of data from both observers / raters.
- **Ethical issues to be considered**, eg specific or more general ethical considerations as applied to this study – protection of welfare, confidentiality and deception, respect or integrity.

Examples of possible tasks:

- presentation of findings from a school project
- presentation on ‘My Hobby’
- presentation on ‘My Holiday’.

Examples of suitable non-verbal behaviours include:

- arm movements
- smiling
- speech hesitations
- pointing etc.

50

(a) **[AO2 = 2]**

1 mark appears to support the nature side of the debate.

Plus

1 mark because the concordance rate is stronger in the identical twins where there is greater genetic relatedness (or nurture must also play a role – not 100% concordance).

Full credit can be awarded to answers which argue for mathematical ability being partly due to nurture as both percentage concordance rates are less than degree of genetic relatedness.

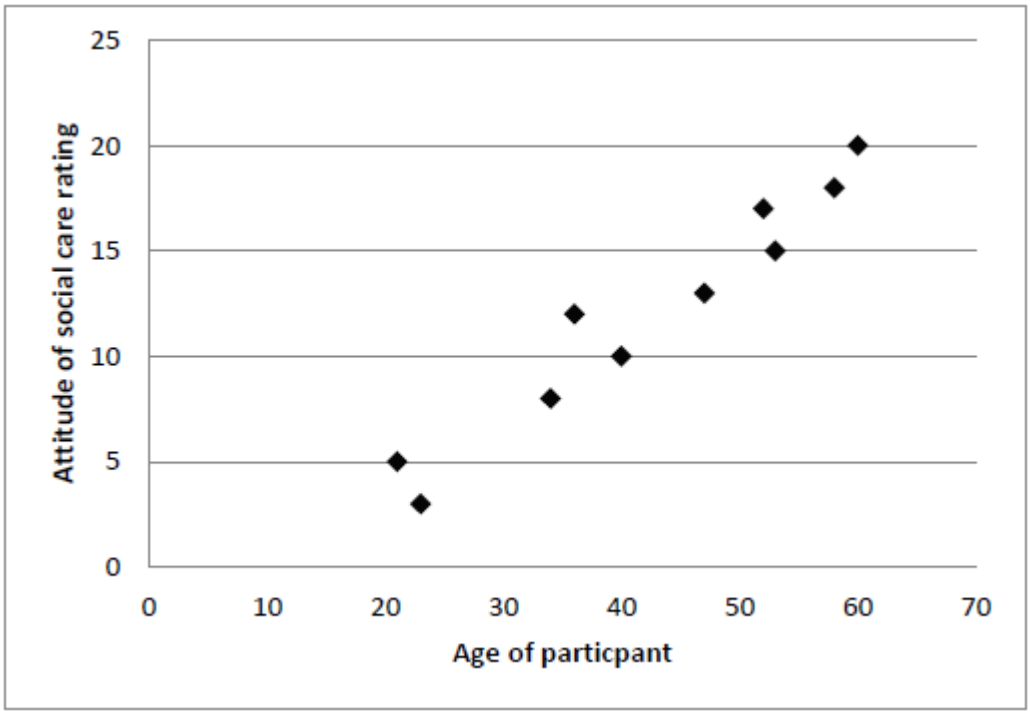
(b) **[AO2 = 4]**

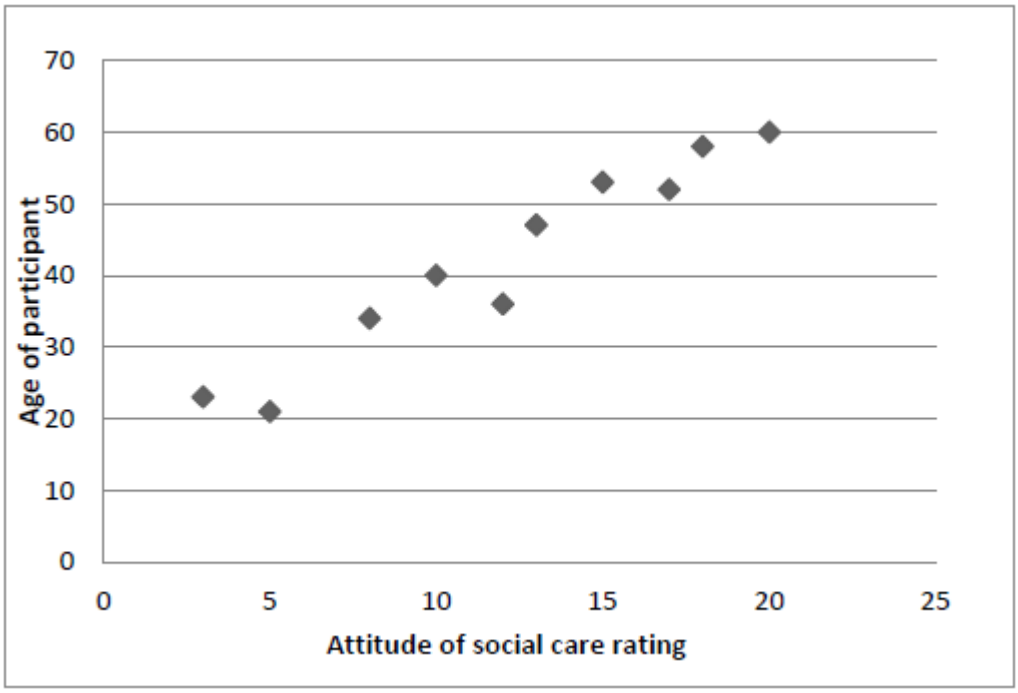
Level	Marks	Description
2	3 – 4	Answer focuses clearly on concurrent validity. How a correlational test would be used to determine the relationship between the two sets of scores is clearly described with reference to calculation of a correlation coefficient and need for a significant positive correlation.
1	1 – 2	Answer focuses on validity. How a correlational test would be used to determine the relationship between the two sets of scores is partly described. The answer lacks accuracy and detail.
	0	No relevant content.

