Hi folks,
Saj Devshi here from www.loopa.co.uk - The best place for AQA Psychology revision.
I have adapted an old booklet I found and updated it with the latest research method questions. I found this immensely useful in helping me achieve an A* grade as what I did was I would go over these questions while cross-comparing against the mark scheme. Over time it began to click and stick in my mind as I kept practicing it.
I have kept the question numbers from their respective papers to help you cross compare against the mark scheme but I do have plans to release a separate booklet that goes with this so it has all the answers in one book and its easier for you to cross compare against (rather than using multiple mark schemes and it becoming a big faff)
If you find this book helpful please be sure to like my Facebook page here: https://www.facebook.com/aqapsychology

This helps spread the word about my blog to others so we can help other students studying AQA Psychology.

All the best, - Saj

All questions from June 2010 to June 2013

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

18) What is content analysis? (2 marks)

19) Explain how the psychologists might have carried out content analysis to analyse the film clips of driver behaviour (4 marks)

20) Explain how the two psychologists might have assessed the reliability of their content analysis. (3 marks)
The Psychologists then designed an experiment to test the effects of using a hands-free phone on driver’s attention. They recruited a sample of 30 experienced police drivers and asked them to take part in two computer-simulated driving tests. Both tests involved watching a three-minute film of a road. Participants were instructed to click the mouse as quickly as possible, when a potential hazard (such as a car pulling out ahead) was spotted.

Each participant completed two computer-simulated driving tests:

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

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

**21) Explain why the psychologists chose to use a repeated measures design in this experiment.** (3 marks)

**22) Identify one possible extraneous variable in this experiment. Explain how this variable may have influenced the results of this experiment.** (3 marks)

**23) Explain how one factor in this experiment might affect its external validity** (3 marks)

**24) Explain one or more ethical issues that the psychologists should have considered in this experiment** (4 marks)

**25) Write a set of standardised instructions that would be suitable to read out to participants before they carry out Test A, chatting on a hands-free phone.** (5 marks)

The computer simulator measured two aspects of driver behaviour:

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

The mean scores for each of these measures is shown in Table 1.
The Psychologists then used an inferential statistical test to assess whether there was a difference in the two conditions.

**26)** Identify an appropriate statistical test to analyse the difference in the number of hazards detected in the two conditions in this experiment. Explain why this test of difference would be appropriate. (3 marks)

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

**27) **Explain why the psychologists did not think that they had made a Type 1 error in relation to the difference in reaction times. (2 marks)

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

Explain why replication of this study would be useful (3 marks)

---

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

17) Write an operationalized non-directional hypothesis for this study. 

(2 marks)

18) Identify an alternative method which could have been used to collect data about happiness in this study. Explain why this method might be better than using a questionnaire. (4 marks)

19) What is meant by internal validity? (1 mark)

20) Describe how the internal validity of the happiness questionnaire could be assessed (3 marks)

21) A Spearman’s rho test was used to analyze the data. Give two reasons why this test was used (2 marks)

<table>
<thead>
<tr>
<th>N (number of participants)</th>
<th>Level of significance for a two-tailed test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Level of significance for a one-tailed test</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>0.025</td>
</tr>
</tbody>
</table>

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

22) The psychologist used a non-directional hypothesis. Using Table 1, state whether or not the correlation between intelligence and happiness at age 11 (+0.42) was significant. Explain your answer. (3 marks)

23) Five years later, the same young people were asked to complete the intelligence test and the happiness questionnaire for a second time. This time the correlation was −0.29.
With reference to both correlation scores, outline what these findings seem to show about the link between intelligence and happiness.

(4 marks)

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

24) What is meant by peer review? (2 marks)

25) Explain why peer review is an important aspect of the scientific process. (4 marks)

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

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

Explain one or more limitations of the student’s proposal and suggest how the investigation could be improved. (10 marks)

JUNE 2012 Psya4 Research Method Questions www.loopa.co.uk
Do Not Copy Or Reproduce Without Consent ©

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

A math’s 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.

