ALevel Psychology Paper 2

Research Methods 2 Past Questions and Mark Scheme

Name:

Class:

Time:

Marks:

Comments:

1

A psychologist studying obedience conducted the following experiment.

A confederate (stooge) approached people in the street and instructed them to pick up a piece of litter and put it in a nearby bin. None of the people approached had dropped the litter.

There were two groups in the experiment.

Group A The confederate (stooge) was dressed in everyday clothing.

Group B The confederate (stooge) was dressed in a uniform.

The psychologist recorded how many people in each group obeyed the instruction of the confederate (stooge).

(a) Identify the experimental design that was used in this study. Briefly explain **one** advantage of using this experimental design in this study.

(3)

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

(2)

(c) Use your knowledge of research into obedience to explain the likely outcome of this experiment.

(3)

(d) Briefly outline **one** ethical issue that might have arisen in this experiment.

(2)
(Total 10 marks)

2

A case study was carried out on Peter whose brain was damaged in a motorcycle accident. Psychologists tested how many numbers he could hold in his short-term memory. They did this by reading him lists of numbers and asking him to recall the numbers immediately in the right order. He could recall a maximum of two items. The psychologists found that his long-term memory was normal.

(a) How was Peter's short-term memory after the accident different from most adults' short-term memory?

_____

_____

_____

_____

(2)

(b)   Does this case study support the multi-store model of memory? Explain your answer.

_____

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

_____

(4)

(c)   Identify **one** ethical issue associated with this case study of Peter. Suggest how
         psychologists could deal with this ethical issue.

Ethical issue _____

_____

How psychologists could deal with this ethical issue _____

_____

_____

_____

_____

_____

(4)
(Total 10 marks)

3 Psychologists have carried out research into the use of cognitive interviews. One possible ethical issue which might arise during this research is protection of participants from harm. Explain how psychologists could deal with this ethical issue.

_____

_____

_____

_____

_____

_____

(Total 3 marks)

4 One situation in which disruption of attachment can occur is when a mother of a young child is admitted into hospital. A researcher decided to study the behaviour of a two year old boy who experienced this disruption of attachment.

She decided to use naturalistic observation of the boy both before his mother was admitted into hospital and after she returned home. Each period of observation lasted for one hour.

(a) Suggest **two** suitable behavioural categories the researcher could use to record the boy's behaviour.

_____

_____

_____

_____

(2)

(b) How might the researcher record the boy's behaviour during the one-hour observation?

_____

_____

_____

_____

(2)

(c)　Explain why the psychologist might want to carry out a pilot study before the main observation.

_____

_____

_____

_____

<div align="right">

(2)
(Total 6 marks)

</div>

5

A psychologist analysed the results of 'Strange Situation' studies from different countries. Some of the results are shown below.

Country (and number of studies)	% of each type of attachment		
	Secure	**Insecure-Avoidant**	**Insecure-Resistant**
Country One (2)	64	7	29
Country Two (18)	65	21	14
Country Three (4)	67	26	7

(a)　Outline what the table above shows about cultural variations in attachment.

_____

_____

_____

_____

_____

_____

<div align="right">

(3)

</div>

(b)   Explain **one** criticism of investigating cultural variations in attachment using the 'Strange Situation.'

_____

_____

_____

_____

_____

_____

_____

(3)
(Total 6 marks)

6   A psychologist was interested in testing a new treatment for people with eating disorders. She put up adverts in several London clinics to recruit participants. Thirty people came forward and they were all given a structured interview by a trained therapist. The therapist then calculated a numerical score for each participant as a measure of their current functioning, where 50 indicates excellent, healthy functioning and zero indicates failure to function adequately. The psychologist then randomly allocated half the participants to a treatment group and half to a no-treatment group. After eight weeks, each participant was re-assessed using a structured interview conducted by the same trained therapist, and given a new numerical score. The trained therapist did not know which participants had been in either group.

For each participant, the psychologist calculated an improvement score by subtracting the score at the start of the study from the score after eight weeks. The greater the number, the better the improvement.

**Median and range of improvement scores for the
treatment group and for the no-treatment group**

	Treatment group	No–treatment group
Median	10.9	2.7
Range	2.1	0.8

(a)   With reference to the data in the table above, outline what the findings of this investigation seem to show about the effectiveness of the treatment.

(2)

(b)   The psychologist used a statistical test to find out whether there was a significant difference in improvement between the 'treatment' and 'no-treatment' groups. She found a significant difference at the 5% level for a one-tailed test ($p \leq 0.05$).

Identify an appropriate statistical test for analysing the participants' scores.
Explain why it would be a suitable test to use in this study.

(4)

(c)   What is the likelihood of the psychologist having made a Type 1 error in this study? Explain your answer.

(2)

(d)   The psychologist assumed that improvements in the treatment group were a direct result of the new type of treatment. Suggest **two** other reasons why people in the treatment group might have improved.

(4)

(e)   The psychologist could have used self-report questionnaires to assess the participants instead of using interviews with the therapist. Explain **one** advantage and **one** disadvantage of using self-report questionnaires in this study rather than interviews.

(4)

(f)   The psychologist needed to obtain informed consent from her participants. Write a brief consent form which would be suitable for this study. You should include some details of what participants could expect to happen in the study and how they would be protected.

(5)

(g)   What is meant by reliability? Explain how the reliability of the scores in this study could be checked.

(4)

(h)   The psychologist noticed that female and male participants seemed to have responded rather differently to the treatment.

She decided to test the following hypothesis:

Female patients with an eating disorder will show greater improvement in their symptoms after treatment with the new therapy than male patients.

She used a new set of participants and, this time, used self-report questionnaires instead of interviews with a therapist.

Imagine that you are the psychologist and are writing up the report of the study. Write an appropriate methods section which includes reasonable detail of design, participants, materials and procedure. Make sure that there is enough detail to allow another researcher to carry out this study in the future.

(10)
(Total 35 marks)

7   A psychologist showed participants 100 different cards, one at a time.

Each card had two unrelated words printed on it, eg DOG, HAT.

Participants in one group were instructed to form a mental image to link the words.
Participants in the other group were instructed simply to memorise the words.
After all the word pairs had been presented, each participant was shown a card with the first word of each pair printed on it. Participants were asked to recall the second word.

The following results were found.

Number of words correctly recalled

	Mental Image Group	**Memorising Group**
Mean Score	80	45

(a)   What is the independent variable (IV) in this study?

_____

_____

_____

_____

(2)

(b)   What is the dependent variable (DV) in this study?

_____

_____

_____

_____

(2)

(c)   What experimental design was used in this study?

_____

_____

(1)

(d)   Explain **one** strength of this experimental design in the context of this study.

_____

_____

_____

_____

(2)

(e) Explain how a psychologist could find out whether these results are reliable.

_____

_____

_____

_____

(2)
(Total 9 marks)

8

By observing interactions between the infants and their mothers in a Strange Situation, Mary Ainsworth was able to identify different types of attachment.

Describe possible demand characteristics in this research.

_____

_____

_____

_____

_____

_____

(Total 3 marks)

9

Research has suggested that institutionalisation can have negative effects on children. In the 1990s, many children were found living in poor quality orphanages in Romania. Luca had lived in one of these orphanages from birth. When he was four years old, he was adopted and he left the orphanage to live in Canada. His development was then studied for a number of years.

(a) Outline possible negative effects of institutionalisation on Luca.

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(4)

(b)   The scenario above is an example of a case study. Outline **one** strength and **one** limitation of this research method.

Strength _____

_____

_____

Limitation _____

_____

_____

_____

(4)
(Total 8 marks)

10

A recent study recorded the amount of time that children spent in day care from birth to four years, and asked each child's mother to rate her child for aggression and disobedience. The study found that, as the time spent in day care went up, the mothers' rating of aggression and disobedience also went up.

(a)   What kind of correlation is this research showing?

_____

_____

(1)

(b)   Outline **one** strength and **one** weakness of using correlational research to investigate the effects of day care.

Strength _____

_____

_____

_____

Weakness _____

_____

_____

_____

(4)
(Total 5 marks)

11

Some psychology students read about an experiment which suggested that organisation is a useful strategy for improving memory. The students carried out an experiment to investigate the effects of organisation on word recall. They made up a list of 50 items that could be bought in a supermarket. The participants were teachers at their school. One group of participants saw the words organised into categories such as fruit, vegetables, dairy products and cleaning materials. The other group saw the same words presented randomly.

The results are given in the **Table** below.

The number of words correctly recalled by participants who saw the organised list and participants who saw the random list

	Organised List	Random List
	20	15
	15	13
	18	19
	45	14
	24	20
	23	10
	28	21
	21	6
	25	22
	30	25
Measure of central tendency		
Measure of dispersion		

(a)   Identify a suitable measure of central tendency that could be used with these data. Justify your answer.

_____

_____

_____

_____

(2)

(b)   The psychology students decided to use a volunteer sample. Suggest **one** way in which this sample could be obtained.

_____

_____

_____

_____

(2)

(c)   Suggest **one** possible extraneous variable in this study.

_____

_____

(1)

(d)   Suggest **one** way in which the students could control for this extraneous variable.

_____

_____

_____

_____

(2)
(Total 7 marks)

12

A psychologist wanted to investigate whether or not the presence of an audience had an effect on the performance of a task.

The task was to shoot netballs through a hoop. Each participant took 20 shots.

The experiment took place in a school sports hall.

The psychologist obtained the sample of participants from a local secondary school for girls. She drew the names of 40 participants at random from a list of girls who all played netball regularly. The first 20 participants drawn took part in the experimental condition and the next 20 participants took part in the control condition.

In the experimental condition, each participant took 20 shots. In this condition, 50 pupils from the school acted as an audience watching the performance.

In the control condition, the other 20 participants performed the same task, but this time without the audience watching.

The psychologist observed each girl's performance and recorded the number of netballs successfully shot through the hoop.

(a)   State an appropriate hypothesis for this study.

_____

_____

_____

_____

(2)

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

The mean number of netballs successfully shot through the hoop in the presence and absence of an audience.

	Mean number of netballs successfully shot through the hoop
Presence of an audience	15
Absence of an audience	9

(b)   What might the psychologist conclude from the data in the table above? Justify your answer.

_____

_____

_____

_____

_____

(2)

(c)   Sketch an appropriate graphical display of the data shown in **Table**.

_____

_____
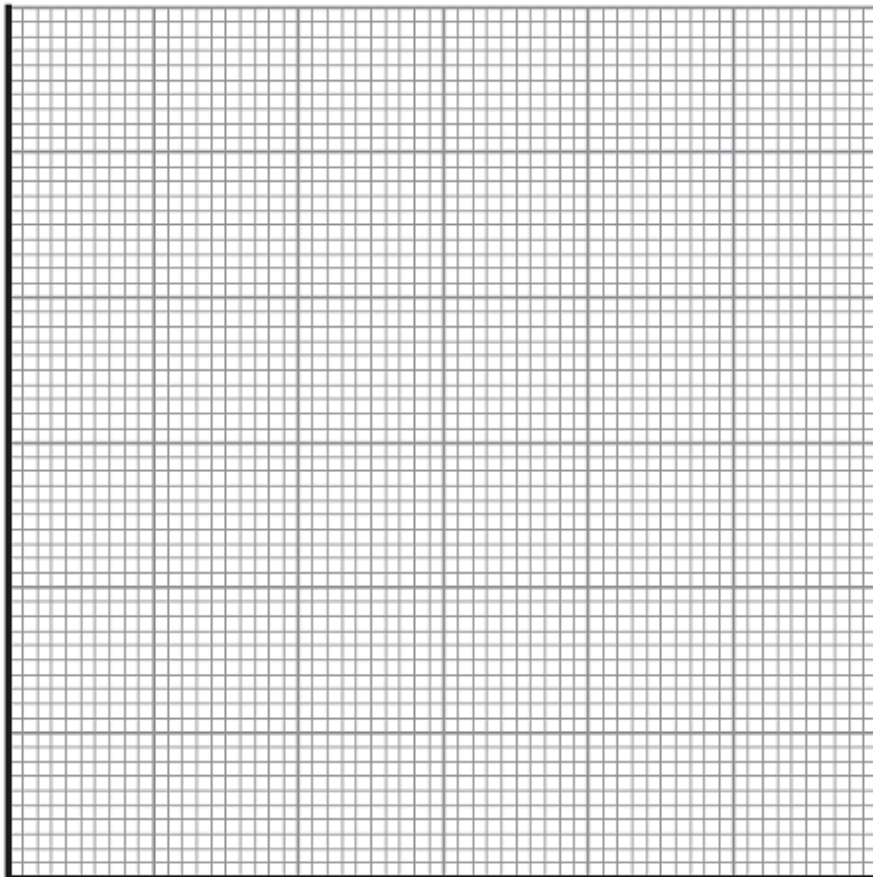
(3)

(d)   The psychologist used random sampling to select the participants in this study.

Explain **one** strength of using a random sample.

_____

_____

_____

_____

(2)

(e)   (i)   The psychologist used an independent groups design in this study.

Briefly explain **one** limitation of using an independent groups design in this study.

_____

_____

_____

_____

(2)

(ii)   Explain how the limitation that you have identified in your answer to **(e) (i)** might have been overcome.

_____

_____

_____

_____

(2)

(f)   (i)   What is meant by an *extraneous variable*?

_____

_____

(1)

(ii)   Explain why it is important to control extraneous variables in experimental research.

_____

_____

_____

_____

(2)

(g)   Write a short set of instructions that the psychologist could have read to the participants in the experimental condition.

_____

_____

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(4)
(Total 20 marks)

13

One measure of the functioning of the immune system is the level of activity of white blood cells.

What does the graph below tell you about the relationship between stress and the level of activity of white blood cells?

Relationship between stress and the level of activity of white blood cells

_____

_____

_____

_____

(Total 2 marks)

14
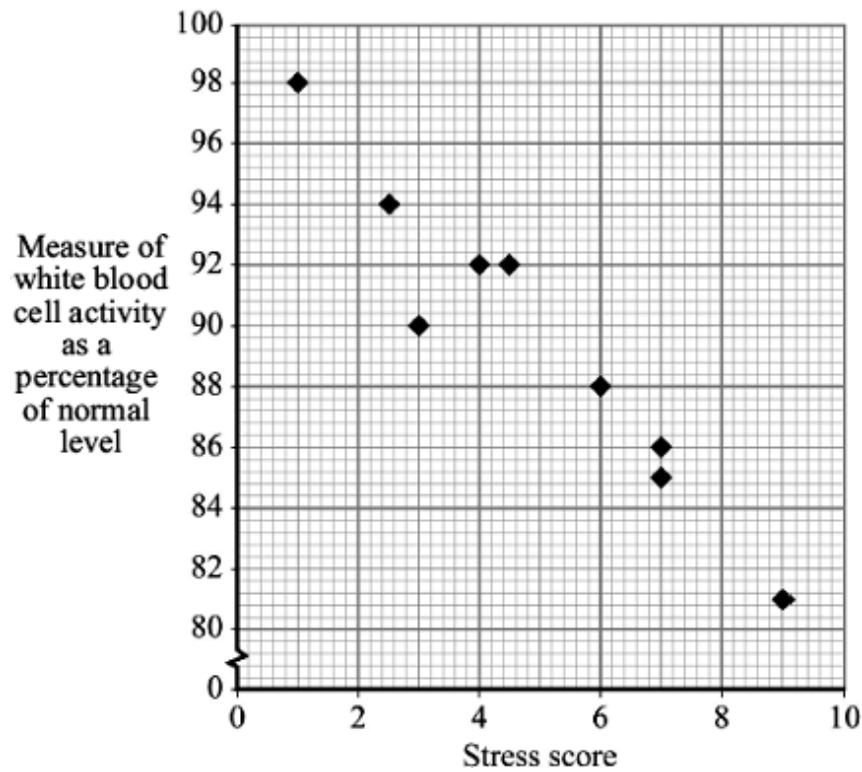
In an experiment into conformity, an experimenter varied both the number of confederates (stooges) and the ambiguity of the task. The bar chart below shows the findings.

Percentages of conformity for different tasks and size of majority
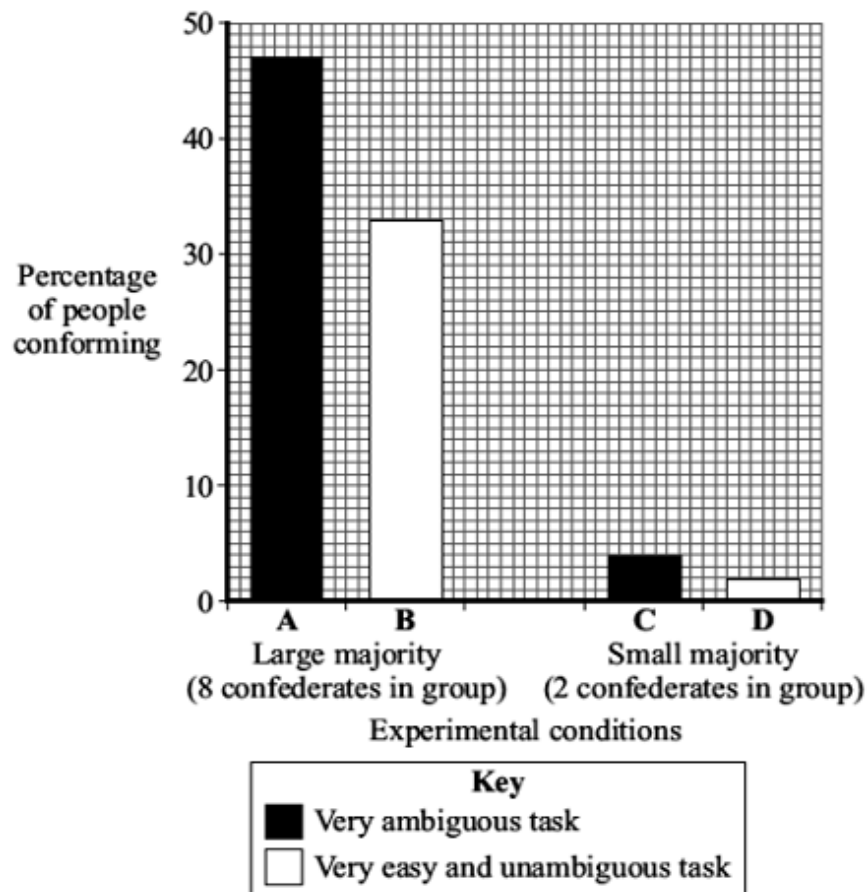
What does the bar chart show about conformity?

_____

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(Total 4 marks)

15 (a) The psychologist was also interested in the effects of a restricted diet on memory functioning and he expected memory to become impaired. The psychologist's hypothesis was that participants' scores on a memory test are lower after a restricted diet than before a restricted diet. He gave the volunteers a memory test when they first arrived in the research unit and a similar test at the end of the four-week period. He recorded the memory scores on both tests and analysed them using the Wilcoxon signed ranks test. He set his significance level at 5%.

His calculated value was $T = 53$.

State whether the hypothesis for this study is directional or non–directional.

(1)

(b) **Table: Extract from table of critical values from the Wilcoxon signed ranks test**

Level of significance for a one-tailed test	0.05	0.025
Level of significance for a two-tailed test	0.1	0.05
N	$T \leq$	
19	53	46
20	60	52
21	67	58
22	75	65

Calculated T must be equal to or less than the critical value (table value) for significance at the level shown

Using the table above, state whether or not the psychologist's result was significant. Explain your answer.

(3)
(Total 4 marks)

16 Read the text below and answer the questions that follow.

A psychologist is using the observational method to look at verbal aggression in a group of children with behavioural difficulties. Pairs of observers watch a single child in the class for a period of one hour and note the number of verbally aggressive acts within ten-minute time intervals. After seeing the first set of ratings, the psychologist becomes concerned about the quality of inter-rater reliability. The tally chart for the two observers is shown in the table below.

Table: Observation of one child – number of verbally aggressive acts in ten-minute time intervals

Time slots	0–10	11–20	21–30	31–40	41–50	51–60
Observer A	2	5	0	6	4	3
Observer B	4	3	2	1	6	5

(a) Use the data in the Table above to sketch a scattergram. Label the axes and give the scattergram a title.

(4)

(b) Using the data in the Table above, explain why the psychologist is concerned about inter-rater reliability.

(4)

(c) Identify an appropriate statistical test to check the inter-rater reliability of these two observers. Explain why this is an appropriate test.

(3)

(d) If the psychologist does find low reliability, what could she do to improve inter-rater reliability before proceeding with the observational research?

(4)
(Total 15 marks)

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

Strength _____

_____

_____

_____

Weakness _____

_____

_____

_____

(Total 4 marks)

18 Outline what is meant by the term *peer review* in psychological research.

(Total 2 marks)

19 Observation in a Strange Situation has been used to investigate cultural variations in attachment.

Give **one** advantage of using observation in psychological research.

_____

_____

_____

_____

(Total 2 marks)

20 Read the text below and answer following questions.

A psychologist was interested in looking at the effects of a restricted diet on psychological functioning. A group of 20 healthy, young adult volunteers agreed to spend four weeks in a research unit. They were kept warm and comfortable but given only water and small amounts of plain food. They were able to socialise with one another and watch television, but they had to keep to strict, set mealtimes and were not allowed to eat anything between meals. The psychologist carried out various tests of emotional and cognitive functioning during this four-week period. One area of interest for the psychologist was the effect of the dietary restriction on the perception of food. He tested this by asking the volunteers to draw pictures of food at the end of each week. When all the drawings had been completed, the psychologist used content analysis to analyse them.

(a) What is meant by the term *content analysis*?

(1)

(b) Explain how the psychologist might have carried out content analysis to analyse these drawings.

(3)

(c) The psychologist needed to be sure that his participants understood the nature of the study so that they were able to give informed consent.

Write a consent form which would be suitable for this study. Make sure there is sufficient information about the study for the participants to make an informed decision.

(5)
(Total 9 marks)

21

A researcher wanted to investigate the effectiveness of therapy as a treatment for obsessive-compulsive disorder in children. Before the therapy started, the mothers of 10 children with obsessive-compulsive disorder each rated the anxiety of their child. They used a rating scale of 1–10, where 1 meant not at all anxious and 10 meant extremely anxious. Each child then attended a programme of therapy. At the end of the programme, each mother rated her child again, using the same anxiety scale. The scores for each child before and after therapy were used to calculate a median anxiety rating.

The data are shown in the table below.

Median ratings of children's anxiety before and after therapy

	Before therapy	**After therapy**
Median rating of anxiety	8.5	4.0

(a)   Identify **two** symptoms of obsessive-compulsive disorder.

(2)

(b)   Name and outline the experimental design used in this study.

(2)

(c)   Explain **one** advantage of this experimental design.

(2)
(Total 6 marks)

22

Questionnaires and interviews are both self-report techniques. Explain **one** advantage **and one** disadvantage of using a questionnaire rather than an interview.

Advantage _____

_____

_____

_____

Disadvantage _____

_____

_____

_____

(Total 4 marks)

23

Explain why peer review is important in psychological research.

(Total 5 marks)

24 A researcher wanted to investigate the effectiveness of a language therapy for children with autism. Before the therapy started, the mothers of 10 children with autism each rated the verbal interaction of their child. They used a rating scale of 1–10, where 1 meant very poor verbal interaction and 10 meant very good verbal interaction. Each child then attended a programme of language therapy. At the end of the programme, each mother rated her child again, using the same verbal interaction scale. The scores for each child before and after therapy were used to calculate a median verbal interaction rating.

The data are shown in the table below.

Median ratings of children's verbal interaction before and after therapy

	Before therapy	**After therapy**
Median rating of verbal interaction	4.0	8.5

(a)   Name and outline the experimental design used in this study.

(2)

(b)   Explain **one** advantage of this experimental design.

(2)

(c)   Explain what the median ratings in the table above indicate about the effectiveness of the language therapy.

(2)
(Total 6 marks)

25 A psychologist investigated the relationship between type of attachment in childhood and success in later adult relationships. He published a questionnaire in a local newspaper. The participants were people who read the newspaper, filled in the questionnaire and sent it to the psychologist. Participants' answers to the questions were used to decide whether they had been securely or insecurely attached as children. The participants who were identified as securely attached children were more likely to have successful adult relationships than those identified as insecurely attached children.

(a)   Identify the sampling technique used in this study. Outline **one** weakness of using this sampling method.

_____

_____

_____

_____

_____

_____

(3)

(b)   Identify **one** ethical issue the researcher would need to consider in this research. Suggest how the researcher could deal with this ethical issue.

_____

_____

_____

_____

_____

_____

(3)
(Total 6 marks)

26 Explain why it might be better to carry out research into eyewitness testimony in the real world, rather than in a laboratory.

_____

_____

_____

_____

_____

_____

(Total 3 marks)

27 Questionnaires have been used by psychologists to investigate stress.

Explain **one** possible ethical issue that might arise when using questionnaires in this area of psychology.

_____

_____

_____

_____

(Total 2 marks)

28 A teacher has worked in the same primary school for two years. While chatting to the children, she is concerned to find that the majority of them come to school without having eaten a healthy breakfast. In her opinion, children who eat 'a decent breakfast' learn to read more quickly and are better behaved than children who do not. She now wants to set up a pre-school breakfast club for the children so that they can all have this beneficial start to the day. The local authority is not willing to spend money on this project purely on the basis of the teacher's opinion and insists on having scientific evidence for the claimed benefits of eating a healthy breakfast.