Content:

- concurrent validity would involve correlating the results on the maths test with results for the same group of people on an established maths reasoning test
- A Spearman’s rho or Pearson’s r test should be used for the two sets of test results
- if the mathematical ability test is valid then there should be a significant positive correlation between the two sets of test scores at the 0.05 level.

51 (a) **Marks for this question: AO2 = 3**





3 marks for the following points:

- Axes correctly labelled as Age of participant and Attitude to social care rating.
- Scales are suitable.
- Points plotted accurately.

(b) **Marks for this question: AO2 = 2**

2 marks for: there is a **positive relationship** between age and interest in social care issues / as people get older their interest in social care increases (1) this is because as the values on one co-variable increase, so do the values on the other co-variable (1) OR as age increases so does attitude to social care **rating / score**.

(c) **Marks for this question: AO2 = 2**

1 mark for knowledge of an investigator effect – this is when the person collecting the data has knowledge of what the research aim is / traits and that knowledge / those traits affect the data obtained.

1 mark for a brief explanation of how investigator effects might have occurred in this study.

If the researchers believed that older people would be more interested in social care they could have just given scores based on the age of the person.

(d) **Marks for this question: AO3 = 2**

2 marks for explaining how investigator effects could have been avoided in the study. The answer needs to explain what could be done and how that would decrease / eliminate the effect.

Possible content:

- Discussion of separate observation by the two researchers and comparison – inter-rater reliability.
- Having 'blind' rating of the discussion by someone who is unaware of the aim or research hypothesis.
- Filming the discussions so there is a permanent record that can be checked by peer review of the data to confirm the scores / ratings.

Credit other relevant procedures.

(e) **Marks for this question: AO2 = 4**

Level	Marks	Description
2	3 – 4	Explanation of how closed and open questions are beneficial is clear. The answer is generally coherent with effective use of terminology.
1	1 – 2	There is limited / partial reference to the benefit(s) of closed and open questions. The answer lacks accuracy and detail. Use of terminology is either absent or inappropriate. **OR** answer only refers to either closed or open questions at Level 2.
	0	No relevant content.

Possible content:

- Closed questions would present participants with options for their response so the researchers would be able to collate and display the information collected easily.
- Closed questions make it easy to compare specific response to questions the researchers wanted answered – they can be sure there will be certain information because they have restricted the options to include that information.
- Open questions allow respondents to interpret the question as they wish to and develop their response with detail or depth – so there is lots of information received.
- Open questions allow the researchers to pursue a line of enquiry that they may not have predicted but which comes to light because of a response by an interviewee.

Credit other relevant procedures.

(f) **Marks for this question: AO2 = 3**

- **1 mark** for an appropriate open or closed question – requiring information about a social care issue.
- **1 mark** for correct identification of this as an open or closed type of question.
- **1 mark** for a suitable explanation for why the choice was appropriate – this could relate to producing a type of data (closed – ease of analysis, open – lots of detail or depth to response / allows respondent to elaborate her / his reasoning for the response given) or it could focus on an issue of social care introduced by the candidate and not in the stem.

(g) **Marks for this question: AO1 = 2 AO2 = 2**

AO2
1 mark: the responses to the open questions in the interview constitute qualitative data.
Plus
1 mark: the attitudes ratings **AND / OR** the collated responses to the closed questions in the interview constitute quantitative data.

AO1
1 mark for an explanation of how the responses to the open questions is qualitative data ie is non-numeric / descriptive / retains detail of actions / thoughts / feelings.
Plus
1 mark for an explanation of how the ratings / collated responses to closed questions is quantitative data ie numerical such as a score / behaviour is represented in the form of a score on a scale.

(h) **Marks for this question: AO3 = 4**

2 marks for each explanation of how the chosen ethical issue could be dealt with.

1 mark for a brief muddled explanation.
2 marks for a clear explanation.

Consent – to be part of what is in essence two studies. Participants should be forewarned – a briefing.

Protection from harm – at the end of participation all will have to be fully aware that they were rated for their social care interest and a low score might indicate they are 'uncaring'. They may wish to withdraw their data.

Right to withdraw – being made aware that they can at any time stop participating and at the end of their participation they can withdraw detail of their behaviour in the research.

The explanation must demonstrate an appreciation that people should be dealt with, with respect and competence.

Credit other relevant ethical issues.

52

(a) **[AO2 = 2]**

2 marks for identification of dependent variable operationalised: number of verbal errors.

1 mark for dependent variable not operationalised: verbal errors or fluency or mistakes.

(b) **[AO2 = 3]**

3 marks for an appropriate non-directional (or directional) operationalised hypothesis:

'There is a difference in number of verbal errors made by participants who perceive / think / believe there are 5 listeners (there is a small audience) and by participants who perceive / think / believe there are 100 listeners (there is a large audience)'.

2 marks for a statement with both conditions of the IV and a DV that lacks clarity or has only one variable operationalised.

1 mark for a muddled statement with both conditions of the IV and a DV where neither variable is operationalised.

0 marks for expressions of aim / questions / correlational hypotheses or statements with only one condition.

Full credit can be awarded for a hypothesis expressed in a null form.

(c) **[AO2 = 3]**

1 mark for identification of **one** appropriate extraneous variable.

Plus

2 marks for explanation of why the variable should have been controlled – for full marks this should include clear explanation of how it would have affected the DV. Award one mark only for muddled or incomplete explanations, eg unelaborated reference to 'avoiding confounding'.

Appropriate variables: can be controlled and need to stay constant to avoid affecting the dependent variable, eg same article / conditions / instructions for each participant.