17) Suggest a suitable non-directional hypothesis for this study. (3 marks)

18) Why might the measure of musical ability used by the teacher lack validity? (3 marks)

19) Explain how the teacher could have checked the reliability of the mathematical ability test. (3 marks)

20) Explain why the teacher chose to use a random sample in this study. (2 marks)

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

Table 1: Mathematical ability test scores and musical ability ratings for 10 students

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematical ability test score</th>
<th>Musical ability rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>

21) In your answer book, sketch a graph to show the data in Table 1. Give the graph an appropriate title and label the axes. (3 marks)

22) Discuss what the data in Table 1 and the graph that you have sketched seem to show about the relationship between mathematical ability and musical ability. (3 marks)

23) 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 marks)

24) In your answer book, draw a table to show how you would record your results. Identify an appropriate statistical test to analyze the data that you would collect. Justify your choice. (3 marks)

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.

16) The psychologists decided to propose a directional hypothesis. Why was a directional hypothesis appropriate in this case? (1 mark)

17) Write a suitable directional hypothesis for this investigation. (3 marks)
18) The psychologists were careful to consider the issue of reliability during the study. What is meant by reliability? (1 mark)

19) Explain how the psychologists might have assessed the reliability of their ratings. (3 marks)

20) This study collected both qualitative and quantitative data. From the description of the study above, identify the qualitative data and the quantitative data. (2 marks)

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.

21) Give two reasons why the psychologists used Spearman’s rho to analyse the data. (2 marks)

22) Using Table 1 below, state whether the result is significant or not significant and explain why. (2 marks)

Table 1: Extract from a table of critical values of Spearman’s rho ($r_s$)

<table>
<thead>
<tr>
<th>Level of significance for a one-tailed test</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
</tr>
<tr>
<td>Level of significance for a two-tailed test</td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>0.10</td>
</tr>
<tr>
<td>$N=29$</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>31</td>
</tr>
</tbody>
</table>

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

23) Explain what is meant by a Type 1 error. (2 marks)

24) Use the information in Table 1 above to explain why the psychologists did not think that they had made a Type 1 error in this case. (3 marks)

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.

25) 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 marks)

26) 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 marks)

27) 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 (e.g. word search).

You must use a repeated measures design.

In your answer you should:

1. fully operationalise the independent and dependent variables
2. provide details of how you would control extraneous variables
3. describe the procedure that you would use. You should provide sufficient detail for the study to be carried out. (10 marks)

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.

17) What were the researchers’ aims in this study? (2 marks)

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

18) What is the purpose of the introduction section of a report? (2 marks)

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

19) What is meant by validity? (1 mark)

20) Explain how one factor in this study might affect its internal validity and how one factor might affect its external validity. (2 marks + 2 marks)

21) 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. (2 marks)

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.

22) Explain why the researchers asked independent judges to rate the toys. (2 marks)

23) Write a set of standardised instructions that would be suitable to read out to participants in this experiment. (5 marks)

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.

24) Write a non-directional hypothesis for this study. (2 marks)

25) Identify an appropriate sampling method for this study and explain how the psychologist might have obtained such a sample. (3 marks)

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 \( \chi^2 = 2.27 \). She then looked at the relevant table to see whether this value was statistically significant. An extract from the table is provided below.

<table>
<thead>
<tr>
<th>Level of significance for a one-tailed test</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level of significance for a two-tailed test</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20</td>
</tr>
</tbody>
</table>

| df   | 1.64 | 2.71 | 3.84 | 5.41 |

\( \chi^2 \) of 2.27 must be equal to or exceed the table (critical) values for significance at the level shown.

26) 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 x 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 marks)
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.

19) 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. (6 marks)

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.

20) The psychologist used a random sampling method. Explain how he could have obtained his sample using this method. (3 marks)

21) Explain limitations of using random sampling in this study. (3 marks)

22) 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 marks)

23) The psychologist used a Mann-Whitney test to analyse the data. Give two reasons why he chose this test. (2 marks)

24) He could have used a matched pairs design. Explain why this design would have been more difficult to use in this study. (2 marks)

25) Other than parental consent, identify one ethical issue raised in this study and explain how the psychologist might address it. (2 marks)

26) 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 marks)

18) Outline what is meant by the term peer review in psychological research. (2 marks)

19) Explain why peer review is important in psychological research. (5 marks)

Read the text below and answer questions 20, 21 and 22.