(a) Explain why the teacher's personal opinion cannot be accepted as scientific evidence. Refer to some of the major features of science in your answer.

A psychologist at the local university agrees to carry out a study to investigate the claim that eating a healthy breakfast improves reading skills. He has access to 400 five-year-old children from 10 local schools, and decides to use 100 children (50 in the experimental group and 50 in the control group). Since the children are so young, he needs to obtain parental consent for them to take part in his study.

(6)

(b) The psychologist used a random sampling method. Explain how he could have obtained his sample using this method.

(3)

(c) Explain limitations of using random sampling in this study.

(3)

(d) Explain why it is important to operationalise the independent variable and the dependent variable in this study and suggest how the psychologist might do this.

(5)

(e) The psychologist used a Mann-Whitney test to analyse the data. Give **two** reasons why he chose this test.

(2)

(f) He could have used a matched pairs design. Explain why this design would have been more difficult to use in this study.

(2)

(g) Other than parental consent, identify **one** ethical issue raised in this study and explain how the psychologist might address it.

(2)

(h)   The psychologist asks some of his students to conduct a separate observational study at the same time on the same group of children. The aim of this observational study is to test the idea that eating a healthy breakfast affects playground behaviour.

Design an observational study to investigate the effects of a healthy breakfast on playground behaviour. Include in your answer sufficient detail to allow for reasonable replication of the study. You should state the hypothesis you are setting out to test.

In your answer, refer to:

•   an appropriate method of investigation

•   materials/apparatus and procedure.

Justify your design decisions.

(12)
(Total 35 marks)

29   Psychologists sometimes use questionnaires to find out about stress. Explain **two** strengths of using questionnaires in research.

Strength one _____

_____

_____

_____

Strength two _____

_____

_____

_____

(Total 4 marks)

30

A psychologist collected data on the effectiveness of three types of therapies for different types of disorders. The results are displayed in the bar chart below.

Percentage of patients who improved according to type of disorder and type of therapy
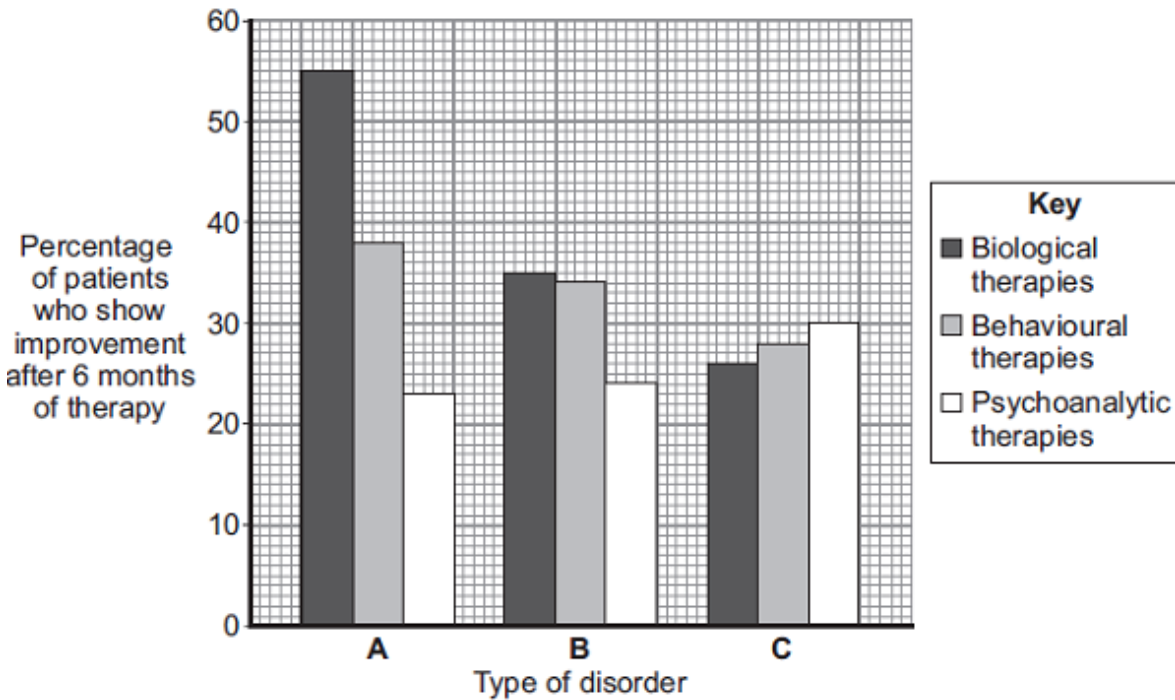
What conclusions about the effectiveness of different therapies can you draw from the bar chart?

_____

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

(Total 4 marks)

31

A psychologist assessed the aggressive behaviour of 100 five-year-old children who were starting school. The children had attended day care for at least 20 hours a week. Fifty of the children had attended day nurseries. The other fifty children had been looked after by childminders. The children who attended the day nurseries were more aggressive than the children who had been looked after by childminders.

(a)   Explain why this is an example of a natural experiment.

_____

_____

_____

_____

(2)

(b)   Suggest **one** way in which the psychologist could have measured the children's aggressive behaviour.

_____

_____

_____

_____

(2)

(c)   Explain **two** ethical issues which the psychologist should have considered when
        carrying out this research.

        Issue 1 _____

        _____

        _____

        _____

        Extra space _____

        _____

        Issue 2 _____

        _____

        _____

        _____

        Extra space _____

        _____

                                                                        (4)
                                                              (Total 8 marks)

32

Psychologists carried out a laboratory experiment to investigate the effectiveness of cognitive interviews. All participants watched the same film of a robbery. They were randomly allocated to **Group One** or **Group Two**. Participants were then asked to recall the robbery. The investigators used a cognitive interview to access recall of participants in **Group One** and a standard interview to access recall of participants in **Group Two**.

The results of this experiment are summarised in the Figure below.

Average number of correct and incorrect statements made by participants under two different interview conditions
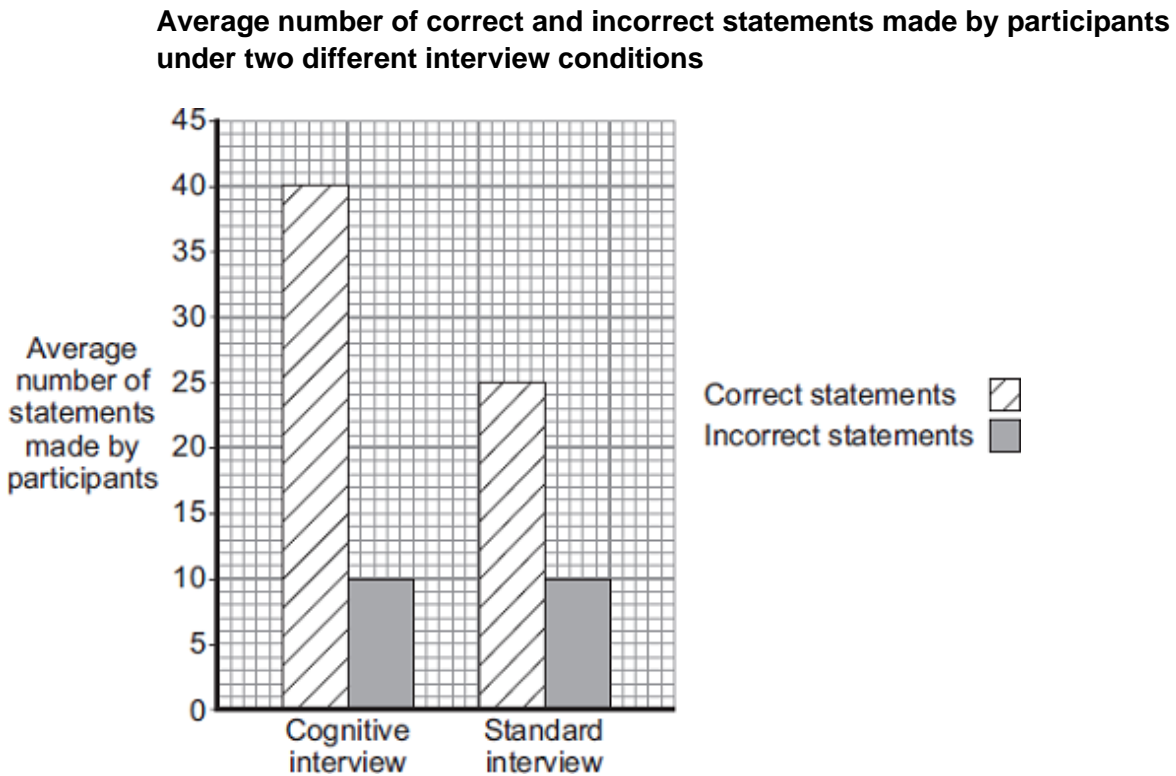
(a)   What experimental design was used in this experiment?

_____

_____

(1)

(b)   Explain **one** limitation of the design that was used in this experiment.

_____

_____

_____

_____

(2)

(c)   Explain what the results suggest about the effectiveness of the cognitive interview.

_____

_____

_____

_____

(2)

(d)   Participants in the standard interview were simply asked to describe what happened in the film.

Suggest **one** way in which participants in the cognitive interview condition could have been asked to recall what happened.

_____

_____

_____

_____

(2)

(e)   What is meant by the term investigator effects? Explain possible investigator effects in this study.

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(4)
(Total 11 marks)

Psychologists often need to select participants to take part in research. The descriptions below are all types of sampling method.

A The psychologist puts an advert in a newspaper, asking for participants.

B The psychologist uses lists of students in a university and selects every tenth student to take part.

C The psychologist asks some of his psychology students to take part in the research.

D The psychologist gives a number to all students in a university, then selects participants in an unbiased way.

In the table below, write which description, **A**, **B**, **C**, or **D**, matches each sampling method.

Sampling Method	Description
Random Sample	
Opportunity Sample	
Volunteer Sample	

(Total 3 marks)

34

Two different drug therapies were tested on a group of patients. All the patients suffered with the same anxiety disorder. Half the patients were given **Therapy A** and the other half were given **Therapy B**. Improvement was assessed on a scale from 0-25, where 0 = no improvement.

The table below shows the improvement made between the start and the end of the treatment.

Average and range of improvement scores

	Average	Range
Therapy A	6.5	2 – 19
Therapy B	6	4 – 9

Explain what these findings suggest about the different therapies?

_____

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(Total 4 marks)

35 Psychological research suggests an association between birth order and certain abilities. For example, first-born children are often logical in their thinking whereas later-born children tend to be more creative. A psychologist wonders whether this might mean that birth order is associated with different career choices. She decides to investigate and asks 50 artists and 65 lawyers whether they were the first-born child in the family or not.

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

(2)

(b) Identify an appropriate sampling method for this study and explain how the psychologist might have obtained such a sample

The psychologist found the following results:

- 20 of the 50 artists were first-born children

- 35 of the 65 lawyers were first-born children.

She analysed her data using a statistical test and calculated a value of x^2 = 2.27. She then looked at the relevant table to see whether this value was statistically significant. An extract from the table is provided below.

Table: Critical values of x^2

	Level of significance for a one-tailed test			
	0.10	0.05	0.025	0.01
	Level of significance for a two-tailed test			
	0.20	0.10	0.05	0.02
df				
1	1.64	2.71	3.84	5.41

Calculated value of x^2 must be equal to or exceed the table (critical) values for significance at the level shown

(3)

(c)   Imagine that you are writing the results section of the report on this investigation. Using information from the description of the study above and the relevant information from the statistical table, provide contents suitable for the results section.

You must provide all of the following:

- an appropriately labelled 2×2 contingency table a sketch of an appropriately labelled bar chart
- identification of the appropriate statistical test with justification for its use
- identification of an appropriate significance level
- a statement of the results of the statistical test in relation to the hypothesis.

(12)
(Total 17 marks)

36

It is thought that colours might affect our performance when carrying out certain tasks. Research in this area has been inconclusive. Some studies have shown that red improves performance but others have found the opposite. It could be that these contradictory results have arisen because red is beneficial only for certain kinds of mental processing. Some psychologists tested this hypothesis in a series of independent-groups design experiments using students at a Canadian university.

The experiments involved computer tasks, with either a red, blue or neutral background appearing on the monitor. The researchers found that participants were better at a word-recall task and a spell-checking task when the screen background was red rather than blue or neutral. However, participants thought of more creative ideas when the screen was blue rather than red or neutral.

The researchers concluded that red is beneficial for tasks that require attention to detail whereas blue aids creativity.

(a)   What were the researchers' aims in this study?

Imagine that you are writing up the report for this series of experiments.

(2)

(b)   What is the purpose of the introduction section of a report?

A psychological report also contains a discussion section. Researchers are expected to consider their findings critically and discuss issues such as validity.

(2)

(c)   What is meant by validity?

(1)

(d)   In the discussion section, researchers are also expected to consider any possible applications of their research. Suggest **one** practical application that might arise from these findings.

In a further experiment, participants were given 20 blue shapes or 20 red shapes. They were then asked to pick 5 shapes and use them to make a toy suitable for a child aged between five and eleven years. They were given a limited time to carry out this task. Participants given red shapes made toys that independent judges rated to be more practical but less original, whereas participants given blue shapes made more creative toys.

(2)

(e)   Explain why the researchers asked independent judges to rate the toys.

(2)

(f)   Write a set of standardised instructions that would be suitable to read out to participants in this experiment.

(5)
(Total 14 marks)

37

Psychologists sometimes use case studies to study children. One example was of a boy who was discovered at the age of six. He had been kept in a darkened room and had had almost no social contact with people.

(a)   How could a psychologist maintain confidentiality when reporting a case study?

_____

_____

_____

_____

(2)

(b)   Psychologists use a range of techniques to gather information in case studies.

        Outline **one** technique which the psychologist could use in this case study.

        _____

        _____

        _____

        _____

                                                                                        (2)

(c)   Apart from ethical issues, explain **one or more** limitations of using case studies.

        _____

        _____

        _____

        _____

        _____

        _____

        _____

        _____

        Extra space _____

        _____

        _____

        _____

                                                                                        (4)
                                                                            (Total 8 marks)

38

Some research has suggested that there is a relationship between the time children spend in day care and their aggressive behaviour. Researchers selected a group of school children who had been in day care. They asked the children's mothers to estimate how many hours a week their children had spent in day care. They measured the same children's aggression.

Relationship between average hours spent in a day care and aggression
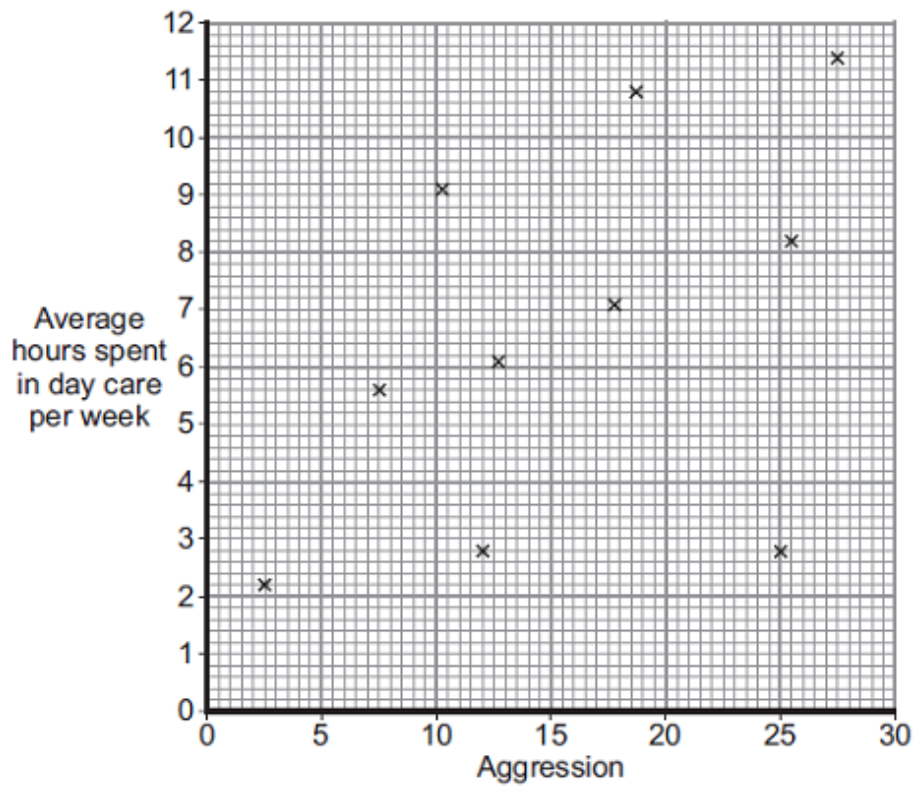
(a)   How many children are represented in this scattergram?

_____

(1)

(b)   Suggest **one** way in which the children's aggression could be measured.

_____

_____

_____

_____

(2)

(c)   In the media, it was claimed that the results showed spending time in day care caused the children to become aggressive.

Explain why this claim may be untrue.

_____

_____

_____

_____

_____

_____

(3)
(Total 6 marks)

39 A psychologist studying the primacy effect in impression formation conducted the following experiment.

Each participant was taken to the same room where they listened to a description of a person called 'Alex'. The participants were randomly allocated to one of two groups in the experiment. The psychologist gave each participant the same information about 'Alex', but the order of the information varied depending on the group.

Group A   Five positive points about Alex's personality were followed by five negative points.

Group B   Five negative points about Alex's personality were followed by five positive points.

After listening to the passage, each participant was asked to state whether they thought 'Alex' was a friendly person or not. The psychologist recorded how many participants in each group stated that Alex was 'friendly'.

(a)   Identify the type of experiment that was conducted. Select **one** option from the list below.

• Laboratory experiment

• Field experiment

• Quasi-experiment

(1)

(b)   Briefly explain **one** advantage of the type of experiment that you have identified in your answer to **part (a)**.

(2)

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

(2)

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

(Total 6 marks)

40   The following scattergram shows the relationship between the number of weeks of treatment with ECT and the score on the Self-Rating Depression Scale (on this scale, a high score indicates depression).

Relationship between weeks of treatment and scores on depression scale

Score on self-rating depression scale

Weeks of treatment
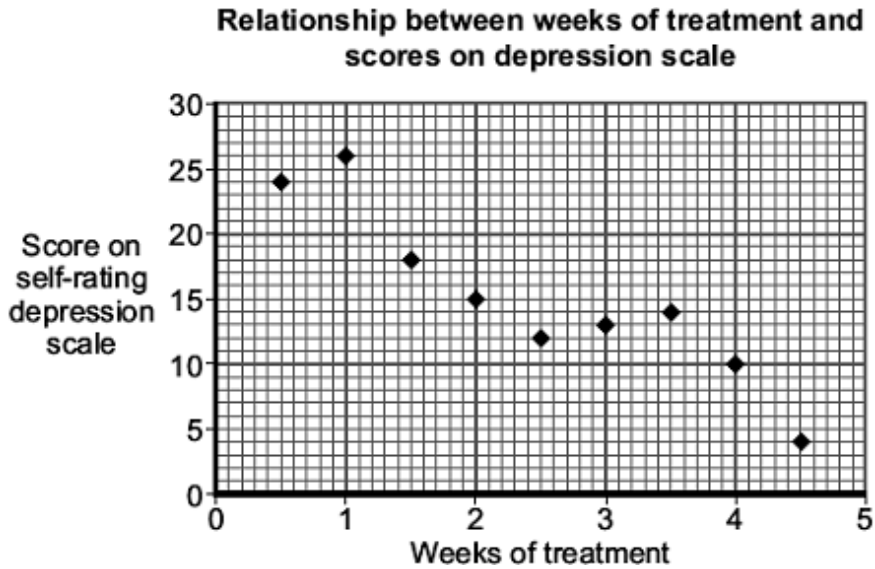
Outline what the scattergram seems to show.

_____

_____

_____

_____

_____

_____

_____

(Total 4 marks)

41 Research has shown that music can affect the ability to concentrate. Design an experiment that could be carried out in a classroom to test the effects of two different kinds of music on a task requiring concentration (eg word search).

You **must** use a repeated measures design.

In your answer you should:

• fully operationalise the independent and dependent variables

• provide details of how you would control extraneous variables

• describe the procedure that you would use. You should provide sufficient detail for the study to be carried out.

(Total 10 marks)

42 A psychologist carried out an experiment using an independent groups design. The psychologist wished to investigate the effectiveness of a strategy for memory improvement. In one condition, participants were taught a memory improvement strategy. In the other condition, participants were not taught this memory improvement strategy. All participants were asked to memorise 10 pictures of familiar objects. For example, the first was a doll, the second was an apple. All participants were then given 50 pictures each, and asked to select the original 10.

The psychologist did a pilot study before carrying out the experiment. The results of the experiment are shown in the table below.

(a) Write a directional hypothesis for this experiment.

_____

_____

_____

_____

(2)

(b) Explain what is meant by an independent groups design.

_____

_____

(1)

(c) Explain **one** strength and **one** limitation of using an independent groups design.

Strength _____

_____

_____

_____

Limitation _____

_____

_____

_____

(4)

(d) Explain why the psychologist did a pilot study.

_____

_____

_____

_____

_____

_____

(3)

**The mean number of pictures correctly identified and standard
deviations for participants with the memory improvement strategy
and without the memory improvement strategy**

	With memory improvement strategy	Without memory improvement strategy
Mean	8	7
Standard Deviation	2.80	0.29

(e)   What do the standard deviations in the table above tell us about the performance of the two groups?

_____

_____

_____

_____

(2)
(Total 12 marks)

43   The psychologists then wanted to see whether the use of diagrams in medical consultations would affect recall of medical information.

In a laboratory experiment involving a medical consultation role-play, participants were randomly allocated to one of two conditions. In Condition A, a doctor used diagrams to present to each participant a series of facts about high blood pressure. In Condition B, the same doctor presented the same series of facts about high blood pressure to each participant but without the use of diagrams.

At the end of the consultation, participants were tested on their recall of facts about high blood pressure. Each participant was given a score out of ten for the number of facts recalled.

(a)   In this case, the psychologists decided to use a laboratory experiment rather than a field experiment. Discuss advantages of carrying out this experiment in a laboratory.

(4)

(b)   Identify an appropriate statistical test that the psychologists could use to analyse the data from the follow-up study. Give **one** reason why this test is appropriate.

(2)
(Total 6 marks)

44   Two psychologists investigated the relationship between age and recall of medical advice. Previous research had shown that recall of medical advice tended to be poorer in older patients. The study was conducted at a doctor's surgery and involved a sample of 30 patients aged between 18 and 78 years. They all saw the same doctor, who made notes of the advice that she gave during the consultation.

One of the psychologists interviewed each of the patients individually, immediately after they had seen the doctor. The psychologist asked each patient a set of questions about what the doctor had said about their diagnosis and treatment. The patients' responses were recorded and then typed out. Working independently the psychologists compared each typed account with the doctor's written notes in order to rate the accuracy of the accounts on a scale of 1 – 10. A high rating indicated that the patient's recall was very accurate and a low rating indicated that the patient's recall was very inaccurate.

(a)   The psychologists decided to propose a directional hypothesis. Why was a directional hypothesis appropriate in this case?

(1)

(b)   Write a suitable directional hypothesis for this investigation.

(3)

(c)   The psychologists were careful to consider the issue of reliability during the study. What is meant by reliability?

(1)

(d)   Explain how the psychologists might have assessed the reliability of their ratings.

(3)

(e)   This study collected both qualitative and quantitative data. From the description of the study above, identify the qualitative data and the quantitative data.

The psychologists used Spearman's rho to analyse the data from their investigation. They chose to use the 0.05 level of significance. The result gave a correlation coefficient of −0.52.

(2)

(f)   Give **two** reasons why the psychologists used Spearman's rho to analyse the data.

(2)

(g)   Using the table below, state whether the result is significant or not significant and explain why.

(2)

Extract from a table of critical values of Spearman's rho (r$_S$)

Level of significance for a one-tailed test		
	0.05	0.01
Level of significance for a two-tailed test		
	0.10	0.02
N=29	0.312	0.433
30	0.306	0.425
31	0.301	0.418

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

(h)   Explain what is meant by a Type 1 error.

(2)

(i)   Use the information in the table above to explain why the psychologists did not think that they had made a Type 1 error in this case.

(3)
(Total 19 marks)

45

A researcher carried out an experiment to investigate misleading information. Participants were shown a photograph in which a man and a woman were talking. The photograph was then taken away and the participants were asked questions about it. Participants were randomly allocated to condition one or condition two.

Participants in condition one were asked:
Question A "How old was the youth in the photograph?"

Participants in condition two were asked:
Question B "How old was the man in the photograph?"

(a)   Why is **Question A** an example of misleading information?

_____

_____

_____

_____

(2)

(b)   Name an appropriate experimental design which could be used in this experiment. Explain why a repeated measures design would be unsuitable to use in this experiment.

Experimental design _____

_____

Explanation _____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(4)

(c)   Explain why it would be appropriate to use a pilot study as part of this experiment.