Do not credit gender (this is controlled) or time to complete task (cannot be controlled).

(d) **[AO2 = 2]**

2 marks for clear and coherent explanation of one advantage of using a stratified sample in this study.

1 mark for a muddled answer with a relevant advantage and some explanation in relation to the study.

Possible advantage: ensures that this sample is truly representative because different types of people (males / females) working in this company are represented in the sample in the correct proportions.

Accept other relevant advantages.

(e) **[AO2 = 3]**

1 mark for each point as follows:

Manual method:

- put all 60 male names in a hat (or similar)
- determine the proportion of males needed to mirror the number of males in the target population as follows: 60%
- calculate 60% of 20 = 12 and draw out 12 names.

Random number table or computer method:

- assign each of the 60 men a number between 1 and 60
- determine the proportion of males needed to mirror the number of males in the target population as follows: 60%
- calculate 60% of 20 = 12 and moving horizontally or vertically through random number tables find 12 numbers between 1 and 60 for the sample **OR** generate 12 numbers between 1 and 60 using random number generation function on computer.

(f) **[AO2 = 4]**

Marks for a clear description of a practical way of randomly allocating the 12 men and 8 women to the two conditions as follows:

- give each man a number 1 – 12 (1 mark)
- put 12 numbers in a hat (1 mark)
- assign first six numbers drawn to Condition A with the remainder for Condition B (1 mark)
- repeat process for women – eight numbers in the hat and draw four for Condition A and remaining four go to Condition B (1 mark).

Accept other valid descriptions that would be practical and produce the same outcome.

53

(a) **AO2 = 3**

1 mark for each of the following points:

- The total observation time for each parent was 10 minutes.
- The psychologist made 20 observations for each parent.
- To generate 20 observations for each parent she must therefore have recorded her observation every ½ minute or every 30 seconds.

(b) **AO2 = 2**

1 mark for the correct answer: 30%.

Plus

1 mark for showing correct workings: 12 divided by 40 multiplied by 100.

(c) **AO2 = 1**

1 mark for primary data.

54 (a) **[AO2 = 2 and AO3 = 4]**

Level	Marks	Description
3	5 – 6	Conclusions in respect of both means and standard deviations are presented with clarity. Understanding of the relevance of each statistic is demonstrated. Justifications for each make good use of the values given.
2	3 – 4	Conclusions and justification in respect of both means and standard deviations are relevant, but there is some lack of clarity in both. Or, one is done well and justified appropriately (most usually this will be the mean).
1	1 – 2	One conclusion is drawn or two are partially correct. Any justification is limited. The answer lacks clarity.
	0	No relevant content.

Means

- Conclusion: when people believe they are presenting to a large audience they are less fluent in their spoken communication than when they believe the audience is small (or vice versa).
- Justification / Application: this is supported by the difference in the mean fluency scores which show more verbal mistakes (on average 6 more mistakes) when the audience is believed to be large (or vice versa).

Standard deviations

- Conclusion: performances of participants in Condition A where audience is believed to be small are less varied / dispersed / spread out than in Condition B where audience is believed to be large (or vice versa).
- Justification / Application: lower SD in Condition A suggests that individual performances in Condition A were more similar to each other and / or all quite close to the mean of 11.1.

(b) **[AO3 = 3]**

1 mark – this would be an improvement because the SD is a measure of dispersion that was less easily distorted by a single extreme score.

Plus

1 mark – one that takes account of the distance of all the verbal error scores from the mean.

Plus

1 mark – not just the distance between the highest verbal error score and the lowest verbal error score.

(c) **[AO2 = 4]**

1 mark for naming the t-test for independent / unrelated groups or a Mann-Whitney test.

Plus

Up to 3 marks for explanation for unrelated t-test. Credit relevant points as follows:

- can assume interval data because verbal errors can be assumed to be of equal size (ie one verbal error is equivalent to any other verbal error)
- the experimental design is independent groups
- the psychologist is looking for a difference between the two conditions.

OR

Up to 3 marks for explanation for Mann-Whitney test. Credit relevant points as follows:

- data should be treated as ordinal. Cannot assume interval data because verbal errors cannot be assumed to be of equal size (ie one verbal error is not equivalent to any other verbal error)
- the experimental design is independent groups
- the psychologist is looking for a difference between the two conditions
- SDs are quite different.

(d) **[AO1 = 2]**

2 marks for a clear and appropriate definition as follows:

This means that there is a less than 5% likelihood that this difference would occur if there is no real difference between the conditions **OR** the researchers would have a 95% confidence level.

1 mark for a less clear answer which shows some understanding, eg this means the researcher can conclude that the difference was not due to chance.

Accept any other valid answer.

(e) **[AO2 = 2]**

2 marks for a clear and detailed explanation applied to this study.

1 mark for a partial or muddled explanation or one that is only loosely applied to the study.

Credit answers based on any type of validity. Most answers will refer to either face or concurrent as follows:

- asking other people if verbal errors are a good measure of verbal fluency (face validity)
- giving participants an alternative / established verbal fluency test and checking to see that the two sets of data are positively correlated (concurrent validity).

55 **[AO3 = 4]**

Level	Marks	Description
2	3 – 4	Discussion is relevant, well developed and well explained, with focus on improvements to be had by using controlled observation. The answer is generally coherent with effective use of specialist terminology.
1	1 – 2	Discussion is relevant although there is limited explanation / development and / or limited focus on the issue of improvement. Specialist terminology is not always used appropriately. Award one mark for answers consisting of a single point briefly stated or muddled.
	0	No relevant content.

Possible content:

- controlled environment affords the opportunity for control of extraneous variables
- examples of extraneous variables that might be controlled and how / why they could affect the outcome of a study if not controlled
- exclusion of extraneous variables allows for greater inference about cause and effect
- exclusion of extraneous variables means researcher can replicate the observation to check for reliability of the effect.