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.

20) What is meant by the term content analysis? (1 mark)

21) Explain how the psychologist might have carried out content analysis to analyse these drawings. (3 marks)
22) 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 marks)

23) 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 mark)

24) Table 1: 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 Table 1, state whether or not the psychologist’s result was significant. Explain your answer. (3 marks)
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 Table 2.

Table 2: Observation of one child – number of verbally aggressive acts in ten-minute time intervals

<table>
<thead>
<tr>
<th>Time slots</th>
<th>0–10</th>
<th>11–20</th>
<th>21–30</th>
<th>31–40</th>
<th>41–50</th>
<th>51–60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer A</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Observer B</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

25) Use the data in Table 2 to sketch a scattergram. Label the axes and give the scattergram a title. (4 marks)

26) Using the data in Table 2, explain why the psychologist is concerned about inter-rater reliability. (4 marks)

27) Identify an appropriate statistical test to check the inter-rater reliability of these two observers. Explain why this is an appropriate test. (3 marks)

28) If the psychologist does find low reliability, what could she do to improve inter-rater reliability before proceeding with the observational research? (4 marks)

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

Award 1 mark for a brief statement and a further mark for elaboration.
**Question 19**

AO3 = 4 marks

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

Credit variations in so far as they explain the process.

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

**Question 20**

AO3 = 3 marks

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

2 marks for some explanation/elaboration:

‘The two psychologists could carry out content analysis of the films separately and compare their answers’ or ‘they could re-code the films at a later date and compare the two sets of data’.

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

**Question 21**

AO3 = 3 marks

Candidates can cover one reason explained in detail here or several reasons in less detail. A repeated measures design was chosen in this experiment:

- To remove the effects of individual differences in reaction times which would occur if an independent groups design was used
- To avoid the potential difficulties involved in matching participants
- To reduce the number of participants required for the experiment.
Question 22
AO3 = 3 marks
This is a repeated measures design and is counter-balanced hence points about order effects and individual differences will not gain credit. There are a range of potential extraneous variables here including:
• The nature and content of the conversation with the psychologist on the hands-free phone
• Interaction between the sex of the psychologist and sex of participant which could influence the type of conversation
• The number of hazards in the computer-based test, hence difficulty of the tests
• The presence of the hands-free headset could have produced distraction.

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

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

Question 23
AO3 = 3 marks
External validity refers to how far the findings of the experiment can be generalised to real-life situations. The most likely answer here is that the hazard perception test was done using a computer test, which does not resemble real-life driving situations. (No noise, stress, etc.)

Award 1 mark for a brief answer (test lacks ecological validity) and 2 further marks for appropriate explanation contextualised within the scenario.

Question 24
AO3 = 4 marks
There are several potential ethical issues here. Candidates can focus on one in detail or several in less detail.
• Protection of participants from harm whilst studying the effects of a hands-free phones on driving. Two key issues here are the use of a computer-based test with no risk attached and of an experienced sample of police drivers.
• Informed consent: Participants should be given full information about the nature of both tasks before deciding whether or not to participate.
• Debriefing: A full debriefing should take place at the end of the experiment. This should provide feedback on performance and allow participants to ask questions if they wish to.
• Freedom to withdraw: Participants should be made aware of their freedom to withdraw before and during the experiment. They should be made aware of their right to withdraw their data after the experiment.
• Confidentiality: Individuals should not be identified, but should retain anonymity (use of numbers or initials instead of names).

Lists of ethical issues with no elaboration 1 mark.

**Question 25**
AO3 = 5 marks
The standardised instructions should include the following information:

a. You will take part in a simulated driving test, which will last for three minutes.
b. Your task will be to identify potential hazards on the road ahead.
c. When you see a hazard, you should press the mouse button as quickly as possible.
d. Whilst you are doing the test, I will chat to you on a mobile phone and I would like you to reply using the hands-free mobile phone headset.
e. Do you have any questions?

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

**Question 26**
AO3 = 3 marks
Students are required to identify an appropriate test and are asked to justify their choice. Award 1 mark for identification of the Wilcoxon (signed ranks) test. Candidates could receive credit for Sign test or related t test. Note that reasons/justification must be correct for the test supplied.