_____

_____

_____

_____

_____

_____

_____

Extra space _____

_____

_____

_____

(4)

(d)   In this experiment, participants were asked to look at a photograph rather than watch a live
        conversation. Explain **one** strength and **one** limitation of carrying out the experiment in this
        way.

Strength _____

_____

_____

Extra space _____

_____

Limitation _____

_____

_____

_____

Extra space _____

_____

(4)

(e)   Describe **at least one** other research study into misleading information. In your answer you
        should include details of what participants were asked to do and what results were found.

        _____

        _____

        _____

        _____

        _____

        _____

        _____

        _____

        _____

        _____

        _____

        _____

        Extra space _____

        _____

        _____

        _____

        _____

        _____

(6)
(Total 20 marks)

46

A researcher investigated the effect of age of starting day care on levels of aggression. Four-year-old children attending a day nursery were used. Each child was assessed by the researcher and given an aggression score. A high score indicated a high level of aggression. A low score indicated a low level of aggression. The maximum score was 50.

Mean aggression scores for four-year-old children who had started day care before the age of two or after the age of two

	Started day care before the age of two	**Started day care after the age of two**
Mean score	25	23

(a) Identify the operationalised independent variable **and** the operationalised dependent variable in this study.

Operationalised independent variable _____

_____

Operationalised dependent variable _____

_____

(4)

(b) What do the mean scores in the table above suggest about the effect of age at which children started day care on children's aggression?

_____

_____

_____

_____

(2)

(c) Name **one** measure of dispersion that the researcher could have used to describe the data.

_____

_____

(1)

(d)   Draw an appropriate bar chart to display the data presented in the table above. Correctly label your bar chart.
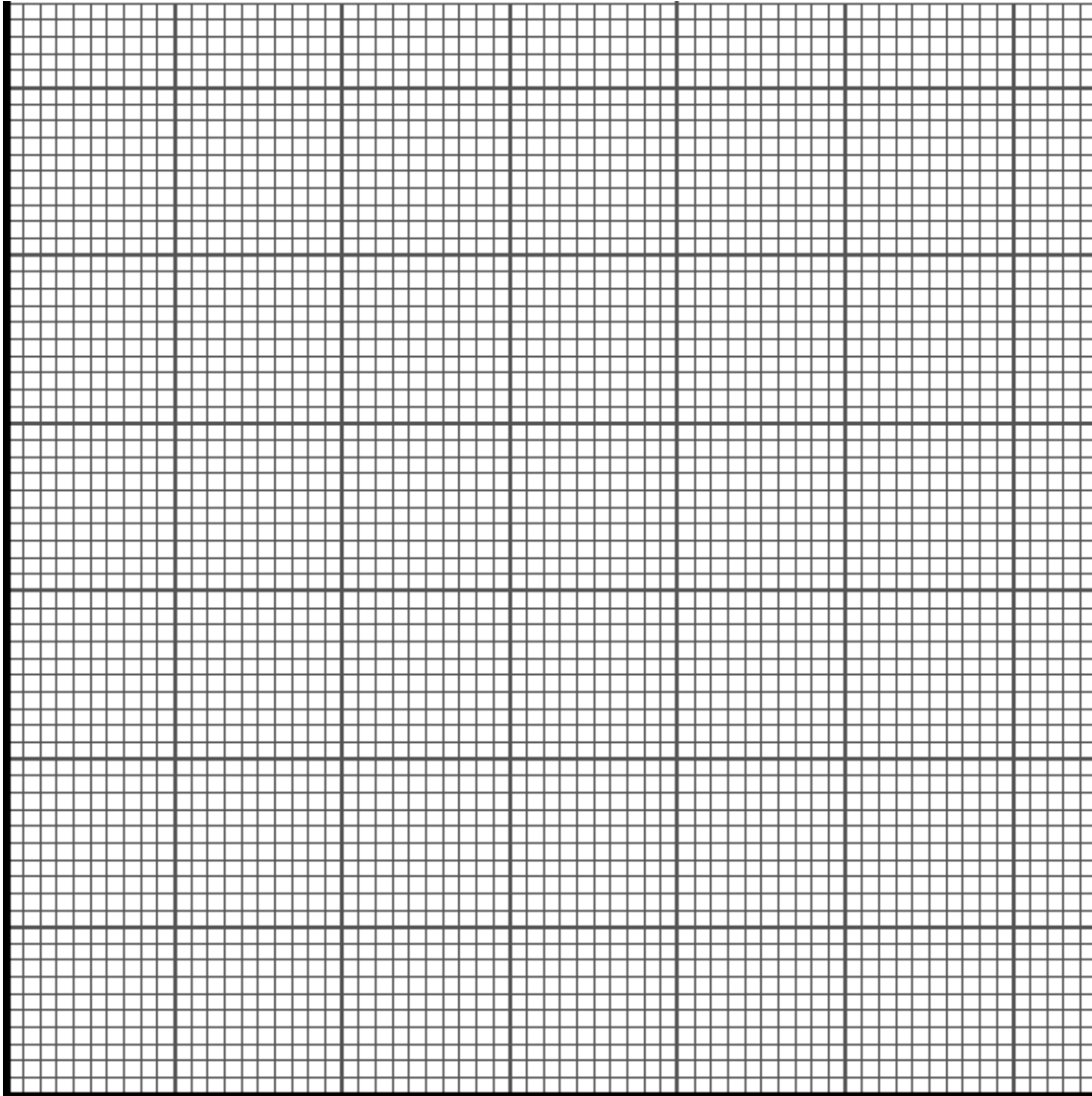
(3)

(e)   State an appropriate directional hypothesis for this study.

_____

_____

_____

_____

(2)
(Total 12 marks)

47   Give **two** reasons why Behaviourists do not collect qualitative data in their investigations.

(Total 2 marks)

48

A study was carried out to test the effectiveness of a new anger management programme. The programme had been designed by a team of psychologists working in a young offenders' institution.

Fifteen male offenders aged 17– 21 years took part in the programme. An anger score for each offender was obtained before the start of the programme. This score was based on a questionnaire designed by the psychologists. The questionnaire had 10 items. The maximum score was 50; the higher the score, the greater the level of anger.

The month-long programme of anger management involved 8 two-hour sessions.

Throughout the programme, the offenders were told to keep a diary of situations that made them angry and to record their anger in these situations. After the programme had ended, they were told to continue to keep their diary.

Two weeks later, after the programme had ended, a second anger score was obtained for each offender. The same questionnaire was used.

The data obtained are shown in **Table 1** below.

Table 1: Median anger scores and the ranges before and after the programme

	Before	After
Median	35	24
Range	15	17

(a) Explain why measures of dispersion are often used in addition to measures of central tendency to summarise data. Refer to the results of this study in your answer.

(2)

(b) A Wilcoxon signed ranks test was used to test for a significant difference between the anger scores at the start of the programme and after the programme had ended.

The calculated value of T was found to be 22.

Table 2: Critical values of T

Level of significance for two-tailed test	0.1	0.05	0.02
Critical value of T (when $N = 15$)	30	25	19

T must be equal to or less than the critical value to be significant.

Using **Table 2** above, explain whether or not the result is significant.

(2)

(c) Explain why the psychologists decided to use a Wilcoxon signed ranks test to analyse the data.

(3)

(d) Explain **two** possible reasons for asking each offender to keep a diary.

(4)

(e)   An independent researcher reviewed the design of the study and noted that there was no control group.

Explain how having a control group could have improved this study.

(3)

(f)   The independent researcher was also concerned that the psychologists had not checked the reliability and validity of the questionnaire used to measure the level of anger.

Outline how the psychologists could check the reliability **and** the validity of the questionnaire.

(5)
(Total 19 marks)

49   Outline what is meant by *replication* in research. Explain why replication forms an essential part of the scientific approach.

(Total 4 marks)

50   When investigating the conditioning of behaviour, Behaviourists such as Pavlov and Skinner collected quantitative data.

Outline what is meant by *quantitative data*. Give an example of quantitative data collected by a psychologist who has studied conditioning.

(Total 2 marks)

51   Case studies are a commonly used method of investigating abnormality.

(a)   What is meant by a case study?

_____

_____

_____

_____

(2)

(b)   Outline **one** weakness of using a case study as a method of investigation.

_____

_____

_____

_____

(2)
(Total 4 marks)

52

A maths teacher wondered whether there was a relationship between mathematical ability and musical ability. She decided to test this out on the GCSE students in the school. From 210 students, she randomly selected 10 and gave each of them two tests. She used part of a GCSE exam paper to test their mathematical ability. The higher the mark, the better the mathematical ability. She could not find a musical ability test so she devised her own. She asked each student to sing a song of their choice. She then rated their performance on a scale of 1–10, where 1 is completely tuneless and 10 is in perfect tune.

(a)   Suggest a suitable non-directional hypothesis for this study.

(3)

(b)   Why might the measure of musical ability used by the teacher lack validity?

(3)

(c)   Explain how the teacher could have checked the reliability of the mathematical ability test.

(3)

(d)   Explain why the teacher chose to use a random sample in this study.

The results of the study are given in the table below.

Mathematical ability test scores and musical ability ratings for 10 students

Student	Mathematical ability test score	Musical ability rating
1	10	10
2	2	9
3	9	3
4	6	6
5	3	9
6	10	2
7	2	1
8	1	8
9	8	4
10	4	7

(2)

(e)   In your answer book, sketch a graph to show the data in the table above.
       Give the graph an appropriate title and label the axes.

(3)

(f)   Discuss what the data in the table above and the graph that you have
        sketched seem to show about the relationship between mathematical ability
        and musical ability.

(3)

(g)   The teacher noticed that most of the students who were rated highly on
        musical ability were left-handed. The teacher is aware that her previous
        definition of musical ability lacked validity.

        Design a study to test whether there is a difference in the musical ability of
        left-handed students and right-handed students. You have access to a sixth
        form of 200 students.

        You should:

        •   identify the design that you would use

        •   explain an appropriate sampling method and justify your choice

        •   describe the procedure that you would use, including details of how you
               would assess musical ability

        •   write a suitable debrief for these participants.

(10)

(h)   In your answer book, draw a table to show how you would record your results.
        Identify an appropriate statistical test to analyse the data that you would collect.
        Justify your choice.

(3)
(Total 30 marks)

A psychologist wanted to investigate whether or not people are influenced by the opinions of others.

The psychologist selected 100 pupils from a secondary school to be participants in the study.

The psychologist showed participants a cake which weighed 350 grams. The task for the participants was to estimate the weight of the cake in grams.

Participants were allocated randomly to one of two groups.

In **Group A**, 50 participants were asked individually to estimate the weight of the cake. The psychologist gave each participant a blank piece of paper on which to write his or her estimate.

In **Group B**, the other 50 participants were asked individually to estimate the weight of the cake. This time, the psychologist gave each participant a piece of paper which contained a list of five weights (493 grams, 512 grams, 502 grams, 485 grams and 601 grams). The participants were told that these were the estimates given by five people and that they should write their own estimate below these other estimates.

The psychologist expected that participants in **Group B** would be influenced by the five other estimates. She expected that they would write down a weight similar to the five estimates on the piece of paper.

The median estimates for the weight of the cake are shown in the table below.

The median estimate for the weight of the cake (in grams) in Group A and Group B

	Group A **Estimate written on a blank piece of paper**	**Group B** **Estimate written below the list of five other estimates**
The median estimate of the weight of the cake (in grams)	348	510

(a)   What might the psychologist conclude from the median scores shown in the table above? Explain your answer.

_____

_____

_____

_____

_____

_____

(2)

(b)   Identify the independent variable and the dependent variable in this study.

Independent variable _____

_____

Dependent variable _____

_____

(2)

(c)   Explain how stratified sampling might have been used to select the participants in this study.

_____

_____

_____

_____

_____

_____

(3)

(d)   (i)   The psychologist allocated the participants randomly to the two groups that were used in this study.

Explain how the psychologist could have allocated the participants randomly to the two groups.

_____

_____

_____

_____

_____

_____

(2)

(d) (ii) Briefly explain **one** reason why random allocation of participants is important.

_____

_____

_____

_____

(2)

(e) The psychologist used an independent groups design in this study.

Explain **one** reason why it would not have been appropriate to use a repeated measures design in this study.

_____

_____

_____

_____

_____

_____

(2)

(f) Identify and briefly explain **one** ethical issue that the psychologist should have considered in this study.

_____

_____

_____

_____

_____

_____

(3)

(g) (i) After the study, the psychologist interviewed some of the participants in **Group B**.

Outline **one** type of interview that she might have conducted.

_____

_____

_____

_____

_____

_____

(2)

(g) (ii) Explain **one** limitation of the type of interview that you have outlined in your answer to (g)(i).

_____

_____

_____

_____

_____

_____

(2)
(Total 20 marks)

54 Explain what is meant by replicability. Why is replicability an important feature of science?
(Total 5 marks)

55 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)

56

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)

Mark schemes

1

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

One mark for identifying independent measures / groups / samples / unrelated design.

Up to 2 marks for an explanation of any relevant advantage of using this design in this study.
1 mark for an advantage, 1 mark for application to the study.

Possible answer:
As participants will either be approached by a confederate wearing uniform or a confederate in everyday clothing / as participants only take part in one condition (1), they are unlikely to guess the aim of the study / there are no order effects (1).

One mark can be awarded for an advantage that corresponds to an incorrectly identified design.

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

Independent variable: whether the researcher was dressed in everyday clothing or a uniform / type of clothing.

Dependent variable: whether participants pick up litter / obey (or not) / the number of people who picked up a piece of litter.

No credit for "obedience" or "level of obedience" or "amount of obedience".

- Award both marks for correct IV and DV that are not labelled but are in the order of the question.
- Award 1 mark for correct IV and DV that are not labelled and are not in the order of the question ie DV then IV.
- No credit for either IV or DV **alone** (if not labelled).

(c) **[AO1 = 1, AO2 = 2]**

AO1

1 mark for knowledge of likely outcome:
More people will pick up litter in Condition B than in Condition A / fewer people will pick up litter in Condition A than in Condition B.

AO2

2 marks for an explanation of the results based on application of obedience research to the scenario.

Possible answer: the confederate's uniform (1), increased the legitimacy / authority / status of the demands or order given (1) or similar. Credit use of evidence as part of the explanation eg Bickman, Milgram.

(d) **[AO2 = 1, AO3 = 1]**

AO2

1 mark for plausible application of the issue to the study.

AO3

1 mark for knowledge / identification of a relevant ethical issue. Likely issues: (lack of) consent, opportunity to withdraw, deception, treating people with respect, protection from harm, confidentiality, debriefing.

Possible answer: An ethical issue in the study is lack of consent (1) because the researcher did not ask the participants if they wanted to be in the study (1).

2

(a) **AO2 = 2**

Digit span is normally considered to be 7+ / –2, so Peter's was much shorter.
1 mark for simply stating his digit span was shorter than normal.
Second mark for an explanation of the difference, eg Peter's digit span of two items was much shorter than the average span of around 7 items.

(b) **AO2 = 4**

The MSM suggests there are separate ST and LT stores. Peter's short-term memory was impaired, but his long-term memory was not. This supports the idea of separate ST and LT stores, because one was damaged but not the other.

One mark for some reference to separate ST and LT stores. Three further marks for elaboration of the explanation.

Alternatively, candidates could suggest the evidence goes against MSM. If memory has to pass through the ST store to reach the LT store, it is likely that damage to the ST store would impair the transfer. Candidates could legitimately refer to evidence both for and against the model.

(c)   **AO3 = 4**

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, confidentiality or respect.
Candidates may point out that as the man has brain damage, his ability to give informed consent might be in doubt.
One mark 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.
For example: confidentiality (1 mark); keep the man's details private (1 mark); the psychologists should not use the man's name in published work, but could use his initials instead (2 further marks).

3

AO3 = 3

Simply identifying or naming one or more potentially relevant ways of dealing with the ethical issue – maximum 1 mark. For example, confidentiality, anonymity, debrief. Further marks for explaining how psychologists could deal with this ethical issue.

For example:

Right to withdraw (1 mark)
Participants should be reminded of their right to withdraw from the research (2 marks)
If participants are showing signs of distress, the psychologist should remind the participants of their right to withdraw (3 marks).

4

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**

Behaviour must be operationalised. Suitable behavioural categories could include crying, clinging to mother, smiling, playing independently etc.
One mark for each suitable behavioural category.

(b)   **AO3 = 2**

Candidates may refer to time sampling, CCTV and later analysis or ticking a box when the behaviour is shown. Unstructured observation could also be relevant.

One mark for a brief explanation. This could include demonstrating some understanding of the use of behavioural categories, eg draw a table and tick boxes. A further mark for elaboration, eg drawing the table and / or indicating when the boxes would be ticked.

(c) **AO3 = 2**

One reason for the psychologist carrying out a pilot study would be to check cameras were positioned appropriately. Another would be to check the suitability of the behavioural categories. Alternative relevant reasons should be credited.
One mark for a brief reason eg to check equipment. A further mark for elaboration as above.

5 (a) **AO3 = 3**

Candidates may point out that the % of secure attachment in all three countries is very similar, but that insecure attachments vary. Country one has the lowest % of insecure-avoidant but the highest of insecure resistant. Country three has the lowest % of insecure-resistant but the highest of insecure-avoidant.

One mark for a brief outline of one point. Two further marks for accurate elaboration of one point in detail or more than one point more briefly.

(b) **AO3 = 3**

Candidates may refer to limitations of the strange situation as a way of identifying attachment type. Alternatively they may focus on the differences in number of studies in each country.
They could also criticise the use of meta analysis.
One mark for a brief outline of a relevant criticism. Two further marks for elaboration.
For example, the findings for country two come from 18 different studies. We can't be sure that all of the studies were carried out in the same way (second mark). It is possible that the 'Strange Situation' or ways of categorising types of attachment were different in the different studies (third mark).
If candidates give more than one criticism, the best should be credited.

6 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 / 3 = 2**

One mark for one brief finding and a further mark for appropriate elaboration or for two brief findings or one mark for a slightly muddled answer.
On average, the treatment group showed greater improvement after the treatment than the no-treatment group. The average improvement score for the no-treatment group was very low suggesting that the treatment gains for the treatment group were not simply a result of the passage of time.

There was some variation in both groups as shown by the ranges but it was wider in the treatment group. The low range in the no-treatment group suggests that most people in this group had similar low improvement scores.

(b) **AO1 = 1, AO2 / 3 = 3**

One mark for identification of a suitable test and 3 further marks for an appropriate justification.The specification only requires knowledge of non-parametric tests. However, if a candidate names an independent t-test and justifies its use, this is perfectly acceptable. It is likely that most candidates will identify a non-parametric test. The most appropriate test is the Mann-Whitney and the justifications for its use are:

- independent groups design
- at least ordinal data
- differences.

(c) **AO2 / 3 = 2**

One mark for correctly identifying the likelihood and one further mark for an appropriate explanation or one mark for a slightly muddled answer.

The likelihood of making a Type 1 error is 5%. A Type 1 error occurs when a researcher claims support for the research hypothesis with a significant statistical test, but in fact, the variations in the scores are due to chance variables. If the level of significance is set at 5%, there will always be a one in twenty chance or less that the results are due to chance rather than to the influence of the independent variable or some other factors.

(d) **AO2 / 3 = 4**

Two marks for each reason. One mark for a basic identification and one further mark for elaboration.

Possible reasons include:

Expectations – the patients might expect the treatment to do them some good and it becomes a self-fulfilling prophesy.

Biased sample – even though the participants were randomly assigned to groups, the treatment group might, by chance have included more people with milder symptoms that were more likely to respond to treatment.

Other support – we do not know what other support/ treatment that the participants might have had over the 8 week therapy period.

(e) **AO2 / 3 = 4**

Two marks for the advantage and two marks for the disadvantage. One mark for simply identifying an advantage / disadvantage and the further mark for elaboration in the context of the study. Answers which are not set in context cannot achieve full marks.

Advantage: Much quicker to administer and to score – could all have been given out at the same time whereas the therapist has to conduct 30 time-consuming interviews; cheaper than interviews, ie in terms of the therapist's time; people might be more comfortable, and, therefore, more honest, if they have to write responses rather than face an interviewer (could work the other way as well – see disadvantages).

Disadvantage: Self-report questionnaires might not yield as accurate data as an interview – questions can limit range of answers and there are no additional cues, eg body language, participants might be less honest on a questionnaire than in a face-to-face interview.

Marks can be awarded for any appropriate advantages / disadvantages.

(f) **AO2 / 3 = 5**

Candidates should demonstrate understanding of some of the requirements of a good consent form. For full marks, it should be succinct, clear and informative.

It is likely to include some of the following information: treatment programme that is noninvasive; requirement to be assessed on current level of functioning; use of a trained therapist to conduct interviews; duration of the programme; requirement for re-assessment at the end of the programme; random allocation to a treatment or no-treatment group.

It should show awareness of ethical considerations, eg

- no pressure to consent – it will not affect any other aspects of their treatment if they choose not to take part
- they can withdraw at any time
- they can withdraw their data from the study
- their data will be kept confidential and anonymous
- they should feel free to ask the researcher any questions at any time
- they will receive a full debrief at the end of the programme.

For full marks, candidates must include a range of both procedural and ethical points.

AO2 / 3 Mark Bands (5 marks)
5 marks Effective Consent form demonstrates sound knowledge and understanding of research ethics.
4 – 3 marks Reasonable Consent form demonstrates reasonable knowledge and understanding of research ethics.
2 marks Basic Consent form demonstrates basic, superficial knowledge and understanding of research ethics.
1 mark Rudimentary Consent form is rudimentary demonstrating very limited understanding of research ethics.
0 marks No creditworthy material is presented.

(g) **AO1 = 2, AO2 / 3 = 2**

AO1: One mark for brief description, eg 'consistency' and one further mark for elaboration. Reliability refers to consistency over time. If a test, questionnaire, etc, is reliable, people tend to score the same on the test if they take it again soon afterwards.

AO2 / 3: One mark for a very brief answer, eg 'do another test' or 'test them again' or 'use another interviewer to check'. Two marks for some elaboration.

Reliability could have been checked by administering a valid and reliable questionnaire to the participants as well as interviewing them and then comparing the scores on the two measures. If the interview score was reliable, there would be strong positive correlation between the scores.

The interviews could have been filmed and given to another trained therapist to assess. A strong correlation between the scores given by each therapist would demonstrate reliability.

(h) **AO2 / 3 = 10**

For full marks, the method section should be written clearly, succinctly and in such a way that the study would be replicable. It should be set out in a conventional reporting style, possibly under appropriate headings. Examiners should be mindful that there are now different, but equally acceptable reporting styles. For example, candidates should not be penalised for writing in the first person. The important factor here is whether the study could be replicated.

There should be reasonable detail with regard to:

- design
- participants
- materials
- procedures.

AO2 / 3 Mark Bands (10 marks)
10 – 9 marks Effective Effective method section that demonstrates sound knowledge and understanding of investigation design. The design decisions are appropriate and the description provides accurate detail of the design, participants, materials and procedure of the study. Effective and appropriate report style.
8 – 6 marks Reasonable The method section demonstrates reasonable knowledge and understanding of investigation design. The design decisions are generally appropriate and the description provides reasonable detail of the design, participants, materials and procedure of the study. Generally appropriate report style.
5 – 3 marks Basic The method section demonstrates basic knowledge and understanding of investigation design. Some aspects of the design are appropriate. The description provides basic detail of some features of the study or rudimentary outline of the main features. Expression lacks clarity.
2 – 1 mark Rudimentary The method section demonstrates rudimentary knowledge or understanding of research. The report is weak, muddled or incomplete. Deficiency in expression results in confusion and ambiguity.
0 marks No creditworthy material is presented.

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 = 2**

The independent variable is the type of memory strategy or whether the participants were instructed to form a mental image to link the words, or to memorise the words.
One mark for the "memory technique" or "instructions to participants or memory strategy."
Two marks where the IV is operationalised as above.

(b) **AO3 = 2**

The dependent variable is number of words correctly recalled.
One mark for "recall" or a slightly muddled statement.
Two marks where the DV is operationalised as above.

(c) **AO3 = 1**

One mark for correct identification of an independent groups design.
Accept independent design / between-participant design.

(d) **AO3 = 2**

One mark for stating a strength, eg the same word list can be used for both conditions.
A further mark for explaining why this is a strength, eg the words for each group will be equally easy to learn. Credit any acceptable strength.

(e) **AO3 = 2**

One mark for identifying a suitable way of checking reliability, eg do it again. A further mark for elaboration, eg carry out the experiment again and look for a similarity in both sets of results.

8

AO3 = 3

Infants would be too young to respond to demand characteristics.

1 mark for a brief reference to mothers changing their behaviour or the cues in the investigation which lead to the change. 2 further marks for elaboration.

For example, the mothers' behaviour may change (1 mark). The mothers try to guess what the psychologist is looking at (1 mark), so they may be more attentive to their babies than when they are not taking part in this research (1 mark).

9

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) **AO2 = 4**

As Luca was in a poor quality orphanage for four years cognitive impairment is likely. Answers could also refer to Bowlby's MDH and possible consequences such as affectionless psychopathy and problems with later relationships. Reactive attachment disorder and physical effects would also be relevant.