Credit other relevant discussion points.

56 (a) **[AO3 = 2]**

Award **1 mark** for outline of a positive correlation / as one variable increases, so does the other.

and

1 mark – It would not be appropriate because correlation only shows a relationship between the two variables, not cause.

(b) **[AO3 = 2]**

2 marks for a clear, coherent outline of an appropriate way of dealing with a relevant ethical issue.

1 mark for a vague / muddled or incomplete outline of an appropriate way of dealing with a relevant ethical issue.

0 marks if the ethical issue is irrelevant or the way is inappropriate.

Relevant issues would include:

- asking for consent
- preferably written and on more than one occasion
- offering the right to withdraw from the study
- maintaining confidentiality
- treating with respect.

(c) **[AO3 = 4]**

Award **one mark** for each of the relevant points below:

- the same participants would complete the sleep questionnaire on more than one occasion
- each participants' scores from the first occasion should be correlated with his / her results from the later occasion to be shown on a scattergraph to describe the correlation, with scores from the first test plotted on one axis and the scores from the second test plotted on the other axis
- the strength of the correlation should then be assessed using either a Spearman's rho test (or a Pearson's r test)
- the degree of reliability is then determined by comparing the correlation with the statistical table to determine the extent of correlation – there should be a (strong) positive correlation between the two sets of scores.

57 **[AO1 = 2]**

1 mark each for any two of the following points:

- observers / researchers decide on a specific event relevant to the investigation
- relevant event is recorded every time it happens
- in this investigation this may be every time a child in the playground is approached by / talks to / plays with another.

Students may refer to the investigation described in their answer, though this is not required by the question.

58 **AO1 = 2**

1 mark for stating that overt observation is where the observer is clearly visible (not hidden from view).

Plus

1 mark for explanation – people being observed know that they are being observed.

59 (a) **[AO2 = 2]**

1 mark for naming the mean.

Plus

1 mark for justification: the mean is the most sensitive method as it takes all the scores in each data set into account OR there are no anomalous results / outliers / freak scores in either set of scores, so the mean will not be distorted.

(b) **[AO2 = 4]**

Full credit can be awarded for answers based on the mean or the median.
A maximum of **2 marks** can be awarded for answers based on the mode.

Using the Mean

- **For 4 marks**, the **mean** is accurately calculated for both conditions (Group A = 5.6, Group B = 12.5) and calculations are included for both groups, ie totals in both conditions divided by 10 (number of scores).
- **For 3 marks**, there are two correct means and one set of calculations or vice versa.
- **For 2 marks**, there are two correct means and no calculations, **OR** one correct mean with calculations **OR** two sets of calculations but no correct mean.

- **For 1 mark**, there is one correct mean or one set of calculations.

Using the Median

- **For 4 marks**, answers for each condition are correct (Group A = 5.5, Group B = 12.5) and for each condition scores are arranged in ascending order with middle values indicated.
- **For 3 marks**, there is one correct median and two sets of scores correctly arranged as calculations, or vice versa.
- **For 2 marks**, there are two correct medians and no calculations, or one correct median and one set of scores correctly arranged as calculations.
- **For 1 mark**, there is one correct median or one set of scores correctly arranged as calculations.

Using the Mode

- **For 2 marks**, there are correct modes for each group (Group A = 4, Group B = 11 and 14).
- **For 1 mark**, there is one correct mode.

(c) **[AO2 = 2]**

1 mark for stating that this is due to retroactive interference.

Plus

1 mark for either of the following explanation / elaboration points:

- because the material is similar in both conditions
- new / recently learnt / acquired information has disrupted / interfered with / affected the recall of old / previously learnt / acquired information
- response competition has occurred.

60 (a) **AO2 = 2**

2 marks for a clear and coherent explanation of the usefulness of the standard deviation in this study.

1 mark for a weak or muddled answer in which the impact of the difference in the SDs is alluded to.

- Useful to inform about the spread of scores.
- Indicates participant variables – as a group the people in Condition 1 are quite different / are more variable than those in Condition 2.

Credit answers which suggest that the SDs can be used to look for similarity or differences in variance.

(b) **AO3 = 2**

2 marks for a clear, coherent outline of a relevant problem.

1 mark for a weak, muddled or very limited outline.

Possible problems:

- Direct observation of memory is not possible and must be inferred from the results / behaviour of the participants – this inference could be mistaken.
- The task given is rarely how normal memory functioning occurs because it is specifically designed to make measurement possible – the researcher therefore collects data that is only related to memory processing under experimental conditions.

Credit other valid problems.

5

Better students were able to extract the relevant information from the table and use it effectively. They considered the baseline of 65% (no confederates) and then compared it to the other two conditions, 92.5% and 10%, which showed the power of confederates. They were also able to comment that in fact the disobedient confederates seemed to have more power than the obedient ones, perhaps by providing role models or allies.

Since this question only asked about the confederates, reference to the third condition (experimenter in different room) was not credit-worthy. This illustrates the need for students to read the question carefully and select and shape their answer accordingly.

A significant number of students confused conformity and obedience and used these terms interchangeably. They seemed to forget that this data referred to Milgram's experiment into obedience and seemed to think that the confederates were a majority.

6

Most students were able to state correctly that the correlation was a positive one; as scores on a stress questionnaire increased, so did the number of days off work through illness. Credit was also given to any reference to strength (moderate or strong) as well as to the flattening out of the graph.

While most students were able to consider the main weakness of correlations, few could correctly outline a strength. In fact many contradicted themselves by saying that a strength was the ability to show cause and effect on a graph and the weakness was that a causal relationship can never be assumed. Better answers referred to the strength as being the ability to study the relationship between variables that occur naturally; or to measure things that cannot be manipulated experimentally.