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

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

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

Test of difference cannot gain credit.
**Question 27**

AO3 = 2 marks

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

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

**Question 28**

AO3 = 3 marks

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

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

**Question 17**

AO2/AO3 = 2 marks

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

Award 1 mark for a non-directional hypothesis, which is not fully operationalized or lacks clarity (‘there is a relationship between happiness and intelligence’).

Award no marks for a null or directional hypothesis, or one that predicts a difference/link/association/connection.

**Question 18**

AO2/AO3 = 4 marks

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

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

Students could also make a case for the use of observation.
Award one mark for identifying an appropriate method. Award up to three further marks for an explanation of why this method would be better than a questionnaire.

**Question 19**
AO2/AO3 = 1 marks
Internal validity refers to how accurately a test or measuring instrument measures what it says it measures. Award one mark for a clear definition. Award no marks for truth, truth value or accuracy.

**Question 20**
AO2/AO3 = 3 marks
There are a number of ways of assessing the internal validity of the happiness questionnaire. Students can gain credit by explaining one in detail or more than one in less detail:
- Concurrent validity involves assessing how closely the scores on the happiness questionnaire match a different measurement of happiness obtained from the same participants, for example from family/teacher reports.
- Content validity involves asking experts in the field to check the content of the questionnaire to see how accurately it measures happiness.
- Face validity is less rigorous and involves looking at the questions to see if they are genuinely asking about happiness.

Award 1 mark for identifying one or more of these methods and a further two marks for elaboration in relation to the happiness questionnaire.

**Question 21**
AO2/AO3 = 2 marks
Award 1 mark each for any two of the following reasons:
- Study is looking for a correlation (relationship)
- Suitable for pairs of scores
- The data type obtained is ordinal, at least ordinal or interval level
- Linear relationship between scores

**Question 22**
AO2/AO3 = 3 marks
Students should state that the obtained value of + 0.42 exceeds the critical value for a two-tailed test (.362) for N=30. The results are therefore statistically significant (p ≤ 0.05)

Award 2 marks for a student who supplies two pieces of information.
Award 1 mark for a student who states that the results are significant but does not provide an explanation OR the student who states results are significant but uses incorrect values from the table.

Award 0 marks for students who argue that results are not significant.

**Question 23**  
AO2/AO3 = 4 marks  
This question requires students to interpret a further correlation coefficient (this time demonstrating a non-significant negative correlation) and put both findings together. For full marks, answers should cover the two key bullet points below:

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

Students may also make the point that there may be a weak tendency for more intelligent teenagers to be less happy at 16 years of age, although this is not statistically significant.

Students may also refer to the contradiction in the results or provide an overall conclusion.

**Question 24**  
AO1 = 2 marks  
Peer review is the process of subjecting a piece of research to independent scrutiny by other psychologists working in a similar field who consider the research in terms of its validity, significance and originality.

0 marks for ‘other psychologists look at the research’.
1 mark for a very brief outline e.g. ‘other psychologists look at the research report before it is published’. Award one further mark for elaboration.

**Question 25**  
AO2/AO3 = 4 marks  
Peer reviewed research may be accepted, sent back for revisions or rejected. Peer review is an important part of the scientific process because:

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

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

**Question 26**

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

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

In order to gain credit, students are required to identify one or more of these limitations and suggest appropriate modifications. There is a depth - breadth trade off here: students can cover one limitation in detail or consider several limitations in less detail.
Question 16
AO1 = 5 marks
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 e.g. 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.

Question 17
AO2/AO3 = 3 mark
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.
**Question 18**
AO2/AO3 = 3 marks
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.

**Question 19**
AO2/AO3 = 3 marks
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 e.g. ‘repeat the maths test’.

2 further marks for appropriate elaboration.

**Question 20**
AO2/AO3 = 2 mark
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)

**Question 21**

AO2/AO3 = 3 marks
Credit should only be awarded for scatter graphs. Other graphs gain 0 marks.
1 mark for appropriately plotted scores.
1 mark for an appropriate title.
1 mark for correctly labelled axes.
Question 22

AO2/AO3 = 3 marks

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 (i.e. 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 e.g. 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.

Question 23

AO2/AO3 = 10 marks

In question 23, 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 e.g. 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 (e.g. 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.