1 mark or 2 marks for identification of possible negative effect(s), eg Luca may have problems forming relationships. [1 mark for identifying one negative effect, 2 marks for identifying two or more.]
Up to 2 additional marks for some elaboration of two or more effects or a more detailed elaboration of one effect.

(b) **AO3 = 4**

Strengths
Rich data, high ecological validity, investigates a situation which could not be set up for ethical reasons.

Limitations
Selection from large amounts of data may lead to observer bias.
Findings from one individual can't be generalised to others.

1 mark each for identification of a strength / limitation. Second mark for some elaboration. For example, an advantage of a case study is that it provides lots of detail (1 mark). This gives great depth and understanding of this single individual (2 marks).

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 = 1**

Identification of a positive correlation, +, +ve.

(b) **AO3 = 4**

The strength and weakness must be appropriate for investigating day care.

1 mark for stating a relevant strength / weakness. A further mark for some elaboration which would apply to day care.

For example, a strength would be it is an ethical way of collecting data (1 mark) because there is no manipulation of time the child spends in day care (2 marks).
A weakness you can't infer cause and effect (1 mark) because you can't be sure that time in day care causes the child to be disobedient (2 marks).

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.

(a) **AO3 = 2**

Identification of the mean or median.
There are no repeated scores in either list, so the mode would not be appropriate.
Justification for the mean could be that it used all of the available data, that it is a powerful / sensitive measure or that it is suitable for use with interval / ratio data.
Justification for using the median is that it is relatively unaffected by outlying scores.

(b) **AO3 = 2**

The specification names random, opportunity and volunteer sampling. Answers must relate to volunteer sampling which involves participants selecting themselves.

1 mark:
very brief suggestion, eg put up a notice.
2 marks:
some elaboration that could apply to the scenario, eg advertise on the staff room notice board, asking teachers to sign a list.

(c) **AO3 = 1**

Extraneous variables are anything other than the independent variable that could affect the dependent variable. In this study they could include participant differences, environmental variables such as temperature or noise, and experimenter variables including the way in which the research is conducted.

(d) **AO3 = 2**

The control must relate to the variable in (c).
1 mark for a brief suggestion. Second mark for some elaboration.
For example, control for participant differences by making sure both groups are similar (1 mark).
Control for participant differences by randomly assigning the participants to conditions (2 marks).

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.

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 = 2]**

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

For one mark, a statement containing both conditions of the IV and a DV.

No marks for expressions of aim, questions, correlational hypotheses or statements without two conditions.

Possible answers:

Directional: Participants will successfully shoot more netballs through the hoop in the presence of an audience than in the absence of an audience.

Note: Also accept a directional hypothesis opposite to the above.

Non-directional: There will be a difference in the number of netballs successfully shot through the hoop for participants who perform in the presence of an audience and participants who perform in the absence of an audience.

Credit null hypothesis.

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

One mark for an appropriate conclusion eg the presence of an audience has a positive effect on the performance of the task (or similar response) (AO3, 1).

One mark for the justification of the response eg the mean number of netballs successfully shot through the hoop is higher when there is an audience than when there is no audience (or similar response) (AO3, 1).

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

(Mean number of netballs) shot(s) through the hoop in the presence and absence of an audience
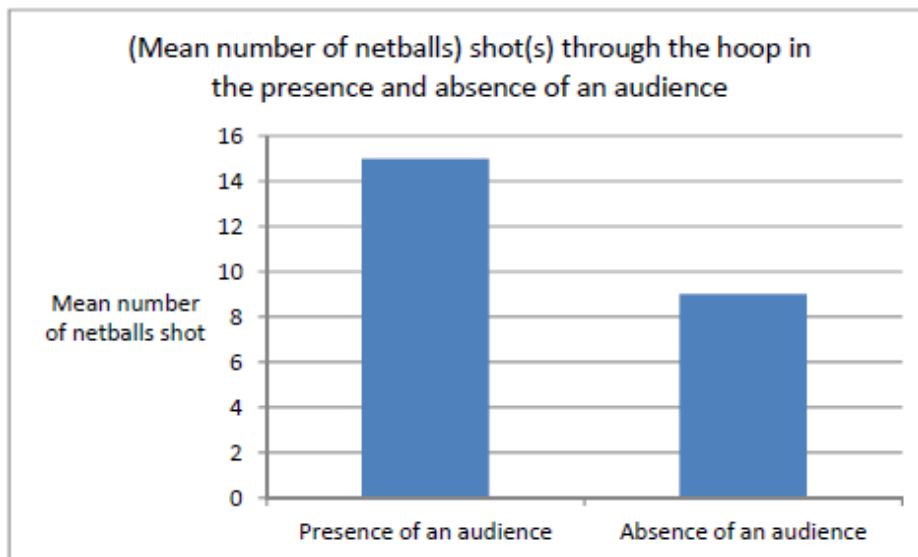
To gain the maximum three marks, candidates must provide the following:

* An appropriate title for the graphical display
* Appropriate axes and labelled eg presence / absence of an audience
* Plotting of data using a sensible scale / no penalty for joined bars.

Line graphs can be credited for title and axes only.

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

One mark for a relevant strength.
One mark for how / why it is a strength.

Likely strengths: more likely to be representative of the target population; able to generalise to the target population; to exclude researcher bias etc.

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

One mark for a relevant limitation of an independent groups design.
One mark for how / why in relation to this study eg individual differences in ability / height etc.

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

Up to two marks for an appropriate explanation of how the problem may have been overcome.

Possible answer: Using a repeated measures design (1) plus explanation of how this could be done or why this would improve the study.

Credit answers that refer to matched pairs / need for random allocation.

Can credit (ii) in respect of incorrect answer to (i).

(f)   (i)   **[AO3 = 1]**

One mark for correct knowledge of the term.

Possible answer: A variable other than the independent variable or an
additional / or another variable that might have an effect on the dependent
variable, only accept the word 'results' for DV (1).

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

One mark for stating that if EVs are not controlled for, then the results may be
confounded (1) the researcher does not know what is causing the effect (1) or
other suitable expansion eg effect on reliability or validity.

Do not accept - results will not be accurate.

Credit answers that refer to the study by way of illustration.

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

Instruction must be written verbatim for more than 1 mark.

Essential points   -   reference to presence of audience
                  -   reference to shooting (hoops) throwing (hoops)

Optional   -   where to stand
                  -   ethics
                  -   introduction of self

4 marks	Both essential and at least one optional point addressed clearly such that completion of the task in the experimental condition would be easily possible. Information should be clear, relevant, sensible and logically structured. Must be verbatim.
3 marks	Both essential points are addressed such that completion of the task in the experimental condition would be relatively easy. There may be deficiencies in clarity, some irrelevance, illogical sequencing or inappropriate content. Must be verbatim.
2 marks	Any two points are addressed. There may be omissions / irrelevancies / muddle such that completion of the task would be very difficult. Must be verbatim.
1 marks	There must be at least one relevant point (optional or essential). Information may be unclear / inappropriate / irrelevant / muddled such that completion of the task would be very difficult.
0 marks	No relevant information. Completion of the task would not be possible.

NB 2 - 4 marks = Verbatim Instructions

13

AO3 = 2

This graph shows a fairly strong negative correlation between stress and white blood cell activity / the immune system. As the stress increases the immune functioning decreases. The following can all receive a mark: direction, strength, and a description of their relationship.

14

AO3 = 4

Conclusions can include: there are two factors that influence conformity, the ambiguity of the task and the size of the majority. A large majority is most influential with an ambiguous task, but still exerts pressure even when the task is easy. However, a small majority has less effect and the type of task does not seem to be an important variable.

> **4 marks Accurate and reasonably detailed**
> Accurate and reasonably detailed answer that demonstrates sound knowledge and understanding of what the bar chart shows about conformity. There is appropriate selection of material to address the question.

> **3 marks Less detail but generally accurate**
> Less detailed but generally accurate answer that demonstrates knowledge and understanding. There is some evidence 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 of 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.

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.

(a) **AO2 / 3 = 1**

1 mark for correct answer – directional (one-tailed is acceptable).

(b) **AO2 / 3 = 3**

1 mark for correctly stating that the result is significant.
2 further marks for an explanation: the calculated value of $T = 53$ which is less than the value of 60 where N = 20 and $p \leq 0.05$ for a one-tailed test.

If the candidate states that the result is not significant, no marks can be awarded.

16

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)   **AO2 / 3 = 4**

For any credit, candidates must sketch a scattergram.
For full marks, candidates should provide an appropriate title for the scattergram, label each of the axes appropriately and plot the data accurately on the scattergram.

Each of the examples below is a full mark answer because:

* it is clearly a sketch of a scattergram

* the data are appropriately plotted

* the labels of the axes and the title taken together show full understanding of the nature of the data.

Relationship of verbal aggression ratings by two observers

Relationship of verbal aggression ratings by two observers
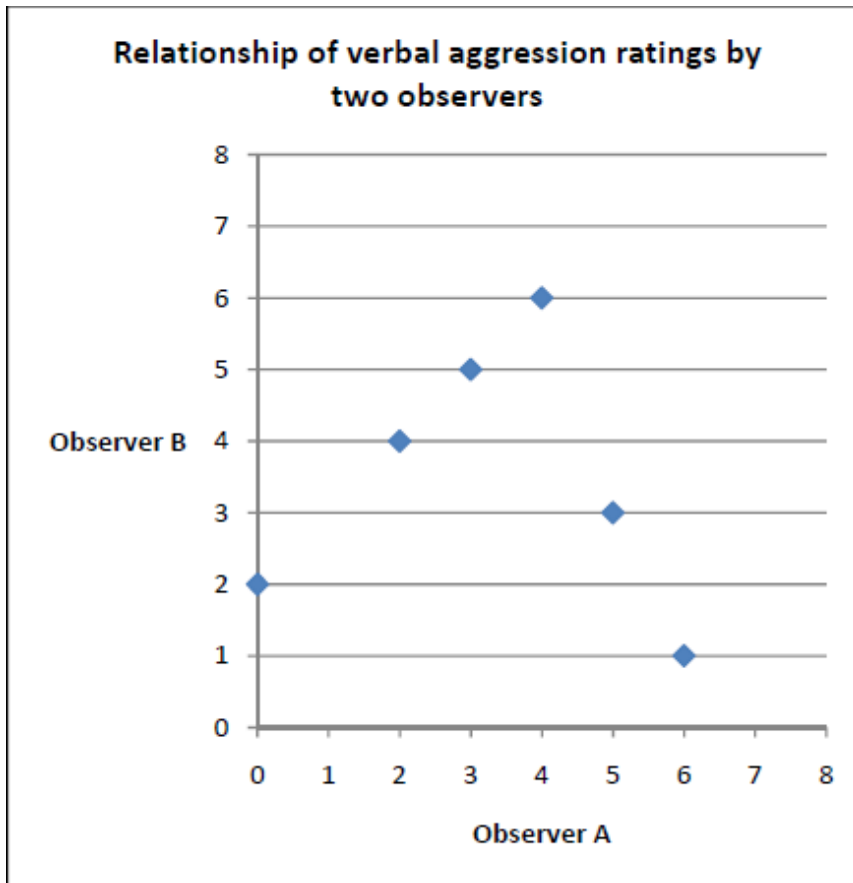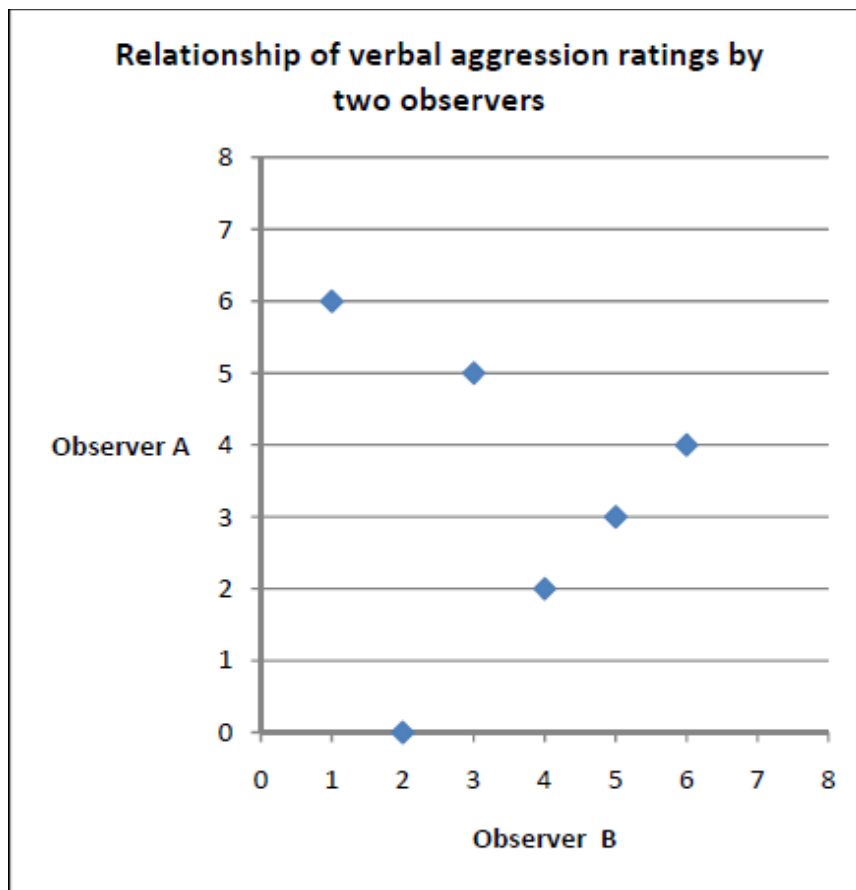
(b) **AO2 / 3 = 4**

For full marks, candidates should give a reasonably detailed explanation eg she is concerned because the observers should both recognise the same types of verbal behaviour as aggressive and you would expect their tallies to be very similar. In this case, the observers disagree in every 10-minute time interval even though they are both watching the same child and should be using the same criteria. In some time slots, there is a really big difference in the number of acts.
This suggests that the observers have interpreted the criteria differently or that, at certain times, one observer was more vigilant then the other (4 marks).

1 mark – 'because the observers do not agree with each other'.
3 further marks for elaboration.
Candidates who simply describe what is meant by inter-rate reliability can gain no marks.

(c) **AO2 / 3 = 3**

1 mark for identifying the appropriate test – Spearman's Rho or Pearson's (with appropriate justification).
2 further marks for explaining why it is appropriate ie the psychologist is testing for a correlation and the data that can be treated as ordinal.
Candidates can gain no marks on this question if their choice of statistical test is inappropriate.

(d) **AO2 / 3 = 4**

1 mark for a very brief answer eg 'better training for the observers'
3 further marks for elaboration.

There is a breadth / depth trade-off here. Candidates can elaborate on one improvement eg explain how the training might be improved or outline several improvements in less detail eg establish clearer criteria for categorising verbal aggression, filming the child so that the observers can practise the categorisation.

17

AO3 = 4

Strength: can study relationships between variables that occur naturally, eg stress from exams and students getting ill. Can measure things that cannot be measured 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 and the immune system, it does not mean that stress directly caused the immune system to become less effective: there may be another variable connecting the two.

Any other appropriate answer can get credit.
For each, 1 mark for a brief outline of the strength / weakness and a further mark for elaboration.

18

AO1 = 2

Peer review is the process by which psychological research papers, before publication, are subjected 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.'
One further mark for elaboration.

19

AO3 = 2

Credit advantages of different types of observation (eg covert) even if this is not explicit in the answer.
The researcher is able to look at the way people really behave. If people are unaware they are being watched they will not be susceptible to demand characteristics.
1 mark for brief reference to an advantage eg it's real behaviour.

2nd mark for some accurate / effective elaboration. Eg it looks at real behaviour rather that what people say they would do.

20

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) **AO1 = 1**

A brief definition of the term is sufficient for 1 mark eg a technique for analysing data according to themes or categories.
Candidates who simply write 'a way of analysing qualitative data' are not meeting the requirement to say 'what is meant by….?'

(b) **AO2 / 3 = 3**

• The psychologist would have identified a number of categories or themes by which to sort the drawings. Such categories/themes might include: the type of food depicted eg carbohydrate, protein; the state of the food eg cooked, raw etc; the portion size; the brightness of the colours used.

• He would have counted examples from each category to provide quantitative data.

• He could then compare the drawings according to these categories to see if there were changes over the four week period.

For full marks candidates can either outline three of the above or outline two with some elaboration.
For 2 marks candidates can either outline two of the above, or one with elaboration.
For 1 mark candidates simply outline one of the above eg "choose a theme like size".
Note: maximum 2 marks if no engagement with the stem.

(c) **AO2 / 3 = 5**

The form would need to contain sufficient information for the participant to make an informed decision about whether to take part or not. The form should contain some of the following:

- The purpose of the study.

- The length of time required of the participants.

- The fact that participants would have to be isolated in a research institute for the duration of the study.

- Details about the diet.

- Right to withdraw.

- Reassurance about protection from harm e.g. the availability of medical supervision.

- The requirement to undertake a series of psychological tests.

- Reassurance about confidentiality of the data.

It is not necessary for candidates to include all of the above points for full marks. However, in order to access the top band, candidates must engage with the study and include sufficient information on both ethical and methodological issues for participants to make an informed decision.
Maximum of 3 marks if no ethical issues are included.

AO2/3 = 5 marks
5 marks Effective The 'consent form' demonstrates sound understanding. Information is given in a clear and concise form and is explicitly relevant. The form includes sufficient information so that participants can make a fully informed decision including the right to withdraw.
4 – 3 marks Reasonable The 'consent form' demonstrates reasonable understanding. Information is given in a reasonably clear and concise form and is mainly relevant. The form includes sufficient information so that participants can make an informed decision.
2 marks Basic The 'consent form' demonstrates basic understanding. There is some lack of clarity and conciseness and material is not always relevant. There are some omissions such that participants would find it difficult to make a decision.
1 mark Rudimentary The 'consent form' is rudimentary and demonstrates very little understanding. Information is not given in a clear and concise form. The form has significant omissions such that a decision is not possible.
0 marks No creditworthy material is presented.

21

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

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

Up to 2 marks for description of both obsessions – recurrent / persistent thoughts / ideas / images / impulses and compulsions – repetitive behaviours / ritual acts / behaviour that reduces anxiety.
Accept physiological symptoms of anxiety.

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

1 mark for naming repeated measures design.
1 further mark for an elaboration of repeated measures design.
Possible answers:
Repeated measures design means that the same participants are used in both conditions of the study.
If the answer is related to the study described: This means that the children whose anxiety ratings are taken in the before therapy condition are the same children as those who provide the anxiety ratings for the after therapy condition.

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

Up to 2 marks for an explanation of one advantage of using repeated measures design.
The advantage of repeated measures design (in this study) is that there will be no participant variables (1) so any differences in performance (the median anxiety ratings before and after therapy) are more likely to be due to the manipulated variables / variables under test (therapy programme) than other variables so the validity of the results is increased.
Answers based on the idea that fewer participants are required than in other designs are relevant.

Note:
If the answer to (b) is incorrect **full credit** can be awarded for (c) if the advantage given matches the experimental design identified in the answer to (b).

22

AO3 = 4

For each advantage and disadvantage 1 mark for stating an advantage / disadvantage.
2nd mark for clear elaboration. Eg an advantage of using a questionnaire is that the data is easier to analyse (1 mark).
Eg An advantage of using a questionnaire is that it is more likely to produce quantitative data which is easier to analyse (2 marks).
Candidates may refer to this study, but this is not necessary.

23

AO2 / 3 = 5

Peer review is an important part of this process because it provides a way of checking the validity of the research, making a judgement about the credibility of the research and assessing the quality and appropriateness of the design and methodology. Peers are also in a position to judge the importance or significance of the research in a wider context. They can also assess how original the work is and whether it refers to relevant research by other psychologists. They can then make a recommendation as to whether the research paper should be published in its original form, rejected or revised in some way. This peer review process helps to ensure that any research paper published in a well-respected journal has integrity and can, therefore, be taken seriously by fellow researchers and by lay people.

AO2/3 marks
5 marks Effective Effective analysis and understanding. The answer is well focused and shows coherent elaboration. Ideas are well structured and expressed clearly and fluently. Consistently effective use of psychological terminology. Appropriate use of grammar, punctuation and spelling.
4 – 3 marks Reasonable Reasonable analysis and understanding. The answer is generally focused and shows reasonable elaboration. Most ideas appropriately structured and expressed clearly. Appropriate use of psychological terminology. Minor errors of grammar, punctuation and spelling only occasionally compromise meaning.
2 marks Basic Basic, superficial understanding. The answer is sometimes focused and shows some evidence of elaboration. Expression of ideas lacks clarity. Limited use of psychological terminology. Errors of grammar, punctuation and spelling are intrusive.
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. Errors of grammar, punctuation and spelling are frequent and intrusive.
0 marks No creditworthy material is presented.

24

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]**

1 mark for naming repeated measures design.
1 further mark for an elaboration of repeated measures design.
Possible answers:
Repeated measures design means that the same participants are used in both conditions of the study.
If the answer is related to the study described: This means that the children whose language interaction ratings are taken in the before therapy condition are the same children as those who provide the language interaction ratings for the after therapy condition.

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

Up to 2 marks for a brief discussion of one advantage of using repeated measures design.
The advantage of repeated measures design (in this study) is that there will be no participant variables (1) so any differences in performance (the median verbal interaction ratings before and after therapy) are more likely to be due to the manipulated variables / variables under test (therapy programme) than other variables / so the validity of the results is increased (1).

Answers based on the idea that fewer participants are required than in other designs are relevant.

Note:
If the answer to (a) is incorrect **full credit** can be awarded for (b) if the advantage given matches the experimental design identified in the answer to (a).

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

Up to 2 marks for a brief explanation of what the results indicate.
As the median verbal interaction ratings have increased (since the therapy programme) (1) that would seem to indicate the programme was effective (1).
Accept other plausible explanations of the difference in the medians.

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.

(a)   **AO3 = 3**

This is a volunteer / self–selected / voluntary sample.
Only people who read this newspaper could take part. The participants have chosen to take part, so it is a biased sample. Findings cannot be generalised to a wider population.
Volunteer sample or self-selected sample (1 mark).
Candidates can receive credit for a correct outline of a relevant disadvantage of volunteer sampling even if they have not identified the correct sampling method.
1 mark for a brief disadvantage. Eg It is biased.

2nd mark for some elaboration eg It is a biased because only some people who read the newspaper respond and they may not be typical of all readers.

(b) **AO3 = 3**

Any relevant issue should be credited.
Likely ethical issues include informed consent, right to withdraw or confidentiality.
1 mark for identification of a relevant ethical issue.
1 mark for a brief mention of a way of dealing with an ethical issue. Further mark for elaboration.
Eg Confidentiality (1 mark). Keep the participants details private (1 mark).
The psychologists should not use the participants' names in published work, or allow them to be identified in any way (2 marks).

26

AO3 = 3

Candidates are likely to refer to the fact that 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 eg to eyewitness testimony in court cases. There are often real consequences / emotional impact in real life which do not occur in laboratory investigations.
In a laboratory participants may show demand characteristics because they know they are in an experiment. This is less likely in real world settings.
Answers which refer to advantages of laboratory research or disadvantages of real world research are not relevant and should not receive credit.
1 mark for a brief explanation eg higher ecological validity.
Further marks for some elaboration as above.

27

AO3 = 2

Ethical issues could include:

- Informed consent
- Right to withdraw
- Protection from harm
- Anonymity
- Confidentiality
- Deception.

Any relevant ethical issue can receive credit, but the question requires candidates to explain the ethical issue.
1 mark for identifying the ethical issue and a further mark for the explanation. For example, confidentiality is an issue because when investigating stress the participant might be disclosing personal information that they want to keep private.

28

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) **AO2/3 = 6**

Candidates need to show that they understand what differentiates opinion from scientific evidence. They could mention some of the following:

- The teacher has only experienced one school in a particular catchment area so she has only observed a very limited number of 5 year-olds (issues of sampling and replicability).

- She has found out that children do not eat anything nourishing simply by chatting with the children. She has no corroborative evidence from eg parents (issues of objectivity).

- She uses vague phrases such as 'decent breakfast' without being clear what this means (operationalisation).

- She has generated a theory and made predictions based on flimsy evidence.

- She has not used any scientific method to lead to her conclusions eg a carefully controlled experiment, survey or observation.

- She has drawn conclusions about the effects of breakfast without considering other variables which might affect reading skills and behaviour.

AO2/3 Mark bands
6 marks Effective Explanation demonstrates sound understanding. Application of knowledge is effective and shows coherent elaboration. Ideas are well structured and expressed clearly and fluently. Consistently effective use of psychological terminology.
5 – 4 marks Reasonable Explanation demonstrates reasonable understanding. Application of knowledge is reasonably effective and shows some elaboration. Most ideas appropriately structured and expressed clearly. Appropriate use of psychological terminology.
3 – 2 marks Basic Explanation demonstrates basic, superficial understanding. Application of knowledge is basic. Expression of ideas lacks clarity. Limited use of psychological terminology.
1 mark Rudimentary Explanation is rudimentary, demonstrating very limited understanding. Application of knowledge 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.