7

(a) Students who scored well often focussed on the anonymity of questionnaires, the lack of investigator effects or the time advantage where questionnaires could be simultaneously completed. Whether students gained full marks depended on how effectively they were able to explain the advantage they had identified. Better answers compared questionnaires to interviews, or referred to the relatively large number of adults in this study. Some students referred to the advantages of analysing data from questionnaires which was not the focus of the question.

(b) Most responses explained the term qualitative data appropriately. A few students described quantitative data; given that the word quantitative can be so easily aligned with number, it is surprising that students get muddled about these terms.

(c) Most responses were appropriate, although a number of questions provided would have produced numerical data (eg how long ?x2, how many ?x2) or categorical answers (usually yes / no responses).

(d) Although most students had no difficulty in identifying two ethical issues, many students were less successful in providing suitable suggestions for how one of these issues could be dealt with. Some students just re-stated the ethical issue. Other students filled up the answer space by explaining how both ethical issues could be dealt with, leaving the examiner to decide which was the more credit-worthy answer.

8

(a) Most students were able to state the independent variable, though some incorrectly emphasised the location in which words were 'learnt' rather than the location in which they were 'recalled'.

(b) Most students were able to state the likely outcome, that participants in Condition 1 would outperform those in Condition 2, and could link this effectively to the notion of context cueing recall. However, for full marks the 'retrieval failure', experienced in Condition 2, and the reason for it, had to be discussed. This final requirement eluded many students who focused on 'recall' rather than 'retrieval failure'.

(c) Answers to this question often lacked precision, for example, it was necessary to 'put the names of **all** the participants in a hat'. It was also necessary to describe how the selection for conditions 1 and 2 would be made, and this information was often vaguely expressed or absent.

(d) Most answers gave a reasonable basis for 'matching' such as 'IQ', but failed to deal with the issue of 'pairs' and how to allocate them to the different conditions.

9

This question required an explanation of peer review. Just under half of students achieved the full two marks and most other answers demonstrated a basic understanding that the process involved getting an expert to look at the report. However, many failed to understand that peer review occurs prior to publication.

10

(a) Students who had carefully read the stem material noted the reference to the age of the participants in the study being estimated by the researcher. They were thus more likely to recognise that one aim was to investigate the effect of age on eyewitness testimony.

(b) Students usually responded by referring to the benefits in terms of validity or fewer demand characteristics. Not all students elaborated their point in order to achieve a second mark.

(c) This question was poorly answered because many students could not accurately identify opportunity sampling and they confused this technique with random sampling. It is important for students to grasp that random refers to where everyone in the sampling frame has an equal chance of being selected and this does not apply in opportunity sampling. The word random has a different meaning in psychology compared with everyday usage.

(d) This was generally answered well with many answers showing a good understanding of extraneous variables. Students were very inventive referring to the effect of both situational (weather, noise levels) and participant (in a rush, alcohol consumption, mental illness) variables.

11

This question proved accessible to weaker students allowing them to access some marks and also provided plenty for the stronger to discuss. Many began by identifying problems with the volunteer sample and the use of an interview to measure happiness in the proposed experiment. Stronger students identified the lack of operationalisation and control at the heart of the proposed study.

Quite a lot of the modifications suggested were weak (for example 'do a random sample') often demonstrating that the students had not thought about the practicalities involved. Better answers were structured around identification and discussion of a limitation (e.g. lack of control) followed by a well-developed argument about different ways to improve this. Some schools / colleges have clearly prepared their students well and many showed an impressive understanding of experimental design. Others who struggled with the question, failed to read the instructions and simply identified a string of speculative limitations such as the; possible lack of a consent form, collecting very few marks often suggesting "improvements" that revealed their lack of practical experience of designing research and collecting data. Teachers are encouraged to do some practical work with students and encourage them to plan 'thought experiments'. It was clear that some students were very familiar with designing experiments and they had a strong advantage here.

12

(a) This was generally answered well.

(b) Students failed to notice that the appropriate conclusion the researchers could draw was that the participants **believed / said** they would help. The actual behaviour was not measured by the questionnaire.

(c) Students are still poor at writing clear hypotheses, with both conditions of the IV present and a measurable DV. Many answers were aims and the expression 'more likely' was often used, as was reference to Condition 1 and Condition 2.

(d) It was evident from many answers that students could not spot obvious variables on which researchers would focus their attention. Instead, answers included reference to the 'heating', with the idea that the room was likely to become so hot / cold that participants would lose the ability to hear what was going on.

(e) Many students lost a mark for this question because they produced a generic explanation for the suitability of the experimental design and did not relate their answers to the study described.

(f) This was not answered well. Many explained what random sampling is or gave advantages or limitations of the method. It was rare to see an answer that identified what must go into the 'hat / random-number generator' and what would then happen. There was a great deal of confusion with random allocation.

(g) Most students suggested a bar chart and could produce a sensible reason for choosing such a display.

(h) Unfortunately, many students did not recognise that, if asked about debriefing participants, then the focus is on telling people everything about the study in which they have just taken part. This is especially important in studies using independent groups design, as participants are not only unaware of the aim of the investigation, but also that other people performed under different conditions. Too many answers merely concentrated on ethical points.

13

This was poorly answered with less than 5% of students achieving the full 4 marks available. Most students had a basic idea that peer review prevents weak or fraudulent research getting into the public domain but could say little beyond this. There were many weak answers referring to ‘checking reliability and validity’ or referring to replication which demonstrated a worrying lack of understanding of the functions of peer review. A number of students became side-tracked into the role of peer review in grading university departments. Some weaker students confused peer review with an ethics panel. Stronger students were able to discuss the importance of peer review where research has clear practical applications, such as therapeutic interventions.

14

(a) Hypothesis writing continues to be a problematic for many students, despite the requirement to do this at AS level. Less than a third students achieved the full 2 marks available and a further third scored no marks at all, having mistakenly written a directional hypothesis or one which predicted a difference between happiness scores and intelligence scores. Many responses were lacking in clarity or failed to operationalise the variables sufficiently. The best answers were concisely and clearly worded such as “There is a correlation (relationship) between pupils’ scores on a standardised intelligence test and their scores on a questionnaire measuring happiness”, which achieved the full 2 marks.

(b) Most students were able to identify an appropriate alternative method to collect data about happiness, the most popular choices being interviews, observations and diary studies. Some were able to provide a clear explanation of why this would be better, but weaker students became side-tracked into describing the possible method in detail (i.e. observational categories that might be used) and lost focus on the question of comparison. Better students were able to refer to precise advantages of their chosen method over questionnaires which were contextualised in relation to measuring happiness.

(c) This question required students to give two reasons why Spearman’s rho was used to analyse the data. Almost all students were able to accurately identify one reason, which is encouraging!

(d) Weaker students still struggle to interpret critical and obtained values appropriately. About a third of students managed to pick up the full 3 marks. Most of the remainder were able to say the result was significant, gaining one mark but went on to select the incorrect critical value from the table. A small number erroneously compared the correlation coefficient (0.42) with 0.5 demonstrating misunderstanding throughout the entire reasoning.

(e) This question required students to interpret a further correlation coefficient (which was statistically insignificant) and put both pieces of information together to draw an overall conclusion to the reported study. This proved challenging and less than 5 % of answers achieved the full 4 marks here. Many students were able to identify that the scores demonstrated different kinds of relationships (positive and negative) but were unable to take this further and think about possible explanations. The better answers focused on the inability to establish cause in correlational research and the role played by other variables, in this case age. Some made use of their knowledge about reliability which was creditworthy.

15

This question was answered well by the majority of students as most were able to refer to three ways in which the study was scientific. Where full marks were not awarded, this was usually because students had only named or identified ways in which the study was scientific rather than providing an outline. Some students only gained one mark, as scientific principles were named but not explained in relation to the study.

16

Most students were able to state correctly that the correlation was a positive one; as scores on a stress questionnaire increased, so did the number of days off work through illness. Credit was also given to any reference to strength (moderate or strong) as well as to the flattening out of the graph.

While most students were able to consider the main weakness of correlations, few could correctly outline a strength. In fact many contradicted themselves by saying that a strength was the ability to show cause and effect on a graph and the weakness was that a causal relationship can never be assumed. Better answers referred to the strength as being the ability to study the relationship between variables that occur naturally; or to measure things that cannot be manipulated experimentally.

17

(a) There were some competent answers to this question, where students knew what was involved in content analysis and how it could be applied in this instance. Some showed confusion between studying the children and studying the diaries. A substantial number of students did not attempt this question at all or scored no marks.

(b) There were some competent answers, especially where students obviously understood the procedure of content analysis. Students were able to focus on the problem of asking the mother to keep a diary and could suggest possible difficulties with demand characteristics, social desirability and problems of recording regularly and in sufficient detail. Some students did not read the stem carefully because their suggestions that the mother would not be at nursery to record the child's behaviour or the recording would be retrospective after two weeks, and thus subject to unreliable memory, were not limitations of this study.

18

Most students could identify volunteer / self-selected sampling. A few responses incorrectly referred to opportunity or random sampling.

Many responses scored the full two marks. Perhaps this was partly due to the fact that the mark was given whether or not the DV was operationalised. Some students had no idea what IV or DV referred to, and a significant minority got them the wrong way round.

Good answers often referred to the diminished likelihood of demand characteristics with independent groups design, or the lack of order effects due to participants taking part in only one condition. A few answers incorrectly muddled the two, ie there would be fewer demand characteristics because participants would not suffer practice effects. Although it was correct to say there would be no order effects, it was not correct to say there would be no demand characteristics. There could potentially be fewer.

Generally students noticed that the stem gave the opportunity to refer to the potential for informed consent, though many had not addressed the right to withdraw. However, some students ignored the requirement to focus their response on the stage of recruiting participants, so answers referring to debriefing were not relevant. The word debriefing was inappropriately used when some students were referring to briefing.

19

This question was generally answered well with most students able to demonstrate some understanding of the problems in making generalisations. Many students referred to generalising from small / unrepresentative samples and the problem of the lack of ecological validity when generalising from laboratory experiments. Where students were not awarded full marks it was usually because either answers were not fully developed or only one problem was outlined.

20

Some students failed to read the stem carefully and found it difficult to answer the question. It was insufficient to identify a technique as 'mental re-instatement'. Students were required to suggest the context was being mentally re-instated. Where students chose to write about mental re-instatement of the context, they needed to remember that they should refer to the context of watching a film of a violent crime. Many focussed on creating mental context within the film rather than of watching the film. Students who chose 're-instate the context' sometimes muddled their instructions to give instructions to 'recall everything'. Generally those who chose to focus on 'recall from a changed perspective', or 'recall in reverse order' found it easier to write instructions for the participants. A few students did not write instructions in direct speech and were limited to a maximum of one mark for the instructions.

Where students did not score full marks for this question, they had usually failed to refer to all of the figures in the table, or to draw any conclusion from the figures.

21

(a) The majority of students gained the mark for this question by outlining the relationship. A small number of students simply stated 'positive correlation' which did not gain the mark as reference to the variables was required for an outline.

(b) Most students were able to name a correct test (Spearman's or Pearson's) although not all could justify the choice of test with reference to the levels of measurement. Some answers simply stated ordinal or interval, which was not enough as it was not clear which variable was being referred to. One could argue that the empathy scale was not equal interval and therefore should be treated as ordinal data. However, the hours spent reading fiction per week could be treated as interval. Some suggested an incorrect test e.g. Wilcoxon or Chi-square.