(b)   **AO2/3 = 3**

In a random sample, every member of the identified population has an equal chance of selection. In this case, the sampling frame consists of the 400 five-year-old children attending ten local schools. In order to obtain a simple random sample, the researcher has to have the names of all 400 children and can then select using one of the following methods:

- **Random number tables** – random number tables are specially devised to meet the following criteria – they contain strings of numbers where each number has the same chance of being selected as any other and each number is independent of the others. Such tables are readily available in statistics text books etc or can be generated by the researcher using a computer program. The researcher assigns each child a number between 1 and 400. He enters the table at any place (he could close his eyes and point with a finger at a starting place) and then moves either horizontally or vertically to produce a string of random numbers. He records all the numbers which correspond to the 400 children until he has recorded a total of 100 non-duplicated numbers.

- **Computer selection** – This is a similar method where the computer does most of the work. A computer can generate an endless string of random numbers ie numbers which have no relationship to one another as a sequence. Each child's name is given a number and a random number generator program is used to produce the required sample size (in this case 100 participants).

- **Manual selection** – Using this method, the researcher has to put each name (or an assigned number) on a separate slip of paper and place them all in a container. The researcher then selects 100 slips from the container. The following conditions could apply: the container should be shaken between each draw; the slips of paper should all be the same size and folded in the same way so that one does not feel different from another; the selector draws 'blind' ie cannot see the actual slips of paper.

A simple definition of a random sample is not credit-worthy since it offers no explanation.\ Similarly, answers which only use the word 'random' as an explanation cannot gain credit eg He would choose 100 participants at random from the children. One mark for a very basic method eg 'he would take names from a hat / computer / random number table'. Two further marks for elaboration.

(c) **AO2/3 = 3**

Candidates could focus on:

- Even if a sample is random, it may not be truly representative of the population eg might all come from the same school, or be all boys or all girls.

- Practical limitations eg the time and effort needed to write out 400 slips for the manual method.

- Difficulties of obtaining a truly random sample eg even if the sample is selected randomly, parents might refuse to allow their children to participate.

Any plausible and appropriate answers should be credited.
Up to 2 marks for identification of limitations. For 3 marks, one or more limitations must be explained in reasonable detail.

(d) **AO2/3 = 5**

There are two requirements to this question, **why** operationalising variables is important and **how** to operationalise the IV and the DV. If a candidate only explains **how / why**, maximum 3 marks.

The terms' 'decent breakfast' and 'reading skills' are vague. It is important from the point of view of objectivity, replicability and control of extraneous variables to make sure that these terms are closely defined.

Suggestions as to how the psychologist might do this could include the following:

The researcher needs to specify the exact composition of the breakfast (possibly by doing a pilot study or a literature search to identify the components of breakfast most likely to bring about behavioural / cognitive change). He probably also needs to specify the time at which it is consumed. The researcher needs to use a standard reading test which should be administered to all the participants at the beginning of the study and at the end – the dependent variable is likely to be the improvement score.

(e)   **AO2/3 = 2**

Reasons are:

- a test of difference

- data (scores from a reading test) are at least ordinal, this would include ordinal / interval and / or ratio

- independent design.

One mark for each appropriate reason (maximum 2 marks).

(f)   **AO2/3 = 2**

It would have been more difficult to use a matched-pairs design because of the number of relevant factors that would need to be controlled (eg gender, intelligence, parental attitudes / income / education, experience of pre-school education, number of siblings in family etc). There is a relatively small pool of children available (ie 400) and it could be difficult to match on all these factors. It would also be very time-consuming; it could be quite expensive to carry out the necessary surveys; it could be quite intrusive collecting such information from parents.

One mark for a basic explanation eg "Because it is difficult to match participants appropriately".
One further mark for elaboration.

(g)   **AO2/3 = 2**

One mark for identifying an appropriate issue and second mark for explaining how it could be addressed.

The most likely issue is confidentiality which could be addressed by ensuring that all scores on reading scales and all personal information are anonymised.
There are also ethical problems involved in denying the control group breakfast although it is more difficult for candidates to suggest a way of addressing this – perhaps to put only those children into the control group who do not eat breakfast anyway, restricting the study length to a short period of time and, if the study results support the hypothesis, to provide free breakfasts to these children for the rest of the academic year.
Parental consent is excluded because it is given in the stem so answers which offer this as an issue cannot gain credit.

(h) **AO3 = 12**

Design should be written clearly, succinctly and with sufficient detail for reasonable replicability.

Candidates will not receive credit for details included in the stimulus material. These include using a random sample of 100 children, gaining parental consent and selection of a Mann Whitney test.

To access marks in the top band candidates must state an appropriate hypothesis in which "playground behaviour" is clearly operationalised. The hypothesis could be directional or non-directional.

Given the wording of the question, a correlational hypothesis is not credit-worthy, however, the rest of the answer should be marked on its merits.

Likely aspects of "playground behaviour" would include activity levels, aggression, cooperative play etc.

An attempt to operationalise "a healthy breakfast" should be credited. However, candidates could assume this had already been done by the psychologist.

As this is an observational study any of the following, together with appropriate justification, would be credit-worthy:

Is the observation covert or overt?

Where are observers positioned? (In playground, watching from window?)

Is a video recording of the children used? How will this be analysed (eg content analysis)?

Do the students who observe know what the children ate for breakfast?

At what times of day does the observation take place?

How many children are observed? (Candidates could justify using a smaller sub-sample of the 100 children in the original study).

How long does each observation last?

Will the observers use a behavioural check list / tally chart?

Will more than one observer observe each child? If so, what training will be given and what checks for inter-observer reliability will take place?

Reference to time sampling or event sampling.

Credit any other relevant material.

AO2/3 Mark bands
12 – 10 marks Effective design A design that demonstrates sound knowledge and understanding of observational research. The selection and application of research techniques is appropriate. The description provides sufficient detail for most aspects of the study to be implemented. Some design decisions are justified effectively.
9 – 7 marks Reasonable design The design is reasonable and demonstrates knowledge and understanding of some aspects of observational research. The selection and application of research techniques is mostly appropriate. The description provides sufficient detail for some aspects of the study to be implemented. Some design decisions are justified.
6 – 4 marks Basic design The design is basic and demonstrates limited knowledge and understanding of aspects of observational research. The selection and application of research techniques are sometimes appropriate. Some basic design decisions/features of the study are described but there may be significant omissions, lack of clarity and possibly some implausible suggestions that severely limit implementation. Justifications of the design are limited.
3 – 1 marks Rudimentary design The design is rudimentary. Design decisions are muddled and or mostly inappropriate and are not justified. Description lacks clarity. The study could not be implemented.
0 marks No creditworthy material.

29 **AO3 = 4**

Strengths of questionnaires:

- Can be given to a large sample of people.

- Participants can answer the questionnaire without the need for the researcher to be present, so reducing experimenter bias.

- Compared with interviews they are easy to use, the researcher doesn't need any special training to use them.

For each strength, 1 mark for identifying the strength explicitly relevant to questionnaires and a further mark for explaining why it is a strength.
The first bullet point is an example of a 1-mark answer as there is no explanation of **why** it is a strength. The other two examples are 2-mark answers as there is some explanation. Candidates could also make reference to the advantage of specific types of questions on the questionnaire, i.e. open or closed.

30

AO3 = 4

Possible conclusions:

- Some therapies are more successful for some disorders than other therapies.

- For some disorders there is little difference between the different types of therapy.

- For disorder A, biological therapies are much more effective than any other therapy.

- Psychoanalytic therapies are not very effective, except for disorder C, but even here, there is little difference between the three types of therapy.

Findings should only be credited if they are explicitly linked to a conclusion.
Candidates could offer several conclusions or only a few, there is a breadth / depth trade-off.

4 marks Accurate and reasonably detailed
Accurate and reasonably detailed answer that demonstrates sound interpretation and understanding of what the bar chart shows about the effectiveness of different types of therapies. 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 interpretation and understanding. There is some evidence of selection of material to address the question.

2 marks Basic
Basic answer that demonstrates some relevant interpretation 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 understanding. Selection and presentation of information is largely or wholly inappropriate.

0 marks
No creditworthy material.

31

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**

This is a natural experiment because the independent variable, whether the children attend nursery or are looked after by a child minder is a naturally occurring variable.
1 mark for brief or muddled explanation eg the children went to nursery anyway.
2 marks for accurate explanation eg the psychologist didn't decide which type of day care the children attended.
Reference to an independent variable should be credited but is not essential for 2 marks.
Answers which refer only to research carried out in a real life situation should not receive credit.

(b) **AO3 = 2**

There are a number of ways in which the children's aggressive behaviour could be measured. These include observation and recording the behaviour in categories. Alternative methods such as interviews or questionnaires given to parents or teachers would be relevant.
1 mark for naming or identifying a relevant method eg observation, using a tally chart.
2nd mark for further detail relevant to measuring aggressive behaviour in children.

(c) **AO3 = 4**

Candidates are likely to refer to informed consent from parents and maintaining confidentiality. Parents' right to withdraw child or data and possible deception would also be relevant. As children are five years old consent from the children could be credited as well as informed consent from parents. In addition to BPS guidelines, issues such as socially sensitive research should be credited.
For each, 1 mark for identification of a relevant issue.
2nd mark for some elaboration eg implications of the issue.

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.

(a) **AO3 = 1**

Independent groups. Accept "independent" but not "individual". Unrelated and between groups / subjects are also credit-worthy.

(b) **AO3 = 2**

There may be differences between the groups as there are different participants in each condition.
More participants are required than for a repeated measures design.
1 mark for very brief or muddled answer eg individual differences or needs more participants.
2 marks as above.

(c) **AO3 = 2**

The graph shows the cognitive interview is effective. There were more correct statements made after the cognitive interview than after the traditional interview. There was no difference in the number of incorrect statements made.
1 mark for a very brief or muddled statement. Eg It shows it's effective.
2 marks for some elaboration with reference to either correct statements, incorrect statements or both.

(d)   **AO2 = 2**

The answer should clearly relate to one or more of the main techniques used in a cognitive interview:

> Context reinstatement
> Recall from a changed perspective
> Recall in reverse order
> Report everything.
> The main additional features of the enhanced cognitive interview:
> Encourage to relax and speak slowly
> Offer comments to help clarify their statements
> Adapt questions to suit the understanding of individual witnesses.

1 mark for simple identification of a relevant cognitive technique, or a very brief suggestion eg "tell me everything you saw."
Further mark for application or elaboration. Eg "Please tell me everything you can remember about the robbery from the film you have just seen", or "Report all the details you can remember even if they don't seem very relevant." 2 marks.
Answers which could not relate to the film or robbery should be restricted to a maximum of 1 mark.

(e)   **AO3 = 4**

Investigator effects occur when the researcher's behaviour or characteristics influence the research in some way. This includes the way the presence of the researcher may influence the participants.
0 marks for an incorrect answer or one which simply re-states the words eg "how the investigator effects research".
1 mark very brief or muddled answer eg the researcher's influence.
Maximum 2 marks for an accurate understanding of investigator effects but no reference to this experiment.
Candidates whose explanation of possible investigator effects in this experiment shows understanding of the term can be awarded 4 marks, even though there is no separate definition.
Eg The psychologist may expect the cognitive interview to be more effective than the standard interview. This may be unconsciously communicated to the participants though mannerisms such as smiling or frowning. (4 marks). There are different routes to 4 marks. Candidates could be stronger on definitions or stronger on application.

33   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

D
C
A.

34

AO3 = 4

- The two averages are very similar, suggesting that both therapies are as good as each other.

- The range of each group is very different. This suggests that for some people Therapy A was very beneficial, but for others it had little benefit. For Therapy B, there was a much smaller range, suggesting that it has a similar effect on improvement for all the patients.

4 marks Effective interpretation of data Effective interpretation that demonstrates sound knowledge of what the data shows, with reference to both the average and the range.
3 marks Reasonable interpretation of data Reasonable interpretation of what the data shows; or effective interpretation of either the average or the range.
2 marks Basic interpretation of data Basic interpretation of what the data shows.
1 mark Rudimentary interpretation of data Rudimentary, muddled interpretation of the data, demonstrating very limited knowledge. Or reference to, for example, larger range/higher average/similar range.
0 marks No creditworthy material.

35

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

'There is an association between birth order and choice of career' = 2 marks.
A directional hypothesis is not credit-worthy. Reference to a relationship / correlation cannot gain credit.

Although technically, the psychologist is looking for an association, candidates can gain credit for expressing the hypothesis in terms of a difference eg 'There is a difference in career choice depending on birth order.'

2 marks for a clear hypothesis, 1 mark for a hypothesis which lacks clarity.

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

One mark for identifying a sampling method.
One mark for a brief explanation of how to obtain the sample eg 'by advertising for lawyers or artists to come forward'. One further mark for elaboration eg 'by explaining that adverts would have to be placed in appropriate journals etc to attract these particular categories of participants'

Candidates who identify a sampling method but describe it incorrectly can be awarded 1 mark.

(c) **AO2 / AO3 = 12**

This is a 12 mark question but marks are allocated to each of the required components as follows:

• An appropriately labelled table = 2 marks

1 mark for a table that displays the data in the question.
2 marks for a table which includes data relating to non first-born children. Totals are not required for the 2 marks.

Table: Table to show the career choices of first born and non-first born children

	Artists	Lawyers	Totals
First born	20	35	55
Not first born	30	30	60
Totals	50	65	115

• a sketch of an appropriately labelled bar chart = 3 marks

For 3 marks, candidates need to display the data relating to first born and non-first born career choices on a bar chart. They should label axes correctly and draw the columns to the correct approximate height for a sketch.

For 2 marks, candidates display data as above but labels are missing or lack clarity. For 1 mark, candidates graph the data supplied in the question relating to first born career choices only.
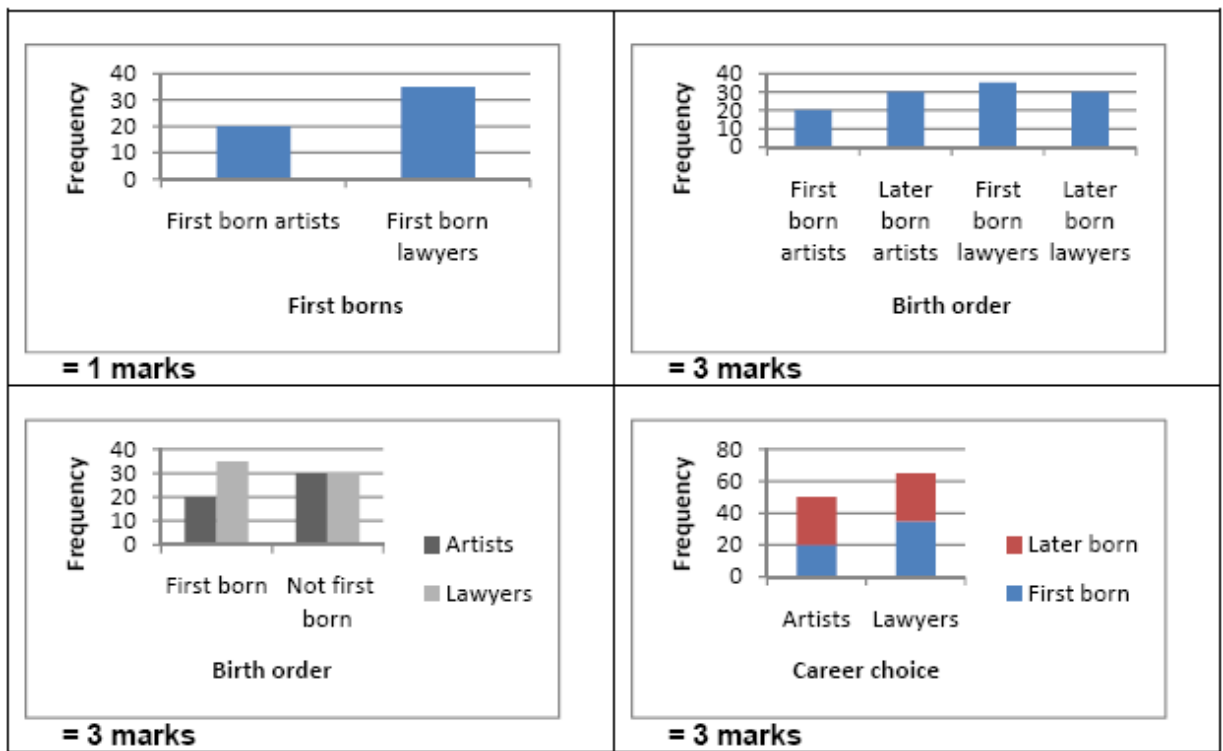
NB Labelled axes but no bars = 0 marks.

- identification of appropriate statistical test and justification = 1 + 2 marks

An appropriate test here is the Chi-squared.
Justification gains 2 marks. Any two correct reasons from:

- data are independent
- level of measurement is nominal
- test of association / difference is required.
- identification of appropriate significance level = 1 mark.

The most likely significance level is 5% ($p ≤ 0.05$). Candidates are not asked to justify their choice. Candidates who choose a more stringent level can achieve marks but they must then follow this through when they make their statement of results.

Candidates who erroneously report 0.05% or $p = 0.5$ do not gain credit for level of significance but can achieve credit for the statement of results in relation to the hypothesis.

- a statement of the results of the statistical test in relation to the hypothesis = 3 marks.

For full marks, the candidate should state whether or not they can accept the hypothesis (or they can express this in terms of rejecting the null hypothesis) at a given significance level and refer to the observed and critical values.

Where candidates choose an inappropriate value from the table but interpret that value correctly, they can gain 2 marks.

The critical value for x^2 (df =1 p 0.05 (two-tailed)) is 3.84. As the observed value of x^2 2.27 is less than the critical value, we cannot reject the null hypothesis. There is not an association between birth order and career choice.

36

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

They wanted to clarify some of the issues raised by previous research where some studies had shown that red facilitated tasks and other studies had shown the opposite. They believed that one way to reconcile these different findings was to look at particular cognitive tasks eg ones which required attention to detail and to compare them with tasks which tap into very different skills eg creativity and thus to narrow down the benefits of providing red backgrounds.

One mark for a brief answer eg 'they wanted to investigate the effects of colour on performance.' One further mark for elaboration, in relation to colour and / or performance.

(b)   **AO1 = 2**

Candidates need to show understanding of reporting conventions. The introduction is an important part of the report that provides background information on theories and studies relevant to the investigation. One mark for a brief explanation of the purpose eg 'It provides background information', and one further mark for elaboration or for other detail such as reviewing methodological issues or how the current aims / hypothesis were derived.

(c)   **AO1 = 1**

In this question, candidates are not required to relate validity to this particular study so a general definition of validity is acceptable. Definitions of specific types of validity (eg population validity) can also gain credit.
Validity refers to how well a test or a piece of research measures what it says it measures = 1 mark.
Answers such as 'truth' or 'whether it is true' 'legitimacy' or 'accuracy' = 0 marks.

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

The Canadian researchers who actually undertook this study suggested the following possible practical applications:

- to help decide what colour to pick for an educational facility

- to help decide what colour enhances persuasion in a consumption context

- to help decided what colour enhances creativity in a new product design process.

Any plausible practical applications are credit-worthy.
1 mark for identifying an application and 1 further mark for elaboration.
'You could use particular colours for pages in textbooks' = 1 mark.
'Red might be used in textbooks covering analytical subjects like maths' = 2 marks.

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

If the researchers had judged the toys themselves, they might have been biased in favour of their hypothesis. There are no objective criteria for what makes a toy either practical or original.
Independent judges would be able to decide between themselves on a set of criteria and then apply them to the toys made by the participants.

Some candidates might interpret 'independent judges' in this question to mean judges who do not confer with one another. In this case, an acceptable answer would be that they could not conform with one another when making their judgement.

One mark for a brief explanation, eg to avoid experimenter bias, and one further mark for elaboration, eg if the researchers judged the toys themselves.
An answer explaining the value of rating the toys should be credited.

(f) **AO2 / AO3 = 5**

Candidates need to use the details in the description of the study to write an appropriate set of instructions for potential participants.

The instructions should be clear and succinct. They must:

- explain the procedures of this study relevant to participants

- include a check of understanding of instructions.

They should also use language appropriate for a formal document and be as straightforward and courteous as possible.

This is not a consent form so explicit references to ethical considerations are not necessary for full marks. However, it is perfectly acceptable to include comments such as 'you are free to withdraw from the study at any time.'

AO3 Mark Bands
5 marks Effective The **standardised instructions** provide accurate detail of the procedure and go beyond the information given in the question eg provide details of time allowed.
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.

37

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 = 2**

Confidentiality could be maintained by making sure individuals are not identifiable when reporting the case study. This could be done by using a different name or initials, avoid publishing details of address, schools etc.

1 mark for identification of a relevant way.
2nd mark for some elaboration (which could be an example) or for identification of a second way of maintaining confidentiality.

(b) **AO2 = 2**

Psychologists may use psychological tests eg IQ testing. They could observe his behaviour in different situations. They might interview people, such as family members, to find out the circumstances of his early life.

1 mark for simply naming any appropriate techniques such as IQ test, observation or interviews.
2nd mark for some elaboration.

(c) **AO3 = 4**

The main limitation is that each individual, and their experience, is unique and the results cannot therefore be generalised to others. Evidence from an individual's past may be difficult to verify.
Researchers may get to know the individual well, which may lead to loss of objectivity.

Although description of specific case studies is not relevant, candidates may refer to examples as part of and explanation of limitations.

AO3 Knowledge of limitations of case studies
4 marks Accurate and reasonably detailed Accurate and reasonably detailed answer that demonstrates sound knowledge of at least one limitation.
3 marks Less detailed but generally accurate Less detailed but generally accurate answer that demonstrates relevant knowledge of at least one limitation.
2 marks Basic Basic answer that demonstrates some relevant knowledge of one or more limitations, but lacks detail and may be muddled.
1 mark Very brief / flawed Very brief or flawed answer demonstrating very little knowledge of limitations.
0 marks No creditworthy material.

38

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**

10 children.

(b) **AO3 = 2**

Number of aggressive acts shown during a 10-minute observation. Aggression score on a questionnaire completed by parents.

1 mark for brief or slightly muddled answer, eg ask children's mothers how aggressive they are.
2nd mark for accurate elaboration, eg ask children's mothers how aggressive they are on a scale from 1 to 10.

(c) **AO3 = 3**

This is a correlational study, not an experiment, and correlation does not prove a causal relationship. Children who are already aggressive may be put into day care for longer, or another variable, such as parents' divorce, may contribute to high scores for both time in day care and aggression.

1 mark eg correlation doesn't prove cause or simply stating that aggression may be a result of other factors.
2 further marks for accurate elaboration as above.
Credit any explanation that focuses on causal inference.
An alternative approach is to argue that some research contradicts this claim. 1 mark for there is contradictory evidence and further marks for accurate reference to relevant research.
0 marks for reference to media exaggeration or individual children.

39

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 identification of laboratory experiment.

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

Up to 2 marks for an explanation of an advantage of a laboratory experiment.

Possible answer: As the research takes part in a controlled environment, the researcher can eliminate the possible effect of extraneous variables.

Answers are likely to focus on advantages based on increased control of variables / increased causality / replicability.

The advantage can be credited if it corresponds with the answer in 6.

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

Independent variable: whether the list of points was positive then negative or negative then positive / the order of the points / information. Answer must imply that there is more than one condition.

Dependent variable: whether (or not) they said Alex was 'friendly' / the number of participants who said Alex was 'friendly' / number of 'friendly' responses.

No credit for 'level of friendliness'.

- Award both marks for correct IV and DV that are not labelled but are in the order of the quesiton.
- Award 1 mark for correct IV and DV that are not labelled and are not in the order of the question ie DV then IV.
- No credit for either IV or DV **alone** (if not labelled).

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

One mark for identifying independent measures / groups / samples / unrelated design.

40

AO3 = 4

The graph shows a strong negative correlation between score on depression scale and weeks of treatment. The more treatments the lower the depression. However, there also seems to be a plateau, where between 2-3.5 weeks there is very little change in depression.

1 mark for each of the following:

- Strength (it is a moderately strong / strong correlation)

- Direction (negative)

- Description of the relationship (the longer the treatment the lower the depression score)

- Indication of plateau / change in direction.

41

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.

AO2 / AO3 = 10

Candidates are required to design an experiment to test the effects of different kinds of music on concentration. Examiners need to ensure that they read the completed answer thoroughly before starting to award marks.

Candidates are directed to three pieces of material which should be included within their proposed design. They are required to:

• Operationalise the independent and dependent variables

• Provide details of how they would control extraneous variables

• Describe the procedure they would use with sufficient detail for the study to be carried out.

Candidates are told that they **must** use a repeated measures design. If they do not, they can only access marks for the IV and DV.

In this experiment:

IV and DV – 2 marks

• The independent variable is type of music (for example classical and rock). Candidates should suggest two different types of music.

• The dependent variable is a measurement of concentration. Candidates can use the suggested word search task but must state how it is to be measured (for example, time taken to complete a word search or number of errors made). Alternatively, candidates may suggest their own DV.

Award one mark for operationalising each variable.

Controls – 4 marks

An important element in a repeated measures design is the control of order effects.

• Counterbalancing is the most likely procedure to control order effects. Half of the participants should carry out the 1st concentration task with music 1 followed by the 2nd task with music 2. The other half should complete the concentration task with music 2 first and follow this with music 1.