(c) Most students gained at least one mark on this question as they could outline a type of validity. However, some just named a type which was not enough for even one mark. Where students gained both marks, they generally referred to either face or concurrent validity (although occasionally to predictive or criterion validity). For full marks, answers had to explain how the type of validity would have been implemented. Many failed to do this part of the question and thus gained only half marks. Some students confused validity and reliability, often referring to split-half or test-retest, and a noticeable number tried to answer the question with reference to pilot studies.

(d) Most students were able to identify a limitation of the study and answered the question well. The majority of the problems identified referred to the sample. A number of answers referred to ethical issues and therefore gained no credit.

(e) Answers which simply stated that correlation looks for a relationship and experiments investigate differences (or similar) gained only one mark. Students that did access further marks were able to explain issues around ethics and manipulation of variables in such a study.

(f) This question was answered well by those students who had clearly had practice at designing and implementing their own investigations. It was pleasing to note that some students gained full marks on this question and it was evident that some schools and colleges had prepared students well. Students who did not gain high marks on this question usually:

- wrote very brief answers or ran out of time
- did not address all 5 bullet points. This was a shame because a number of the points were often addressed comprehensively but then students failed to address certain sections, in particular materials. The best answers addressed each point in turn in a structured and comprehensive manner
- provided little or no justification. The question specifically asked for justification of design decisions, but some answers gave little or no justification, e.g. for design, sampling techniques etc. Without justification the question was not properly addressed and this led to lower marks.

22

Many students answered this question well, often gaining full marks. There were a number of students who clearly demonstrated knowledge of how psychodynamic psychologists neglect the rules of the scientific approach but failed to link the points made to the stem. The question clearly required 'reference to the study above' and without linking the concepts (e.g. implications of generalising from small samples, bias, subjectivity, etc) to the scenario outlined in the 'dream diary', no marks could be accessed. This was a shame and is indicative of the need of students to read the question carefully and address the requirements of the question.

23

This was the first time on this paper that students had been required to sketch a graph and it was done very well indeed, in spite of the fact that many did not have a ruler or pencil. The main weakness was in correctly labelling the y-axis.

24

(a) In this part of the question students were asked to 'Describe one way…', yet descriptions of a way of investigating were sometimes very vague. Of the many students who chose to write about the Strange Situation, a good number failed to mention a stranger, the key element of the procedure. Although the study did not have to be identified by name, most were identifiable from the detail of the method. A small number of answers gained no credit because they consisted of little more than vague references to behaviours such as imitation, cuddling or motherese. A few students used animal studies despite the explicit instruction not to do so.

(b) There were many well-applied three-mark answers to this part. Responses consisting of generic evaluation points without explicit application to the study were limited to one mark.

25

(a) This was generally answered well although a surprising number of students failed to include both elements of ‘order’ and ‘middle’ in their answers. Some students just presented the partial formula of n(+1) / 2 with no explanation of what this meant.

(b) Students failed to notice that the DV was an ‘estimation’ not the median score.

(c) Students are still poor at writing clear hypotheses with both conditions of the IV present and a measurable DV. Many answers were aims and the expression ‘more likely’ was often used.

(d) Students were still confused about the term ‘experimental design’, often thinking it referred to conducting a field or laboratory experiment. Others did not realise that in both the first and second parts of the investigation the researcher was looking for a difference between the performances of men and women meaning that on both occasions the experimental design was independent groups.

(e) Overall, there was a lack of clarity in some responses to this question. Unless students made it clear exactly how the researcher would end up with a sample of 15 men and 15 women from the factory, they could not gain full credit. References to random number generators / computers often did not describe how the names would become numbers.

(f) This was answered well with the majority of students gaining at least one mark.

(g) This was answered reasonably well with many students gaining at least 2 marks.

(h) Unfortunately, once again, students did not recognise that if asked about debriefing participants then the focus is on telling people everything about the study in which they have just taken part. Too many answers merely concentrated on ethical points or added information which would not normally be available at debriefing, such as, details of all the results. Also, some accounts stated that participants could withdraw from the study, even though they had finished participating.

(i) Students often limited themselves to one mark because they failed to develop the reason they proposed as an explanation for why researchers conduct pilot studies. In general, it is accepted that the purpose of a pilot study is to identify possible flaws so that they can be eliminated / to ensure the data collected is appropriate / to ensure time is not wasted.

26

(a) The majority of students could correctly identify the IV, though some gave the DV.

(b) Few students gained both marks for this question. A large number of students were under the impression that all quasi-experiments are conducted in a natural setting, such that extraneous variables cannot be controlled. Many students recognised that quasi-experiments do not involve manipulation of the IV, but failed to apply this understanding to the context of the question, that OCD would be pre-existing or naturally occurring.

(c) Many of the variables that students suggested were appropriate, with most opting for age as a criterion for matching participants. A few students, incorrectly, wrote about whether they had OCD or not despite having mentioned this in part (a).

(d) Many students believed that matched pairs designs remove or eliminate participant variables rather than merely controlling them, and yet students were unable to link this advantage to the stem of the question: that the researcher could be more confident that the results found were due to the existence of OCD, than other differences between participants.

27

(a) This question required a definition of content analysis which proved challenging for many students. Almost half of the answers achieved no marks at all. This was made more remarkable by the fact that most were able to gain some marks on part (b) where they were asked to explain how to carry out a content analysis for the data in question.

(b) Most students were able to gain some marks here despite poor performance on part (a) and could identify in a basic way how to carry out a content analysis on the video recordings. Some were able to provide a clear description of the process but few appreciated that behavioural categories need to come from somewhere, whether that is from pilot work or previous research.