• The two concentration tasks should be matched for difficulty.

• Alternatively candidates could argue for randomisation or a time delay between the tasks.

Other relevant controls eg volume of music, time allocated for task should be credited. Answers which make no reference to the control of order effects maximum 2 marks.

<u>Procedure – 4 marks</u>

Procedural information should provide detail of how to go about conducting the study (ie what participants are required to do). Candidates could approach this task at a macro level ie from getting consent to debriefing or at a micro level ie the specific procedure for one participant. Other creditworthy material could include:

- Dealing with ethical issues

- Sampling

- Details of conditions and allocation to them

- Standardised instructions

- Data collected.

Note: there are only 4 marks available for the procedure and therefore candidates do not need to address all of the above to gain full credit.

42

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**

0 marks for a non-directional or correlational hypothesis.
The DV in this experiment is number of pictures correctly identified. Hypotheses where the DV is incorrect (eg number of participants who identified 10 pictures) = 0 marks.
1 mark if not fully operationalised, eg Participants who used the memory improvement strategy did better.
2 marks Participants who use a memory improvement strategy will correctly identify more pictures / objects than participants who do not use a memory improvement strategy.

(b) **AO3 = 1**

In an independent groups design a different group of participants is used in each condition.

1 mark =   Different participants / people in each condition / group

                Different / separate groups

                Random allocation to groups / conditions.

0 marks =   Different / separate conditions

                Independent participants / people

                Different experiments.

(c) **AO3 = 2 + 2**

Strength

The participants are naÏve because they take part in only one condition, so are less likely to show demand characteristics. There are no order effects such as practice or fatigue because participants take part in one condition.

Limitation

Individual variation, because there are different participants in each condition. More participants are needed than if a repeated measures design was used.

In each case 1 mark for very brief or slightly muddled strength or limitation, 2nd mark for appropriate elaboration of explanation.
0 marks for simply stating there are different participants in each condition.

(d) **AO3 = 3**

A pilot study is used to check aspects of the research such as whether participants understand standardised instructions, whether timings are adequate etc. It allows the researcher to try out the study with a few participants so that adjustments can be made before the main study, so saving time and money.

1 mark for a very brief explanation. Further marks for appropriate elaboration or identification of other reasons. Eg

To check it works. 1 mark
To check the standardised instructions are clear. 2 marks
To check the standardised instructions are clear enough for the participants to understand what they are required to do in the experiment. 3 marks

This question requires an explanation of why a pilot study was used, so a description of what a pilot study is (small scale study carried out before the main research) is not credit-worthy on its own. Candidates do not have to refer to a specific aspect of this experiment.

However, to gain full marks the answer must be relevant, so reference to checking sound levels for example would not be relevant.

(e) **AO3 = 2**

The standard deviation (spread of scores) is larger in the condition with the memory improvement strategy.

Candidates who use the word 'range' to suggest spread should be credited.

1 mark The standard deviation is larger in the condition with the memory improvement strategy.
2 marks The data shows the dispersion or spread of scores is larger in the condition with the memory improvement strategy.

43

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

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

Up to four marks are awarded for discussing advantage(s) of using a laboratory experiment in this case.

The most likely advantages of the laboratory setting in this experiment include:

- Control over extraneous variables. The lab setting meant that extraneous variables could be minimised. In this experiment, outside factors such as waiting time, noise and stress (which would be difficult to control in a field experiment) were removed.

- Ethical issues. In this case, the testing of memory in a field experiment would have involved ethical issues including deception of patients or withholding of information.

Candidates may also refer to other advantages of the laboratory setting such as replicability. These can receive full credit if they contextualised within the scenario.

Award four marks for an answer which provides accurate and detailed discussion of relevant advantage(s) with a clear link to the scenario.

Award two or three marks for an answer which includes discussion of relevant advantage(s), with some reference to the scenario.

Award one mark only for an answer which merely identifies one or more relevant advantage(s) of a laboratory experiment appropriate to this scenario.

Advantages of laboratory experiments which are not relevant to this study cannot gain any credit eg use of technical equipment.

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

- One mark for correctly identifying the Mann Whitney U test or independent t test.

- One mark awarded for an accurate reason for choice (for Mann Whitney these are: test of difference, independent groups design / independent data or data which can be treated at an ordinal level).

44

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

(a) **AO2 / AO3 = 1**

One mark for an accurate reason: The decision to use a directional hypothesis was based on findings of previous research which pointed to an effect in a particular direction ie memory is poorer with age.

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

A suitable directional hypothesis would be 'There is a negative correlation (relationship) between age and recall accuracy rating'.

- 3 marks for a fully operationalised hypothesis as above

- 2 marks for a directional correlational hypothesis that identifies age and recall as the two variables but is not fully operationalised

- 1 mark for a directional hypothesis where the variables are not identified ('there will be a negative correlation') or where the hypothesis lacks clarity.

Award zero marks for a non-directional or null hypothesis or any hypothesis predicting a difference or association.

(c) **AO1 = 1**

One mark for an accurate definition: The extent to which results or procedures are consistent or simply 'consistency'.

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

One mark for identification of a way of ensuring reliability. By far the most likely answer here is inter-rater reliability.

Two marks for some explanation/elaboration: using two separate psychologists and comparing them.

Three marks for an accurate and clear explanation: using two separate psychologists to rate the typed accounts for accuracy and comparing / correlating the ratings to see how similar they are.

Candidates could make a case for test retest which would involve the same psychologist re-examining the ratings after a period of time.

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

Award one mark for correct identification of one of each type of data.

- Qualitative data: the patient's responses, the typed accounts, the doctor's notes.

- Quantitative data: the ratings of recall accuracy on a scale of 1 – 10, ages of patients.

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

One mark for each accurate reason given:

- the researchers are testing for a correlation or a relationship between two variables.

- the data is to be treated as ordinal because the recall accuracy is in the form of ratings.

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

One mark for stating that the result is significant.

Second mark for explaining that -.52 exceeds .306 ($p \leq 0.05$, n=30 for a one-tailed test).

(h) **AO1 = 2**

One mark for a brief or muddled answer which hints at rejecting HO / accepting the H1 in error.

Two marks for explaining the term: where the researcher rejects the null hypothesis (or accepts the research / alternative hypothesis) when in fact the effect is due to chance – often referred to as an error of optimists.

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

3 marks for a clear explanation which is based on comparison of the calculated value of rs with the critical value at the 0.01 level of significance and indicates competence in use of statistical tables as follows:

- A Type 1 error is unlikely because the calculated value of rs (-0.52) exceeds the critical table value at both the 0.05 and 0.01 level for a one-tailed test.

- The chance of a Type1 error occurring is therefore less than 1%.

- This means that the researchers can be 99% certain that the results obtained are not due to chance.

Award one mark for a brief explanation (it is significant at 0.01).
Award two further marks for an explanation which refers to two of the above points.

Award one mark for stating that the obtained value (-0.52) exceeds the critical value (0.306) by a reasonable margin.

45

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)   **AO2 = 2**

This is an example of misleading information because the word "youth" suggests the man was young.
1 mark for a brief or muddled answer eg identifying the use of the word "youth" or "it refers to age".
2 marks for some accurate elaboration eg the answer clearly states that the man was young or a youth or suggests that his age may influence the answer.
Credit answers which state that the information is misleading because the question suggests there was a youth in the picture, when in fact there was only a man and a woman.

(b)   **AO3 = 1 mark + 3**

1 mark for independent (groups, measures, participants or subjects or between subjects or participants) design or unrelated design. 0 marks for individual.
1 mark for matched (groups, measures, participants, subjects).
A repeated measures design could not be used because participants would take part in both conditions. This would be inappropriate because their answer to one question would affect their answer to the other question. Candidates may point out this would make it easy to work out the aim of the experiment and so could lead to demand characteristics.
1 mark for a very brief / muddled answer eg "they couldn't answer both questions." "It could lead to demand characteristics."
Further marks for accurate detail. "It could lead to demand characteristics because they would know what the experiment was about." 2 marks
"Participants couldn't take part in both conditions because their answer to one question would affect their answer to the other question." 3 marks

(c)   **AO3 = 4**

In this experiment it could be used to check how long the participant should be given to look at the picture so that the timing could be changed if it was too long or too short. It could check the participants understand the questions asked and what they are required to do. It could also be used to ask a few participants about their experience of taking part.

Credit any appropriate answer which could apply to this investigation. No marks are awarded for a definition of a pilot study. Explanations which do not relate to this investigation maximum 2 marks.

AO3   Application of knowledge of research methods
4 marks Accurate and reasonably detailed Accurate and reasonably detailed explanation that demonstrates sound understanding of why a pilot study would be appropriate in this study.
3 marks Less detailed but generally accurate Less detailed but generally accurate answer that demonstrates sound understanding of why a pilot study would be appropriate in this study.
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.

(d)   **AO3 = 4**

One strength of using photographs in the investigation would be control of variables eg the same pictures could be shown for the same amount of time. Candidates may refer to a limitation of the live conversation.
One limitation is lack of validity. The findings cannot be generalised to real life situations where other factors such as changing facial expressions and gestures could be relevant.

For each strength and limitation 1 mark for stating a strength / limitation. 2nd mark for accurate elaboration.

(e)   **AO1 = 6**

Candidates must select a research study (studies) which relates to misleading information / leading questions, so research into weapon focus should not be credited.
Candidates are likely to refer to Loftus and Palmer's (1974) experiment where the verb in the critical question was changed (smashed, collided, bumped, hit or contacted.) Other relevant research would be Loftus and Palmer asking participants "Did you see any broken glass?" and Loftus et al's (1978) study using a red Datsun and Stop or Yield signs.
Research into anxiety and EWT is not relevant unless the candidate refers to misleading information such as Yuille and Cutshall where the witnesses to a real-life shooting appeared resistant to misleading information.
Research relating to age could also be relevant. Eg Warren et al (2005) found children were more likely to be influenced by misleading information than adults.
Credit any relevant research.

Examiners are reminded that there is a depth / breadth trade-off.

AO1   Knowledge and understanding
6 marks Accurate and reasonably detailed Accurate and reasonably detailed answer that demonstrates sound knowledge and understanding of the procedures and findings of one or more relevant research studies.
5 – 4 marks Less detailed but generally accurate Less detailed but generally accurate answer that demonstrates relevant knowledge and understanding of the procedures and findings of one or more relevant research studies.
3 – 2 marks Basic Basic answer that demonstrates some relevant knowledge and understanding of the procedures and findings of one or more relevant research studies but lacks detail and may be muddled.
1 mark Very brief/flawed Very brief or flawed answer demonstrating very little knowledge of the procedures and findings of one or more relevant research studies.
0 marks No creditworthy material.

46

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 = 4**

The independent variable is age at which the children started day care, or whether the children started day care before or after age 2.
1 mark for age.
2 marks where the IV is operationalised as above.
The dependent variable is aggression score as assessed by the researcher.
1 mark for aggression.
2 marks for aggression score, measure of aggression, level of aggression.

If either IV or DV is identified but not entirely clear – 1 mark.

(b) **AO3 = 2**

1 mark for pointing out the difference is small or the age of starting day care didn't make much difference to mean aggression score.
1 mark for stating the children who started day care before age 2 had a higher mean score than those who started after the age of 2.
1 mark for saying both groups mean score was approximately half the maximum.
Maximum 2 marks.
Eg "The mean aggression score was slightly higher for children who started day care before the age of 2." 2 marks
Candidates can gain 2 marks by two brief points or one point elaborated.

(c) **AO3 = 1**

1 mark for: range
                semi-interquartile range
                interquartile range
                standard deviation or
                variance.

Do not credit: deviation or interquartile.

(d) **AO3 = 3**

0 marks if the candidate has not drawn a bar chart.
1 mark if the candidate has drawn a bar chart but the scale is clearly inappropriate and not correctly labelled.
2 marks if the candidate has drawn a correctly labelled bar chart but the scale is clearly inappropriate; or the candidate has drawn an appropriate bar chart but the labelling is incomplete eg vertical axis refers to mean score or aggression score rather than mean aggression score.
For full marks the bar chart should indicate a small difference. Both bars and the vertical axis should be correctly labelled.

(e) **AO3 = 2**

0 marks for a non-directional hypothesis or a correlational hypothesis.
1 mark if either variable is not operationalised eg day care makes children more aggressive or the answer is slightly muddled.
2 marks for eg Children who start day care before age 2 have higher aggression scores than those who start day care after age 2, or Children who start day care at a younger age will be assessed as more aggressive than children who start day care at an older age.
Credit a directional hypothesis in the opposite direction.

47

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 = 2]

One mark for each reason applied to the behaviourist approach. For two valid points only - allow 1 mark. Likely answers will probably include: subjectivity of qualitative data / open to interpretation; cannot be replicated; not open to quantification and statistical analysis; specific so not amenable to generalization; not associated with the scientific approach. Credit relevant comparison to quantitative data collection.

48

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 an explanation.
Inform researcher of spread of scores.
One mark for link to the study.
The programme did not seem to affect people in each condition differentially as spread of scores in each condition is similar / large.
The offenders' anger behaviour showed a wide variation both before and after the programme.
Accept spread is greater in the after condition than before.

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

One mark for statement that the result is significant.
One mark for rationale: the calculated value of T is 22 and is less than the critical value of 25 (at the 0.05 level of significance).
Can accept not significant at 0.02 level.

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

Maximum of 3 marks can be obtained from: one mark for each reason or two marks for each reason with explanation.

- Reason – ordinal level of measurement / non-parametric
  Explanation – self reports / estimated scores of anger; data might not be normally distributed
- Reason – design of the study is related / repeated measure
  Explanation – same people before and after
- Reason – research involving differences between the 2 sets of scores
  Explanation – anger scores before and after

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

Up to two marks for each reason and explanation. Likely points: as an aid to memory; a qualitative measure to supplement the quantitative data collected; to check the validity of the questionnaire; part of the therapeutic process / increased self-awareness.
Accept other valid reasons.
One mark for an appropriate reason and one mark for an explanation of the reason.

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

Up to three marks for outlining how a control group could have improved this study: it is not possible to tell if the programme has caused the improvement; improvement could have been due to the programme or due to spontaneous recovery; by using a control group would make it more scientific; scores can be taken at the same times (pre-programme / post-programme) as in an experimental condition; post programme differences between the groups can inform if programme is effective; can be more confident in inferring cause and effect .

Allow a maximum of one mark for the general purpose of a control condition: acts as comparison / baseline measure where nothing changes

Accept 'scientific' and 'validity' only if justified.

(f)   **[AO3 = 5]**

Up to 5 marks for addressing both reliability and validity. One of these marks must be for reference to statistical testing.

A maximum of three marks if only one of these is addressed.

One mark for identifying a type of validity: face validity; concurrent validity. Accept also content validity; criterion validity; predictive validity.
Only accept identification mark if it matches how the assessment would be carried out.

One mark for outlining how the assessment would be carried out. For example in concurrent validity, scores from the questionnaire are compared with those from an established but similar questionnaire known to have good validity to see if the results are similar.

One mark for the statistical testing (checking for a <u>positive correlation</u> / applying Spearman's rank order correlation).

One mark for identifying a way of assessing reliability. Most likely is test-retest but accept split-half reliability and item analysis.
Only accept identification mark if it matches how the assessment would be carried out.
Do not accept inter-rated / inter-observer reliability.

One mark for outlining how the assessment would be carried out. For example in test-retest, the same group of young offenders would be tested using the same questionnaire at a later date to see if the findings remained consistent.

One mark for the statistical testing (checking for a <u>positive correlation</u> / applying Spearman's rank order correlation).

The one mark for statistical testing can only be credited once.

49 **[AO3 = 4]**

Allow one mark for outlining what is meant by replication: being able to copy the original experiment / research using the same methods and procedures in order to see if results are same or similar.
One mark for reference to method (doing the same way).
One mark for reference to results (if similar or not).

Up to two marks are for explaining the role of replication in the scientific approach. Candidates may focus on two reasons briefly or may choose to elaborate on a single reason. Possible points: increases confidence in results; theory strengthened through repeated attempts at refutation / falsification; reliability / consistency of findings supported; a condition of validity; generalisability increased.
Credit answers embedded in an example.

50

[AO3 = 2]

One mark for knowledge of what is meant by quantitative data: data in numerical form; data representing how much there is of something.
The second mark is for an example of a specific behaviour: such as Pavlov - a measure of the amount of saliva produced by the dog; or Skinner, the time taken to respond to a stimulus (to peck a button / press a lever / jump over a barrier / reach a goal box etc) or number of responses to a stimulus (pecks, lever presses in a given time period).

51

(a) **AO3 = 2**

A case study is an in-depth study of one person or a group of people over time. It is usually carried out in the real world. They are idiographic and very individualistic.

1 mark for a basic answer and a further mark for elaboration. Note that the answer does not have to be linked to psychopathology.

(b) **AO3 = 2**

Weakness of case study:

- Findings cannot be generalised very easily to other individuals.

- The information gathered is often based on retrospective data, which might not be accurate.

- Because it is very difficult to replicate a case study they lack reliability.

- As case studies are unique situations it is difficult to generalise (to other situations).

1 mark for a basic answer and a further mark for elaboration.
Note that the answer does not need to be linked to psychopathology.
Reference to ethical issues are not creditworthy.

52

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

A suitable non-directional hypothesis would be 'There is a correlation (relationship) between pupils' scores on a test of mathematical ability and pupils' scores on a test of musical ability'.

3 marks for a fully operationalised non-directional hypothesis.
2 marks for non-directional hypothesis that identifies both variables but does not operationalise them.
1 mark for non-directional hypothesis where the variables are not identified.
No marks for a null or directional hypothesis or one referring to association or difference.

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

The main issue is that the teacher has made up her own test:

- This involved subjective judgement on the part of the teacher who rates the students' musical ability. Her judgement may not reflect real differences in musical ability and is likely to differ from other people's judgement and / or any absolute criteria for tunefulness.

- Lack of reliability in rating musical ability would compromise the validity of the measure.

- As the students can choose the song they will sing, the rating of ability could reflect the teacher liking / dislike of the song rather than the student's ability.

- The rating may be invalid as the students selected songs which varied in difficulty so the tunefulness reflected the difficulty of the song not the students' ability.

- Operationalising musical ability as tuneful singing is a very narrow measure. Someone can have musical ability such as playing an instrument which would not be reflected by this measure.

1 mark for identifying an appropriate reason.
2 further marks for elaboration, explanation of why it is a problem, how it might affect the result or for further reason(s).
Note that 3 marks can be awarded for one reason elaborated or more than one reason in less detail.

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

In the case of the maths test candidates could refer to split half or test retest as methods of checking reliability. They could also refer to checking the reliability of scoring by using two separate markers for the test and comparing the scores. Credit any other appropriate suggestion.

1 mark for identifying an appropriate method or a brief explanation eg 'repeat the maths test'.
2 further marks for appropriate elaboration.

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

The teacher chose to use a random sample because it would probably be more representative of the whole GCSE group than if she had used an opportunity or volunteer sample. Candidates could also say that she had ready access to her target population making it convenient for her to select a random sample.

No credit for definition of a random sample.
1 mark for a brief or muddled reason (it is not biased).
2 marks for a reason that clearly points to an advantage of random sampling. This could be achieved through a comparison with another method (it is less likely to be biased than a volunteer sample).

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

Credit should only be awarded for scattergraphs. Other graphs gain 0 marks.

1 mark for appropriately plotted scores.
1 mark for an appropriate title.
1 mark for correctly labelled axes.

Relationship between scores on a mathematics test and scores on a test of musical ability

Relationship between scores on a mathematics test and scores on a test of musical ability
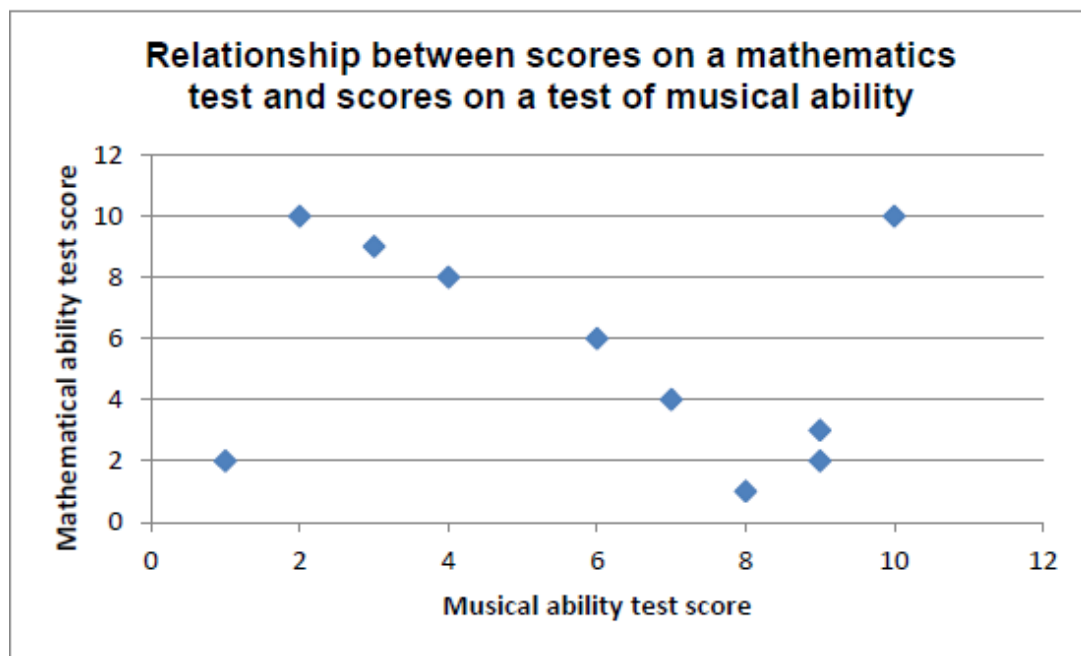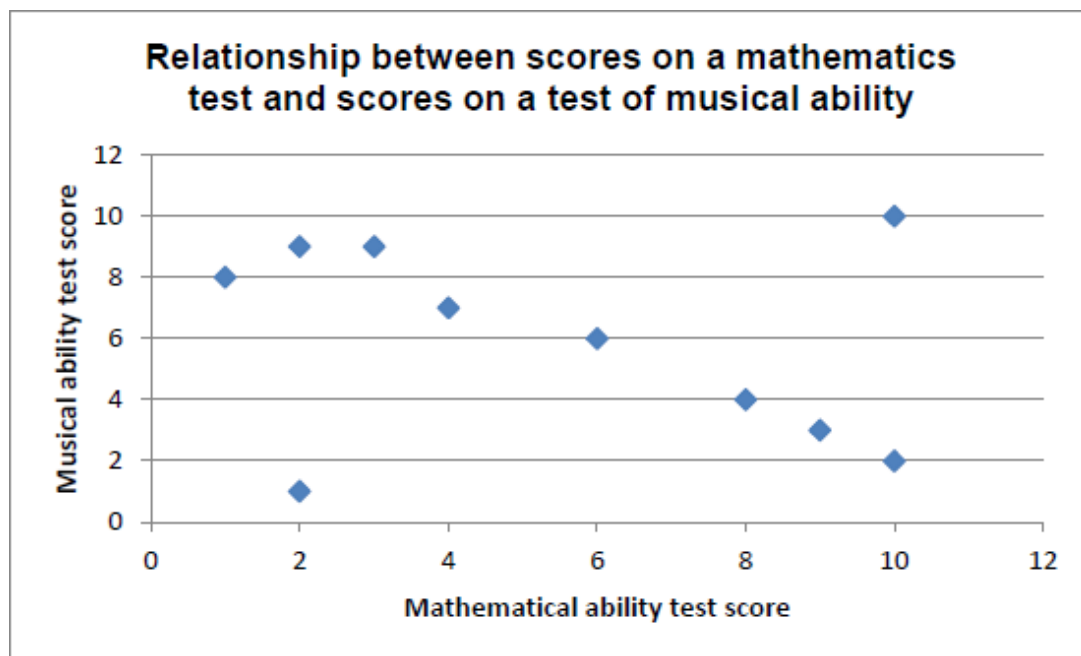
(f)   **AO2 / AO3 = 3**

Up to 3 marks for a discussion of the relationship between mathematical and musical ability. Likely points include:

- The graph seems to show a negative correlation between mathematical and musical ability.

- This means that high scorers in mathematical ability tend to achieve low scores on musical ability and vice versa.

- The presence of two strong outliers, means that the actual correlation is very weak and closer to zero.

- Comment on the small sample size which limits the conclusions that could be drawn.

- Credit can be achieved for plausible interpretations of the strength of the correlation which are justified (ie looks moderate to strong or the outliers make it weak in practice) or those based on rough calculations (around -0.2).

1 mark for a very brief answer eg negative correlation or zero correlation.
2 further marks for elaboration/discussion this could be focused on one point in detail or several points in less detail.

(g) **AO2 / AO3 = 10**

In this question, candidates are asked to design a study to test if there is a difference between left-handed and right-handed students in musical ability.

Design – 1 mark

• Award 1 mark for identification of an appropriate design (independent measures or matched pairs).

Sampling – 2 marks

• Award 1 mark for explaining an appropriate sampling method and 1 further mark for justifying why this method would be appropriate. As left-handed people are less common in the population than right-handed people this needs to be addressed in the sampling method.

Procedure and assessment of musical ability – 4 marks

Award 1 mark for procedure, 1 mark for assessing musical ability and two further marks for elaboration of either or both of these.

• Description of the procedure eg each participant will be given a standardised musical ability test, participants should be tested within a controlled environment, with minimal noise or distraction.

• Students are required to suggest a plausible alternative method of assessing musical ability to the one in the stem (eg singing a short, novel phrase played on the piano). Further credit could be given for stating that the test should be identical for all students or for explaining how it will be assessed.

Debrief – 3 marks

• Award up to 3 marks for writing a debrief. This could include the aim of the study, thanking participants for taking part, asking if they have any questions, relevant ethical considerations.

• If this is not suitable to be read out to participants, maximum 1 mark.

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

Award 1 mark for a clear table appropriate for the study described in (h).

Musical ability scores:

Participant number	Left handed	Right handed
1		
2		
3		

Award 1 mark for the identification of an appropriate statistical test for the proposed design.
Award 1 mark for one correct justification eg a test of difference, at least ordinal level data.

53

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 an appropriate conclusion. One further mark for explanation / justification of the conclusion.

Possible answer: Participants in Group B were influenced by / conformed to / went along with / were led by / affected by the other estimates they had seen (1 mark for conclusion) because the median estimate for Group B was closer to the scores on the list (1).
Accept 'higher than' Group A.

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

One mark for identification of the IV, ie whether participants saw 5 other supposed estimates or not

OR

whether participants wrote their estimates on a blank piece of paper or below 'other' estimates.

One mark for identification of the DV, ie the estimated weight of the cake. Accept 'estimate'.

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

Up to 3 marks for an explanation of how stratified sampling is achieved. Answer may be expressed in terms of pupils from a secondary school.

One mark each for reference to the following points:

- identification of strata
- proportions to be worked out
- random selection (within the strata)

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

Up to 2 marks for an explanation of how the psychologist could have allocated the participants randomly to the two groups.

Possible answer: Each participant in the sample could be given a number which is put into a hat. The first 50 numbers drawn out of the hat would be allocated to one of the groups.
Or
50 A's and 50 B's could be put in a hat and each participant draws out a letter. The letter designates the group.

Credit alternative method

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

One mark for an appropriate reason. One further mark for an explanation of that reason. Random allocation will reduce the likelihood of individual differences affecting the outcome.

Possible answer: Random allocation reduces researcher bias (1) and makes it more likely that the manipulation of the IV caused the results (1).

Accept increasing validity but do not accept reference to increasing reliability.

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

One mark for a plausible reason.

Likely answers: to prevent the participants guessing aim of the study / naivety; to prevent order effects.
Credit reference to the illogical nature of using the same participants in both groups in this study.

One mark for relating the reason to the study.

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

One mark for knowledge of an appropriate ethical issue.

Likely answers: deception; consent; confidentiality, etc.
One mark for elaboration of the ethical issue or why the issue is important.One mark for relating the issue explicitly to the study.

Note: INFORMED consent is only to be given **any** credit if candidate explains how it would be totally inappropriate in this study. Such a response could achieve full marks if candidate goes on to explain how use of debrief would be a way of compensating for lack of information beforehand.

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

Up to 2 marks for an outline depending on detail.
Likely answers: a structured interview (1) in which the questions asked are prepared and in a set order. (1)
An unstructured interview (1) in which the researcher has an aim but no predetermined questions. (1)

Accept other types of interview, eg face-to-face, group, one-to-one, telephone.

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

Up to 2 marks for a limitation of the chosen interview method that is elaborated.

Structured interview: possible answers – no deviation from predetermined questions is allowed (1) this means that an unexpected response cannot be pursued for more detail. (1)
Or, structured interviews are more formal (1) and this might make it a more uncomfortable experience for the interviewee. (1)

Unstructured interview: possible answers – the data are difficult to analyse (1) as all the responses are individual and therefore hard to collate / summarise. (1)
Or, unstructured interviews may produce irrelevant information (1) because it is easy for the interviewer and / or interviewee to go off the focus or point of the investigation (1).

Accept other valid responses eg comparison with other types of interview.

54

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.

AO1 = 5

Replicability is the ability to check and verify scientific information.

Candidates could explain replicability as:

- the ability to repeat the method to assess if similar findings are achieved

- the ability to achieve similar findings.

Award 1 mark for a brief definition of replicability.

Tautological definitions eg merely stating that 'replicability is the ability to replicate' should not be awarded credit.

Replicability is an important part of the scientific process. Scientific method involves defining a problem and formulating a hypothesis which is tested with empirical research. Research findings are an important part of this process. If we wish to draw conclusions from research studies, the procedures and findings should be repeatable. Unrepeatable results may imply flaws or lack of control within the method used and are of limited use in theory construction.

Award up to 4 marks for an explanation of why replicability is important. This is likely to be contextualised within a description of the scientific method.

AO1 mark bands	
4 marks	Explanation is sound
3 marks	Explanation is reasonable
2 marks	Explanation is basic
1 mark	Explanation is rudimentary
0 marks	No creditworthy material

55

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.

56

(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).

Examiner reports

1

(a) The majority of students were able to successfully identify 'independent measures' as the experimental design, although many confused 'design' with 'type of experiment', usually 'field', and failed to access any marks. 'Repeated measures' was also often stated though in these cases, students were at least able to access one mark if the advantage corresponded with the design they had given. Many students failed to make an appropriate link to the study described in their answer. However, some managed this by virtue of the fact that they went on to define 'independent measures' as 'participants only take part in one condition', thus the application in their answer was implied.

(b) This question was generally well answered with the Independent Variable (IV) and the Dependent Variable (DV) appropriately operationalized. Unfortunately, several students got these the wrong way round. Students should be reminded that it is best to state variables in an 'operational' form eg the DV could be given as 'number of people who picked up litter' but 'obedience' or 'level of obedience' would not gain credit.

(c) This question was generally very well answered with many students referring to the idea that a 'uniform' gives 'legitimacy to the authority figure / the orders', alongside an accurate likely outcome. However, some students did not state the likely outcome in terms of groups A or B in their answer. Other students offered only the outcome and so failed to access additional marks for 'explanation'. Some simply cited relevant evidence (usually Milgram or Bickman) without developing these into a coherent argument.

(d) The vast majority of students gained both marks for this question. Of those that did not, 'psychological harm' proved to be a less suitable choice of issue as it was difficult to link to the information available in the stem. Some students correctly identified '(lack of) consent' as an issue but then did little more than re-state the same phrase as part of their 'link' to the study / experiment described.

2

(a) Most answers scored two marks. Very few candidates failed to score at least one mark.

(b) Candidates generally either answered the question effectively, relating the multi-store model to the stem about Peter, or they failed completely to apply their knowledge as required. Concise, apposite answers scored full marks.

(c) There were many appropriate and clear answers but also some which identified one issue but then described how a different issue could be dealt with. Some candidates failed to clearly state the issue merely saying how an unnamed issue could be dealt with. Also, identifying 'withdrawal from the study' as the issue and then saying 'Peter should be told he could withdraw from the study' as a way of dealing with the issue, is not going to gain more than minimal credit.

A few candidates offered problems of methodology rather than ethics.

3

Some candidates tackled this question well and were able to offer sensible suggestions for protecting research participants from harm. Weaker answers were very brief, eg 'confidentiality', 'debrief them', etc without making it clear how this strategy would protect the participant from harm.

4 (a) Many candidates did not understand behavioural categories. Even those candidates who answered the rest of this question well struggled with part (a). Many candidates suggested broad behaviours, eg reunion behaviour, which could not be operationalised.

(b) The requirement to explain how the researcher might record the boy's behaviour was often ignored. Better answers described how behaviour categories could be recorded in a tally chart.

(c) Most candidates appeared to have some idea of why pilot studies are useful. However, a surprising number seemed to think that pilot studies are used to 'give you the idea for the main study'. Answers sometimes did not contain sufficient explanation to gain the second mark.

5 (a) This was a straightforward question and many candidates gained full marks. Some candidates inappropriately used their knowledge of cross-cultural research to write about different parenting styles in various countries instead of simply reading the table as the question required. It would be worth pointing out to candidates that when a question says, "Outline what the table shows", that is exactly what is required.

(b) Most candidates were able to identify a criticism although they often failed to explain the criticism for full marks. Some candidates ignored the requirement to 'explain one criticism' and, instead, identified two or three criticisms without elaborating any of them. Other common errors were explaining a criticism of the strange situation with no reference to cultural variations, or explaining a criticism of investigations of cultural variation with no reference to the strange situation.

6 Perhaps surprisingly, many candidates achieved higher marks for this section than for their other two questions. Most candidates attempted all parts of the question although a significant minority did not complete part (h) suggesting that they might have run out of time. This was a pity since this part carried 10 marks.

(a) This was a straightforward question and many candidates accessed full marks but a surprising number were confused by median and range and some did not understand what the range indicated about the data.

(b) Many candidates were very well prepared and got full marks here but some wrote very confused answers showing little understanding eg 'Spearman's Rho because it was nominal data and repeated measures'.

(c) There was a centre effect here. Some candidates had a good understanding of Type 1 error while others had clearly never heard of it. Some understood what is meant by the term and offered a definition but were not able to apply their knowledge to answer the question.

(d)   Candidates offered a wide range of answers, although some were a little bit too brief or poorly explained to get both marks. There was some confusion about what is meant by a placebo and some candidates offered two explanations which were essentially the same as one another. Other candidates offered factors which could apply equally to the treatment and non-treatment group. It is important in this kind of question to read the stem carefully. Some candidates said that the therapists might have been biased in favour of the treatment group, but the stem clearly states that the therapist did not know who had been in which group.

(e)   A lot of candidates missed the point that the advantage / disadvantage needed to be in comparison to interviews. Many candidates gave advantages/disadvantages that could apply equally well to both self-reports and interviews. This was acceptable only if the candidate made it clear ie '*People are less honest in a questionnaire*' was not credit-worthy because it could apply to both interviews and questionnaires. However, '*People are less honest in a questionnaire because they are anonymous and feel they can lie about themselves without being found out. In an interview where they are face-to-face with the interviewer, they might find it more difficult to lie*'.

(f)   Some candidates wrote excellent consent forms containing both ethical and procedural information and expressed them in appropriate language. Some candidates had a very vague understanding of what needed to be included here, either only focusing on all the ethical issues (*you will have the right to withdraw your data, yourself etc*) with no mention of the procedures, or vice versa. Many adopted a rather inappropriate tone eg '*Once you have signed this form, you are committed to being in the study*' or '*You have to subject yourself to an interview*'. It was surprising to see that a few candidates seemed to think a consent form acted as some kind of legal disclaimer – '*you may suffer harm but if you sign this you can't sue us.*'

It was notable on this question that candidates who were able to express themselves clearly and succinctly were much more likely to access full marks. Many answers were so poorly constructed that the content was difficult to understand. Many switched confusingly between pronouns eg *They will have to have an interview. You can withdraw at any time. I agree to be part of this study.*

(g)   There were some very muddled answers to this question. Candidates often didn't read the question carefully, and wrote something like '*Reliability means if you do the study again you will get similar results*' for their definition and then didn't know what to write for the next part of the question. Those candidates who explained it in terms of inter-rater reliability generally gained full marks. Some candidates did not read the question carefully and did not relate their answer to checking the scores in this particular study. Many candidates thought incorrectly that test-retest involved using different participants. Some candidates suggested split-half methods indicating a lack of thought about the question. Some candidates confused reliability with validity.

(h)   Many candidates showed limited awareness of a conventional reporting style. While it was not necessary to divide the method section into sub-sections, this strategy might have helped candidates to include all the relevant details. Weaker answers made no mention of gender or eating disorders and simply repeated details from the stimulus material. Many candidates completely lost sight of the fact that gender differences were being investigated and suggested randomly allocating participants to groups. A lot of time was wasted in including aims / hypotheses and statistical analyses which do not form part of a method section. Better answers included appropriate detail of IV, DV, design, sampling method, materials / equipment and procedure which would have enabled replication to take place. As in (f), poor expression and grammatical errors often obscured meaning.

7   There appeared to be a centre effect on this question. Some candidates were very well prepared and achieved high marks. However, there was a significant minority who achieved very few marks on the whole question. It was particularly surprising that part (c) presented such difficulty. A common mistake was to think that the design was a 'laboratory study', but there was a range of odd suggestions, eg 'random sample', 'controlled groups', 'peg-word method', etc. Candidates who did answer part (c) correctly usually gave a good answer to part (d), showing that they had genuine understanding of the design and its strengths. Perhaps, surprisingly, most candidates, even those who had not been able to answer the other parts of this question, did seem to have some understanding of the idea of replication in order to test reliability for part (e).

8   This question was usually well done by candidates who understood the term 'demand characteristics'. Some candidates clearly did not know what the term meant.

9   (a)   This question was well answered when candidates focused on two or three possible negative effects which could apply to Luca. Sometimes effects were just listed with no elaboration. In other cases, research into effects of institutionalisation was described with insufficient application to Luca.

(b)   Some candidates got confused with longitudinal and observational research and gave answers that did not apply to a case study at all.

10   (a)   Most candidates were able to identify this as a positive correlation. There were very few incorrect answers.

(b)   Candidates found it harder to identify a strength than a weakness. Good answers often explained the advantage when it would be unethical to manipulate variables. For the weakness most candidates focused on causality issues.

11

(a) Some candidates failed to recognise the term 'central tendency'. Others identified a correct measure but struggled to identify why the measure they had chosen should be used. A number of candidates explained how to calculate the measure or simply defined it, which could not receive credit. Candidates who had been well prepared for this question gave clear and concise answers such as "The median, because it is unaffected by outliers."

(b) Sampling method was generally well understood and many candidates were able to apply the method to teachers in the school. Candidates who had not read the stimulus material carefully enough suggested ways to obtain a volunteer sample which would not have been appropriate for this specific example. Others confused volunteer with opportunity or random sampling.

(c)(d) A number of candidates had not read the study closely enough and selected possible extraneous variables to do with the word lists. Good answers often cited individual differences or environmental variables in question (c) and were then able to explain how to control this in question (d), eg matched pairs or keeping noise, temperature constant / controlled.

12

(a) Students found this question very difficult. Many of the answers given were untestable statements because they did not contain two conditions and / or the DV was not operational. Answers were often in the following format – *There will be a better performance in the audience condition.*

(b) The majority of students were able to describe the pattern of the results but failed to use this as an explanation for the conclusion that could be drawn – the presence of an audience affected / increased the number of accurate shots through the hoop.

(c) Titles for the graph were generally good. The y axis sometimes did not contain reference to 'mean' number of successful shots. Plotting was accurate. Many students were able to present clear and accurate graphical displays.

(d) Students needed to be very clear about the strength and then explain why that was a strength of random sampling. Many students merely wrote a definition of random sampling.

(e) (i) Although students were able to identify a limitation of using independent groups – such as there are participant variables, rather than explaining why this is a limitation, they then went on to describe what is meant by participant variables.

(ii) Answers to this question often lacked information about overcoming the limitation. Students could suggest that using matched pairs or repeated measures might be appropriate, but could not expand on what a researcher would actually do.

(f) (i) This was not well answered. There were many very vague definitions of extraneous variables such as, something that cannot be controlled / something that has to be controlled.

(ii)   This was quite well answered although only a few students referred to 'confounding' in their answer.

(g)   This question required students to imagine themselves as the researcher about to read out to participants in the 'audience' condition the instructions for participation.
Some students produced very clear information in verbatim form. There were some students who focussed only on ethical issues.

13   Most candidates offered accurate and detailed answers for this question.

14   This question required candidates to draw conclusions from data, but many could not help going beyond the graph and interpreting the findings in terms of conformity research (Asch, NSI, ISI etc). The question did not require an explanation, but only what the chart showed. Many candidates did not gain many marks because they adopted the wrong approach.

15   (a)   Almost all answers were correct – however, surprisingly, some answers were left blank or the answer 'yes' was provided.

(b)   Many candidates clearly understood how to read the table and to interpret results and so gained the full 3 marks here. Some gained 1 mark for saying that the result was significant but then demonstrated a complete lack of understanding in the rest of their answer.

16   (a)   This question proved to be a good discriminator. Candidates who understood scattergrams were able to make a reasonable sketch with appropriate labels and accurately plotted data and so gained full marks. However, a disappointingly large number of candidates clearly had no understanding of scattergrams and drew a frequency polygon instead for which they could gain no marks. The requirement to present and understand graphs is clearly stated on the AS specification: 'presentation and interpretation of quantitative data including graphs, scattergrams and tables.'

(b)   Some candidates gave full answers in which they made good use of the data contained in the table. However, fewer candidates were able to make use of the information in the scattergram and very few referred to correlation. There were 4 marks available for this question which should have made candidates realise that some detail was required. Answers such as 'she was concerned because the observers gave different ratings' could not gain much credit. Quite a few candidates wasted time by defining inter-rater reliability. Answered included suggestions of how to improve reliability which, of course, was addressed in part (d).

(c)   Relatively few candidates identified an appropriate test - almost every reasonably familiar test was quoted. Experimental designs were often quoted as incorrect reasons for test selection. Many candidates did not even suggest an inferential test but suggested calculating the range, mean or standard deviation. Candidates who did identify the appropriate test were usually also able to offer an appropriate justification.

(d)   This was a good discriminator. Most candidates could offer at least one solution to this issue but many stopped after making their initial point eg 'give them more training'. Some were able to elaborate on this effectively to gain full marks but many showed little understanding. Very common errors were 'get more observers' or 'average the results' or 'only use one observer'.

17   Most candidates could provide an accurate outline of the weakness of using correlations. They were rarely able to give a strength, instead giving a definition.

18   Many candidates seemed not to have heard of this term, and attempts to guess were unsuccessful eg 'working with your friends'. A common misconception was that it was a marking exercise to give feedback during the research process. There were also many tautological answers such as 'getting a peer to review your work'. Many candidates appeared to have an idea of what peer review was, but were unable to articulate it in the way that would get them full marks.

19   There were some very good answers to this question but poor expression in some answers obscured meaning. A number of candidates simply described observational studies rather than offering an advantage.

20   **(a)**   Many candidates seemed unable to say what is meant by content analysis. In some cases, this was because of poor expression and the inability to define terms clearly. In others, it was simply that they did not know the term. Teachers and candidates must be aware that the Research Methods section of the PSYA4 specification builds on what was covered at AS. Anything that appears on the Research Methods specification at AS can be examined on PSYA4.

(b)   When candidates understood the term they were able to apply their knowledge effectively. For example, they explained how the psychologist would identify themes or categories in the drawings, count examples of each category to provide quantitative data and compare categories of drawings for changes over the duration of the study.

(c)   Many candidates wrote thorough consent forms using appropriate content and tone. But some just included procedural details with no mention of ethics or vice versa. Some had problems in including enough information to allow the participant to make an informed decision. Specifically there was often insufficient information on the stay in a research unit, the nature of the restricted diet and the need for testing. While some candidates referred to ethical issues, including right to withdraw, many did not. A few actually suggested that participants would be locked in if they agreed to take part.

21   **(a)**   Two symptoms of obsessive-compulsive disorder were accurately described by many – typically, and predictably, the 'obsessions' and the 'compulsions'; though physiological symptoms of anxiety were also deemed creditworthy. Some students gave symptoms that did not adequately distinguish OCD from other disorders, such as 'irrational thinking'.

(b)   Many students could identify the correct experimental design used in the study but fewer could provide an appropriate outline. A considerable number, however, thought the design was 'independent groups' or even 'matched pairs'. Finally, 'quasi-experiment' was an often seen answer.

(c) The advantage of 'repeated measures' was often stated rather than explained, for instance, 'no participant variables' was frequently offered without elaboration. Better, fuller answers tended to be those based on the time and cost-saving benefits of using the same participants twice in comparison to alternative designs. It was possible to gain two marks if the answer in part (c) could be matched to that in part (b), therefore, many students scored full marks in this question for an advantage of independent groups having named it above.

22

Many candidates were able to provide an appropriate advantage and a disadvantage of a questionnaire and to explain it. The main pitfall here was to describe a strength or weakness of a questionnaire without explaining it in comparison to an interview. For example 'people could answer dishonestly because of social desirability bias'. This could equally be true of an interview. Better answers included the explanation that it could be easier to lie on a questionnaire because the respondent is not face-to-face with the researcher. There were some very brief answers such as 'it's quick and easy'. Candidates who have tried to construct questionnaires might realise that while they may be quick to administer they are far from easy to construct.

23

Candidates who understood peer review were able to give a reasonable answer, but not many showed the elaboration needed for full marks. Many expanded on what it is rather than why it is important. A common error was that it enabled peers to replicate the research. While ethical considerations could have been of relevance, some candidates did not understand that peer review is a retrospective process and can only prevent ethical problems being repeated. A surprising minority talked about corrupt peers who would give a negative review to maintain their own interests.

24

The same issues on the *Anxiety Disorders* section were also relevant here. Many gave the wrong design for part (a) or the correct design without the outline. 'No participant variables' was often quoted for part (b) without elaboration. Part (c) posed few problems; the therapy was judged to be 'effective' by most and justified with reference to the data in the table.

25

(a) Most candidates understood that this was a volunteer / self-selected sample and many of these went on to explain a relevant weakness. The most commonly offered was the problem of generalising from a self-selected sample. Some candidates lost a mark by confusing a volunteer sample with an opportunity or random sample. However, they could still gain marks by explaining an appropriate weakness. A surprising number of candidates did not understand what was meant by a sample and identified the method (questionnaire) instead. Unfortunately, this meant that they usually got no marks at all.

(b) This was generally answered well with the most successful answers focusing on issues of confidentiality and protection of harm. There were some very thoughtful responses and it was pleasing to see that some candidates could respond by giving serious consideration to ethical issues. Informed consent was a perfectly acceptable issue but candidates who used this often offered a weak way of dealing with it eg 'informed consent was an issue so the researchers would have had to ask them to give their consent'. A surprising number of candidates did not seem to understand what was meant by an ethical issue and wrote about problems of sampling instead.

26 This question was generally answered well. Most candidates focused on the higher ecological validity provided by real-world studies. Weaker answers simply made this point without any elaboration. Better answers explained why a real-world setting would provide more ecological validity – usually in terms of heightened anxiety and / or consequentiality associated with real-life events. The best answers made reference to EWT studies rather than simply stating the general advantages of real-life studies over laboratory studies.

27 In general candidates were able to do well on this question.

28 (a) A challenging question because candidates needed to apply their knowledge. They often knew about what makes something scientific (objectivity, replicability, etc) but seemed unable to engage with the stem. There were lots of answers involving paradigm shift, which were not relevant to this question.

(b) Most candidates had some idea about how a random sample could be obtained, but often failed to explain the methods fully. They could suggest all the names should be put in a hat, but did not make it clear that the names were then selected "without looking" or "without bias". There was some confusion with systematic sampling.

(c) Many answers displayed some confusion here, eg saying that a limitation was that it was not representative of the whole population, when the point is that is might not be representative of the target population of 400. Some answers referred to problems of allocation to conditions, rather than random sampling. A good point was made by those who said that if some parents did not give consent, the psychologist would have to select again, and that would not be random.

(d) This question was not answered well. Most candidates seemed very unclear about why it is important to operationalise variables. How to actually operationalise the two variables was beyond many candidates. Some effective answers referred to food content eg fat, sugar etc.

(e) Most candidates answered this correctly.

(f) There was some serious confusion about what exactly matched pairs design is. Few could go beyond "it's time consuming" or "difficult to match on all variables". Some referred back to the random sample and said it would not be possible; others felt that at five-years-old children are either too similar to match or too different.

(g) Most could identify an ethical issue such as confidentiality, the right to withdraw and protection from harm (those who did not get any breakfast or who were embarrassed at their poor reading). Some seemed to forget that they also had to explain how the issue would be dealt with, or they simply repeated that the right to with draw could be dealt with by giving the right to withdraw.

(h) This question was not answered well. Many candidates failed to read the question carefully before they attempted it. They were given the information that they were using the same group of children (ie the 5-year olds in the previous study). Despite the fact that the ethical issues and sampling had already been addressed in the plan for the original study many wrote at great length about sampling and ethics. The majority of candidates were unable to write a fully operationalised hypothesis, and often simply restated the aim. Many seemed to think the IV was breakfast versus no breakfast, rather than healthy versus unhealthy breakfast. Some of their ideas were totally impractical, especially given that the children were only 5 years old. In many answers lack of detail would have made any kind of replication very difficult. However, some candidates did understand the need for some sort of training for the observers, the need for clearly identified behaviour categories to record, and the importance of being able to distinguish the two groups in the playground. Designing a study is clearly a difficult task for candidates, and one that they need to practice.

29 This question asked candidates to explain the strengths, not merely describe a characteristic of a questionnaire. There were a wide range of answers, such as if the questionnaire is anonymous, respondents may be more honest; if they are filled in private demand characteristics are reduced. However, answers that simply said that they are 'cheap' or 'quick', gained no credit unless there was some elaboration. For example, questionnaires are quicker than interviews. Describing the type of data as a strength was in itself not credit-worthy, since such data could often be generated by other methods. However, if reference was made to the type of question and the data it would generate, this could be credit-worthy. For example, open-ended questions produce qualitative data that can be rich and varied.