(c) Most students were able to identify an appropriate method of testing the reliability of the content analysis and collect at least one mark. The most popular answers were test-retest and inter-rater reliability. Many failed to gain the three marks available as their explanation of how the method of checking reliability would be carried out lacked detail. A few students became side-tracked into improving reliability and a small number used split half which was inappropriate in relation to content analysis and gained no marks.

(d) This question required students to explain why a repeated measures design was used in the experiment. Many students provided a basic answer referring to the need for less participants or the removal of individual differences but were unable to provide further explanation of why this would be important in this experiment. Students who thought about the scenario and elaborated their explanation with reference to reaction times, concentration or driving skills, achieved full marks.

(e) There was a broad range of answers to this question and about 75% of students achieved no marks at all. Many students contradicted their previous answer to part (d) and referred incorrectly to individual differences in reaction times and a similar proportion referred to order effects which had been controlled by counterbalancing or driving experience. Some students picked up on the possibility of differences in the nature of the ‘chat’ on the phone which was encouraging. However, few students showed any awareness of the need to match the two hazard perception tests (stimulus materials / tasks) in this repeated measures design.

(f) Many answers to this question displayed a marked lack of common sense. Despite referring to a simple hazard perception test, which is a key component of the driving test, many students claimed that watching a 3-minute film of a road would be traumatic, leading police drivers to suffer psychological harm. Others referred to possible deception and failed to appreciate that the purpose of the experiment is rather obvious in a repeated measures design. Better answers took issues such as informed consent / right to withdraw and explained how these related to this research.

(g) The question on writing instructions was answered well, with around half of students achieving four or five marks. Some failed to gain full credit as their instructions referenced both conditions or failed to include a check of understanding. Very weak answers failed to refer to the conversation or made no reference to reacting as quickly as possible.

(h) This question required students to identify an appropriate statistical test and justify their choice. About one third of students gained the full marks for identifying the Wilcoxon test with appropriate justification but just under half gained one mark only for identification of the test. Common problems included justification as a test of difference which gained no credit as it was included in the question. Other students were confused about the type of

data required for the Wilcoxon test and many answers referred to ‘not nominal’ data.

(i) There is still evidence that few students understand the concepts of statistical error and well over half failed to gain any marks here. Some became confused between type 1 and type 2 errors and others referred to the number of hazards detected rather than reaction times.

(j) This question was answered reasonably well, with many students referring to the greater potential for generalisation in a larger sample of inexperienced drivers. Some also referred to the general importance of replication to check findings in the context of the experiment, which was creditworthy.

28

Answers need to explain the strength rather than merely state it. So to state that questionnaires are “quick” gains no credit because there is no explanation as to why they are quick. Better answers compared questionnaires to other methods such as interviews. So to say that a questionnaire is quick in comparison to an interview becomes a creditworthy comment. Since the question asked about the strength of the methods, answers that referred only to the type of data did not receive credit, unless they were explicitly linked to the type of question that would generate such data. Some students gave generic answers that could apply to either questionnaires or interviews, and these received minimal credit.

30

This was a data response question, asking students to interpret the data in the table. The focus here must be on these results and not on prior knowledge. Answers that went beyond the data and discussed the findings of Bickman’s study, or Milgram’s findings did not gain credit. Better answers included comment on both what the confederate was wearing (smart versus casual) and the task involved (easy versus difficult). Some students appeared to think that the terms obedience and conformity are interchangeable.

32

(a) Some students clearly understood what behavioural categories are while others had no idea. A number of students were unable to suggest operationalised behaviour categories. Kicking and swearing would be suitable, physical and verbal aggression would be unsuitable. Given the material in the stem, just suggesting children would be observed in the playground, with no further elaboration did not attract credit. A few students did not take account of the stem material and suggested how children could be observed in the home at the age of two.

(b) Most students could identify an ethical issue but some chose issues that were difficult to apply to this scenario. Many ethical issues could be made relevant; the need for informed consent and the need for confidentiality were apt. The effectiveness of the elaboration distinguished between students' marks. Students who chose to write about protection from harm needed to make an appropriate case for this.

(c) Students who responded to this question in the context of the stem material were able to contextualise the comparison between interviews and questionnaires and that helped in producing an appropriate answer. A number of students produced good answers referring to the interviewer's opportunity to clarify questions and answers and the benefits that might accrue from being face to face with the parents. Students needed to explain why parents might be more likely to lie in a questionnaire or in an interview, since a case could be made for either. Some students who did not read the question carefully wrote about the advantages of questionnaires. A few students thought the children, rather than the parents were being interviewed.

34

(a) There were some excellent answers referring to checking the procedure of the words / words and picture study and making changes if necessary. There appeared to be three main areas of misunderstanding; that a pilot study should be used to see if the hypothesis was supported, to see whether a different experimental design should be used or to check participants. In relation to the final point there seemed to be an incorrect assumption by some students that the same participants would be used in the pilot study as the experiment.

(b) Students were required to operationalise the IV and the DV in order to score full marks. “There will be a difference in the number of words correctly recalled when words are presented with pictures and without pictures” is operationalised. “There will be a difference in words recalled in condition one and condition two” is not operationalised and would therefore attract only one of the two marks available. Many answers referred to a correlation, using the term ‘relationship’ or ‘link’ when they meant a difference. A few students produced a null hypothesis in response to this question.

(c) The comparison of experimental designs proved difficult for some students. Students who considered the designs in relation to the stem material were more likely to produce an appropriate answer. Those who just repeated rote learnt problems of repeated measures designs were less successful. For example, in the context of this study, (learning 20 words / 20 words with pictures) fatigue was unlikely, given the minimal requirements of the tasks.

(d) Students were required to show some understanding of the outcome of the experiment. This could be achieved by reference to the median and range in such a way that it was apparent that the students understood the terms, eg referring to the average or spread of the scores. Students could also show understanding by drawing an inference from the figures. Simply repeating the contents of the table showed no understanding.