Many candidates misread the question and provided one strength and one limitation.

30 Drawing conclusions from data seems to be a topic that candidates find hard. Perhaps teachers might emphasise the difference between a conclusion and a finding (although findings can be used to support a conclusion). Also they need to highlight to candidates that these types of questions are straightforward and are not trying to trick them. Often very able candidates seemed to think that they had to go beyond the data and make links to theory etc.

31 (a) Many candidates showed some understanding of a natural experiment but a substantial number confused it with studies in a natural environment.

(b) Most answers were credit-worthy although some merely named rather than outlined a method. The best responses suggested a way which involved both observation and recording of the behaviour. Few candidates suggested the use of interviews or questionnaires, although these were credit-worthy.

(c) This question was usually well answered and there were very thoughtful responses, for example, where candidates suggested that consent from the children as well as informed consent from parents would be appropriate, given the age of the children.
It is encouraging to see that candidates are able to give consideration to ethical issues. However, worryingly, a few did not seem to be aware of what ethical issues are.

32 (a) Most candidates accurately reported independent groups / measures, unrelated or between subjects / groups design. Candidates seemed better prepared to identify the experimental design than in previous examination series, although there were some who incorrectly identified the type of experiment (eg laboratory experiment).

(b) If candidates got (a) wrong (laboratory experiment) then their limitation was usually wrong, so there were some incorrect references to low ecological validity. When candidates selected a correct limitation eg participant variables they did not always elaborate it well enough for two marks. Examples of individual differences were most plausible when they focussed on an appropriate characteristic for this study eg memory differences.

(c) Candidates who looked at the results on the graph and commented appropriately on the pattern of correct and incorrect statements, scored full marks. It is advisable for candidates to read the labels on the axes of a graph. A number of candidates failed to do this and so reported inaccurately on the average number of *participants* who had made statements.

(d) Answers generally showed good understanding of the cognitive interview approach. Candidates who just named a process, eg *report everything*, did not score full marks. Some elaboration or application to the stem was necessary to obtain a second mark.

(e) Responses to the term *investigator effects* were varied. A few candidates wrote excellent responses identifying cues the investigator might produce in the context of the study, which might lead to the investigator's expectations being fulfilled. Quite a number of candidates inappropriately discussed leading questions and often cited Loftus's work.

33 This was answered correctly by most candidates.

34 It was encouraging to see that this cohort of candidates was able to go beyond simply describing the findings. They were able to make suggestions about what they showed. For example, that both therapies showed some improvement, as there were no scores of zero; that in fact neither showed much improvement as the average was only 6.

However, it was also clear from the responses that a minority of candidates had no real understanding of what range tells us about data.

35 (a) Hypothesis writing is still a problematic area for many candidates – despite the requirement to do this at AS level. Many candidates achieved zero marks on part (a), having mistakenly written a directional or a null hypothesis. Many responses were lacking in clarity or failed to include an operationalised DV so only achieved 1 mark. The best answers were concisely and clearly worded responses such as "There will be an association between birth order and career choice", which achieved the full 2 marks.

(b) Virtually all candidates identified an appropriate sampling technique. However, a large number did not score full marks because their account of how to obtain the sample was confused or insufficiently linked to the study in question on artists and lawyers. Candidates who chose a random sample needed to explain how the target population would be identified.

(c) Some centres had clearly prepared their candidates very well and many showed an impressive understanding of inferential statistics scoring 11 or 12 marks. However, other candidates struggled with the question and collected very few marks. Some of the most common errors were as follows.

A number of candidates did not know how to express the statistical conclusion of a research study, by referring to observed and critical values and probability. There were errors in correctly identifying the observed and critical values and their relationship to the hypothesis. A large number of candidates did not label the axes of the graph or only showed data relating to first born career choices.

Some candidates chose the wrong statistical test; some did choose the correct statistical test but did not then state the reasons why the test was appropriate.

Yet again, advice to teachers is to do some practical work. It was clear that some candidates were very familiar with the rationale for selecting a test and deciding if an observed value is significant or not. These candidates had a strong advantage on part (c).

36

(a) This question was answered well. Most candidates provided a detailed aim that was awarded 2 marks. A minority of candidates provided a more general aim that was credited with just 1 mark, such as "to investigate the effect of colour on performance in cognitive tasks". For both marks some elaboration (either related to colour or performance) was required.

(b) This question was problematic for a lot of candidates. Many candidates confused the introduction with the abstract or the method sections and received no credit. Some recognised the inclusion of aims / hypotheses in the introduction but did not achieve 2 marks because they did not make reference to background information. A few impressive answers showed real understanding and referred to "contextualising the research".

(c) Although this question was worth only 1 mark, many candidates produced lengthy answers. Some distinguished between specific types of validity such as external validity or population validity. A small number of candidates became confused between validity and reliability and provided a definition of the test-re-test method. Just over half of candidates gained the mark.

(d) There were some lovely, imaginative responses to this question which was answered well in general. The majority of candidates achieved 2 marks by including an example of how colour could be used in a real world setting. The most popular answer was use of colour in classroom walls or on textbook pages to aid learning in particular subjects.

(e) The majority of answers to this question demonstrated an understanding that independent judges were required to reduce bias, and in doing so the majority achieved two marks.

(f) Few students achieved full marks on this question, providing little additional information to that included in the question stem. Candidates were too focused on providing details of ethics (which was not required), at the expense of standardised instructions. Some candidates also made an error in their instructions by stating that participants would be given 40 shapes, 20 red and 20 blue, when in fact participants would only be given one colour of 20 shapes. A further common error was writing that participants would be given a limited time to make the toy. Writing a limited time is not a clear standardised instruction and stronger candidates wrote exactly what the time limit would be. Very few candidates checked if participants had any questions at the end of the instructions. Candidates who had conducted research were at an advantage here and produced answers of a higher quality.

37

(a) There were many appropriate suggestions for maintaining confidentiality, such as using pseudonyms, avoiding using photographs and avoiding identifying the location of the case study participant and / or family. A few candidates did not seem to understand the term confidentiality as some answers were based on other ethical issues eg protection of participants from harm.

(b) Most techniques identified were appropriate and there was some thoughtful elaboration related to the case study. Answers referring to the use of meta-analysis were not appropriate.

(c) Many candidates scored at least two marks, usually by focusing on the uniqueness of a case study and the limited possibility for generalisation. A few candidates could elaborate on this whilst others considered a second factor such as the difficulty of verifying evidence from the past. Answers focusing on limitations due to the time involved in longitudinal research were not credited as many case studies are not longitudinal.

38

(b) Candidates were required to identify and elaborate an appropriate methodology for measuring aggression. Many were able to do this.

Unethical answers such as 'provoke the children and observe their response', did not gain credit.

(c) By identifying the lack of causal inference in correlation and suggesting other relevant factors which may affect children's aggression, some candidates obtained full marks. Another appropriate route was to provide evidence from research to create a counterclaim. Trying to answer by using the scattergram was not a successful strategy as the question asks about causal inference and graph shows only the relationship between factors.

39   (a)   Most students correctly identified 'laboratory experiment'.

    (b)   Many students gained a mark by outlining the increased 'control' that is offered by laboratory experiments. Some failed to expand or substantiate the advantage cited, by failing to acknowledge that this would lead to 'more reliable cause and effect relationships', for instance.

    (c)   This question was generally well answered with the Independent Variable (IV) and the Dependent Variable (DV) appropriately operationalized. Unfortunately, several students got these the wrong way round. Students should be reminded that it is best to state variables in an 'operational' form eg the DV could be given as 'number of people who picked up litter' but 'obedience' or 'level of obedience' would not gain credit.

    (d)   This question was almost always correctly answered.

40

Students seem to struggle with the interpretation of graphs and many do not seem to understand the difference between what data show ie findings and what can be concluded from the data. The graph shows a negative correlation, which is moderate-strong; this shows that the more weeks of treatment the lower the depression score. Students could also comment on the apparent "plateau" or change in direction in the middle of treatment. Very few students considered the strength of the relationship and weaker students just read the data off the graph.

41

As with previous high mark research method questions, this question had a range of answers from students that covered marks from 0-10. Some schools and colleges had clearly prepared their students very well and many showed an impressive understanding of experimental design and controls. However, other students struggled with the question and gained very few marks. Some of the most common errors were as follows:

- Ignoring the requirement to use repeated measures and converting the experiment to an independent groups design

- Failing to counterbalance order of presentation of the two types of music

- Producing two concentration tests which were not matched for difficulty

- Testing music v no music

- Focussing on trivial controls (breakfast, temperature) and ignoring important ones (volume of music).

Yet again, advice to teachers is: do some practical work and encourage your students to plan 'thought experiments'. It was clear that some students were very familiar with designing experiments and they had a strong advantage on this question.

42 (a) Many candidates produced an operationalised directional hypothesis. Sometimes the dependent variable in the hypothesis was not operationalised, resulting in less than full marks.

(b) Most responses were appropriate with candidates showing a good understanding of independent groups design.

(c) Strength: many responses focused appropriately on the absence of different types of order effects or the limitation of demand characteristics relative to a repeated measures design. Answers which elaborated a relevant strength gained full marks.

Limitation: Good answers often referred to examples of individual differences between the groups affecting reliability. Appropriate elaboration was often evident.

Lack of clarity of expression sometimes obscured the point being made.

(d) A substantial number of candidates were able to explain the purpose of a pilot study, usually focusing on usefulness for checking and amending aspects of the procedure before the main study. The opportunity to observe and address unexpected ethical issues was also a creditworthy point. Candidates were particularly successful when they used an example to illustrate their answer for example to check the clarity of the pictures. Answers which merely outlined what a pilot study is were not credited.

(e) Responses to this question divided clearly between those who understood and could apply the concept of standard deviation and those who did not understand the concept and hence could not interpret the numbers in the table. A common wrong answer was to claim that the memory improvement group did 'better', presumably because 2.8 is higher than 0.29.

43 (a) In this question, students were required to discuss the advantages of carrying out the experiment described in the stem, in a laboratory. Fewer than half of students made any reference to the stem and the most common mark awarded was one out of four. Those who referred to an advantage (eg control of extraneous variables) and linked it appropriately to the scenario (eg posters on the walls) were able to access the full range of marks. A small but significant minority insisted on writing about disadvantages and achieved no marks. Once again, schools and colleges should advise students to read stems carefully and apply knowledge in Section C.

(b) Most students achieved full marks, identifying the Mann-Whitney as the appropriate test and giving and ordinal data or independent groups as a reason. Some students provided two or three reasons going beyond the requirements of the question. There were a minority of cases where an incorrect answer was given, most commonly Spearman's rho or Wilcoxon's signed ranks test.

44 (a) This question was answered well with most students aware that a directional hypothesis was appropriate due to the existence of previous research. A minority of students provided rather more detail than required for one mark.

(b) Hypothesis writing is still a problematic area for many students, despite the requirement to do this at AS level. Many students achieved zero marks on this question, having mistakenly written a non-directional hypothesis or one which predicted a difference between older and younger patients. Many responses were lacking in clarity or failed to operationalise recall adequately. The best answers were concisely and clearly worded such as "There is a negative correlation (relationship) between age and recall accuracy rating", which achieved the full three marks.

(c) Although this question was worth only one mark, many students produced lengthy answers. Some distinguished between specific types of reliability such as external or internal. A small number of students became confused between validity and reliability.

(d) There was a broad range of answers to this question, with students in roughly equal measure being awarded marks across the full range. The majority had at least a rough idea of ways of assessing reliability (the most common being inter-rater) but found it difficult to select an appropriate method for the study detailed. The weakest answers were those where the student focussed on reliability of the study overall, rather than reliability of the ratings which was what the question required. Answers that achieved the full three marks generally selected the most straightforward idea; to take two independent psychologists who rated the typed accounts separately and then correlated their ratings. Students who achieved only one mark suggested test retest as a method but most were unable to carry this through and indicate that the psychologist would need to return to the data after a suitable interval and re-rate the accounts.

(e) This question was answered well with the majority of students achieving two marks. There was a range of both kinds of data to draw on here including the doctor's notes and the patients responses (qualitative data) and ages and accuracy scores (quantitative data).

(f) Answers to this question demonstrated an understanding of the use of the Spearman's rho statistical test with the majority of students achieving two marks.

(g) This question confused many students who were unaware that the critical value relates to the magnitude of rho not the direction. So negative correlation drops the minus sign when compared with the critical values. About half of students were clearly aware of this and could compare the obtained value with the correct figure from the table. The remainder made a number of errors, some comparing -.52 with 0.05, others claiming that the figure was smaller than .306. Some incorrectly used the values relating to a non-directional hypothesis.

(h) Full marks were achieved by stating that the null is rejected and the experimental hypothesis accepted, when in fact results are due to chance. Good understanding was shown among students who referred to the level of significance being set too leniently or the 5% likelihood of a Type 1 error occurring with the 0.05 level of significance. In about one in three cases students confused Type 1 and Type 2 errors.

(i)   This question was challenging for students with many achieving no credit or not answering the question. Even students who were able to explain what was meant by a Type 1 error were unable to apply this knowledge in this question and compare the obtained value with the 1% significance level. A small number gained one mark for identifying that the obtained value was substantially larger than .306. However, far too many relied on a rote learned response that the 5% significance level avoids Type 1 errors therefore one could not have occurred.

45

(a)   Mostly appropriate answers were provided. Many answers referred to the term 'youth' misleading participants as it suggested the man in the photograph was young and this could influence the answer. Others said the reference to 'youth' was misleading because it suggested an extra person, a 'youth', was part of the photograph. Some answers suffered from poor expression resulting in a muddled answer.

(b)   There were still a number of students who did not seem to understand the term 'experimental design' and who tried to answer by stating types of experiments such as laboratory or field. Even amongst students who understood the concept, there was some difficulty when it came to identifying the appropriate design. The explanation of why a repeated measures design was unsuitable was in some instances prefaced by unnecessary time wasting explanations of why an independent groups design would be a good idea.

(c)   Although most students could explain the benefits of a pilot study, far fewer students gave an explanation in the context of this experiment, as required.

(d)   There were some clear and effective answers to this question where students showed understanding and could apply this to the scenario. Some students however, failed to understand the question and perhaps would have benefited from re-reading the stem.

(e)   Although there were some 'accurate and reasonably detailed answers' there were many more that were just 'generally accurate'. As in previous exam series, when asked to describe a research study, some students did not even know one study sufficiently well to access the top mark band. Some students had a little muddled knowledge of several (usually Loftus' studies) and produced answers where the research was so poorly described it was difficult to identify. There were, however, some good answers which accurately described one or more studies. Some students wasted time evaluating the research.

46

(a)   Attention was drawn to the requirement for the variables to be operationalised in this answer. Some responses met this requirement effectively, other responses were more vague. Inevitably there were some who mixed up the IV and DV.

(b)   Although almost all students recognised that the mean score for aggression was higher in those who started day care before the age of two than after the age of two, fewer students pointed to the magnitude of the difference being small.

(c)   There was a wide range of incorrect answers to this question. Clearly a number of students did not recognise the term "dispersion".

(d)   Most students were able to draw a bar chart. Those who drew separated bars to represent those who started day care before or after the age of two created a better visual impact than those who chose to join the two bars. Students who did not score full marks usually failed to label the axes fully eg labelled the 'y' axis as mean score, rather than mean aggression score.

(e)   In a majority of responses, hypotheses were both appropriate and directional. Some students failed to operationalise part of the hypothesis and so did not score full marks. A few responses were written in the form of a correlational hypothesis, which was not appropriate.

47

Almost all students gained at least one mark. Those who gained full credit did so by appropriately referring to the nomothetic / scientific nature of the behaviourist approach and disadvantages of qualitative data. Some students gave good and relevant justifying commentary for comparisons with quantitative data collection. Students who only achieved one mark, generally did so because they gave two vague or very brief reasons or a generic evaluation of qualitative data not linked to investigations carried out by behaviourists. A minority of students gave one reason only.

48   (a)   Students often struggled with this question. Very few understood why measures of dispersion are used in addition to measures of central tendency and a number used the term 'dispersed'. Applications to the stem lacked the necessary detail to attract a mark.

       (b)   Many students were able to make correct use of the table and draw an appropriate conclusion about the statistical significance of the T value. Stronger students were able to present this information well, with some even correctly stating that the results were not significant at the 0.02 level and explaining why. A few less successful students were confused about the critical and calculated values of T.

       (c)   It was heartening to see so many students being able to explain why this test was used, with many students scoring the full 3 marks. A few students, however, stated that the data was nominal or interval.

       (d)   Most students were able to score at least one mark on this question even if they scored poorly on preceding and subsequent questions. Students had to think carefully about this answer and many were able to suggest one or two sensible reasons for the use of a diary by each offender. Elaboration and explanation of each reason proved more of a challenge, sometimes resulting in overlap and repetition across the two reasons offered. Some students simply repeated the stem.

       (e)   This question produced some answers that showed a lack of understanding of the use of a control group with quite a few students suggesting that a control group would consist of 'a group of non-offenders' or 'normal people / people with no anger issues from a normal population'. More informed students were able to explain why a group of people who would not have the anger management programme would have improved the study although some simply said that 'it would make it more scientific' without explaining why.

       (f)   Answers to this question were most disappointing with almost two fifths of students failing to score a single mark. What should have been a straightforward question proved challenging for many students who seemed unfamiliar with how to establish the reliability and validity of the questionnaire. Some students confused reliability with validity, others simply provided definitions of each and, where an attempt was made to apply reliability and validity to the stem, students very often referred to pilot studies and peer review. Reliability was sometimes mistaken for replicability, the 'split-half method' was frequently explained as dividing the total score in half to see if the two halves correlated and test- retest was sometimes explained as testing one group of offenders and retesting another group of offenders. Face validity was rarely applied to the questionnaire ie anger scores. Even students who correctly addressed the issues of reliability and validity, failed to explain how statistical tests of correlation would be used in this context. The value of carrying out practical activities to enable students to 'think like a psychologist' and apply their knowledge of practical activities in answer to questions such as this one cannot be overstated. It is clear that where students had been presented with such opportunities, they were able to write confidently and in an informed manner.

49

Students seemed to have a reasonable understanding of replication in relation to finding similar results, although only a few mentioned repeating the study in the same way or similar way, such as repeating the same method or procedure. Students who referred to 'replicating' the method or procedure did not gain credit. On the whole, students fared better with the second part of the question which required application to the scientific approach. However, some students failed to grasp that replication is a condition for validity but does not on its own confirm validity.

50 This was a straightforward question for which many students gained at least one of the two marks for referring to numerical data. However, not all students provided a suitably clear example of quantitative data collected by a psychologist who studied conditioning. Those who managed to gain full marks often did so with, for example, reference to the work of Pavlov and Skinner, by referring to the amount of saliva or time taken to respond.

51 Interestingly, students seemed to find it easier to outline a weakness of case studies rather than explain what they are. There was a wide range of answers to (a) and better answers referred to a study of a single person, usually over a period of time, using a range of methods to collect data. Some students also illustrated their answers with relevant examples, such as Freud's Anna O or Little Hans.

Most students were able to offer lack of population validity, since a case study was of a unique person; or the issue of reliability since they could not be replicated.

52 (a) Hypothesis writing continues to be problematic for many students, despite the requirement to do this at AS level. Around 40% of students achieved zero marks on this question , having mistakenly written a directional hypothesis or one which predicted a difference between mathematical ability and musical ability as opposed to a relationship. Many responses lacked 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 test of mathematical ability and their scores on a test of musical ability", which achieved the full 3 marks.

(b) This question was answered well, with most students scoring two or all three marks. Weaker students were able to spot the test was based on a subjective judgement and some also made the point that singing was a poor measure of all round musical ability. Stronger students identified the lack of control (different choices of song) and were able to link this appropriately to investigator bias. Some students also made the point that the test lacked validity as it had not been standardised.

(c) There was a broad range of answers to this question and about 40% of students achieved no marks at all. Some confused reliability with validity, suggesting various methods such as comparing the scores with another measure of maths ability. Few contextualised this by identifying alternate forms which would have been creditworthy. Others made reference to running a pilot study which received no marks.

The remaining 60% had some idea of ways of assessing reliability of the maths test, the most common methods being test-retest and split-half. Some used inter-rater reliability appropriately suggesting that two separate markers could be used for the maths test: others became sidetracked into assuming that the study was observational. Stronger students were able to explain two or three methods of checking reliability in reasonable detail.

(d)   This straightforward question on a random sample caught out quite a few students. Most were able to achieve 1 mark by referring to the method as being likely to yield a more representative sample. The weakest students simply defined random sample and went no further.

(e)   This question required students to draw a scatter graph to display the data. About half achieved all three marks here. Many students failed to gain full marks by inaccurate or missing labels or title. About one third of students drew an incorrect graph, the most common error being to draw a bar chart.

(f)   Most students were able to make some commentary on the generally negative correlation shown in the graph and table. Better students noted the presence of two outliers which weakened the overall strength of the relationship and some commented on the impact of outliers in a small sample. A small number of students made a rough calculation of Rs which was impressive but unnecessary to gain full marks.

(g)   This question had a range of answers from students that covered marks from 0-10. The mark scheme allowed students to argue for different ways of designing the experiment (independent measures or matched pairs) and of generating a sample (volunteer or random selection from the two groups) provided these were workable and justified. Some common errors included:

•   suggesting an inappropriate design (repeated measures) which did not take account of the information relating to left and right handers

•   suggesting a sampling method but not explaining how it would yield an appropriate sample of left and right handers

•   assuming that a maths test also needed to be completed (ie incorrect IV)

•   failing to provide any procedural information

•   producing a debrief which was not suitable to be read out to participants

•   providing standardised instructions and claiming they were a debrief.

Some schools and colleges had clearly prepared their students well and many showed an impressive understanding of experimental design. Others struggled with the question and/or, failed to read the instructions and therefore gained very few marks. Once again, advice to teachers is: to do practical work. It was clear that some students were very familiar with designing experiments and they had a strong advantage here.

(h) This question required students to follow through their design from (g) and give some indication of how the data would be recorded and analysed. Most managed to sketch an appropriate table to record data, although a few misread the question and produced a summary table. Some students were able to follow through their design / data type with an appropriate test which could have been Mann Whitney (independent design) or Wilxocon (matched pairs). Students who had collected nominal data or recorded data in nominal form were credited if they suggested chi square to analyse it.

53 (a) The majority of students scored full marks for this question although weaker responses merely re-stated the pattern of the results given in Table 1 and omitted any reference to an appropriate conclusion.

(b) Answers to this question were relatively poor, with only a third of the cohort able to state the independent and dependent variables with clarity.

(c) This was very poorly answered with more than half of the responses achieving no credit. Students muddled stratified sampling with both random sampling and systematic sampling. Very few could explain in a logical way exactly what a researcher would do to gain a stratified sample of the available target population.

(d) (i) Although some students seemed aware that once selected the participants could be randomly allocated to conditions using a hat or computer, they were very unclear about exactly what should be done. Descriptions suggested 'putting all the names in a hat/computer' but could not describe what would happen next. It seemed quite obvious to examiners that many students have never had a practical opportunity to allocate either people to conditions or words to lists, in a random way.

(ii) Similarly, few students gained full credit for this question as they struggled to explain why random allocation is important, although many were able to produce answers with vague reference to the issue of bias.

(e) Students have clearly learned generic responses to questions about the advantages and disadvantages of experimental designs. However, here they needed to make it clear that repeated measures would not be a suitable experimental design in this case. Some did not notice that having guessed the weight of a cake once, participants would lose their naivety in the study and it would be an illogical request to ask them to repeat the task either with, or without, other estimates on the answer sheet. In many cases they just trotted out answers which suggested the participants would be extremely fatigued having guessed once or that they would get better at guessing so their answers would improve.

(f)   There seemed to be an expectation that as the question asked about ethical issues then the response should include how to address the issue. The question did not require students to do this but many offered such answers. Some stated that informed consent was an issue but then explained the issue of consent/permission to participate in the study, without realising these are different issues. Others identified informed consent but failed to recognise that attempting to gain this would mean it would be impossible to implement the study.

(g)   (i)   Although the Specification only names structured and unstructured interviews as types of interview, there were some very obscure answers to this question. Unfortunately, even when students chose to write about the named interviews they often drifted into answers about type of question. Some thought that structured interviews only contain closed questions and unstructured only open questions and elaborated their description in this way. They then compounded this error in part (g)(ii) by discussing a limitation of closed or open questions. Others suggested that one type of interview would be a questionnaire.

        (ii)   As stated above, many students discussed limitations of types of question rather than types of interview and only about a third of the cohort gained full credit for this question.

54

This question was problematic for the majority of students with an average mark of 2 / 5 and only about 15% achieving 4 or 5 marks. Whilst most were able to provide some definition of replicability, few were able to explain in any detail why replication is an important part of the research process. Weaker students asserted that replication means that a study is reliable and / or valid. Stronger students contextualised replication in a discussion of the scientific method and referred to the importance of repeating studies to check for methodological flaws or investigator biases and some considered the importance of replication in supporting or refuting theories. Those who were able to provide a clear overview of the scientific process fared best.

55

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.

56 (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.