Question 24
AO2/AO3 = 3 marks
Award 1 mark for a clear table appropriate for the study described in question 23.

Musical ability scores:

<table>
<thead>
<tr>
<th>Participant number</th>
<th>Left handed</th>
<th>Right handed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Award 1 mark for the identification of an appropriate statistical test for the proposed design.

Award 1 mark for one correct justification e.g. a test of difference, at least ordinal level data.

**Question 16**
AO2/AO3 = 1 mark

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 i.e. memory is poorer with age.

**Question 17**
AO2/AO3 = 3 marks

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.

**Question 18**
AO1 = 1 mark

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

**Question 19**
AO2/AO3 = 3 marks

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.

**Question 20**
AO2/AO3 = 2 marks

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.

**Question 21**
AO2/AO3 = 2 marks
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

**Question 22**
AO2/AO3 = 2 marks

One mark for stating that the result is significant.

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

**Question 23**
AO1 = 2 marks
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.
**Question 24**  
AO2/AO3 = 3 marks

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.

**Question 25**  
AO2/AO3 = 4 marks

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 e.g. use of technical equipment.

**Question 26**

AO2/AO3 = 2 marks

- 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)

**Question 27**

AO2/AO3 = 10 marks

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:

- Operationalize 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 e.g. volume of music, time allocated for task should be credited.

Answers which make no reference to the control of, order effects maximum 2 marks.

Procedure – 4 marks
Procedural information should provide detail of how to go about conducting the study (i.e. what participants are required to do). Candidates could approach this task at a macro level i.e. 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.
Question 17
AO2/AO3 = 2 marks

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 e.g. ones which required attention to detail and to compare them with tasks which tap into very different skills e.g. creativity and thus to narrow down the benefits of providing red backgrounds.

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

Question 18
AO1 = 2 marks

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 e.g. '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.

Question 19
AO1 = 1 mark

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.
Question 20
AO2/AO3 = 2 + 2 marks

In this question, candidates have to make their answers relevant to this particular study. Candidates need to make it clear which factor refers to internal and which to external validity. Where candidates do not make this clear, examiners should accept the first factor as referring to internal validity and the second to external validity.

For each factor, one mark for a brief explanation and one further mark for elaboration.

Factors that might affect internal validity include:

• Individual differences e.g. colour blindness could have affect the outcome as the studies were all independent groups design
• Possibility of experimenter bias in judging the creativity of the ideas.

'Individual differences' = 1 mark.
'Because the researchers used an independent groups design, there could be a problem with individual differences' = 2 marks.

Factors that might affect external validity include:

• Sampling bias – all participants were university students
• Cultural bias – study took place in Canadian university – response to colours might well be affected by cultural factors.

'Sample bias' – 1 mark
'There was a sampling bias. Although all the participants were university students the investigators drew more general conclusions = 2 marks'

Question 21

AO2/AO3 = 2 marks
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 decide what colour enhances creativity in a new product design process.
Any plausible practical applications are creditworthy.
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.

**Question 22**
AO2/AO3 = 2 marks

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.

**Question 23**
AO2/AO3 = 5 marks

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.'
**Question 24**  
AO2/AO3 = 2 marks

There is an association between birth order and choice of career'  
=2 marks

A directional hypothesis is not creditworthy. 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 e.g. '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.

**Question 25**  
AO2/AO3 = 3 marks

One mark for identifying a sampling method.  
One mark for a brief explanation of how to obtain the sample e.g. 'by advertising for lawyers or artists to come forward’. One further mark for elaboration e.g. ‘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.

**Question 26**  
AO2/AO3 = 12 marks

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

<table>
<thead>
<tr>
<th></th>
<th>Artists</th>
<th>Lawyers</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>First born</td>
<td>20</td>
<td>35</td>
<td>55</td>
</tr>
<tr>
<td>Not first born</td>
<td>30</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Totals</td>
<td>50</td>
<td>65</td>
<td>115</td>
</tr>
</tbody>
</table>
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% ($p \leq 0.05$). 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.

**Question Stem**

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.

**Question 19**

AO2/3 = 6 marks

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 (operationalization).
• 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.

**Question Stem**
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-olds from ten local schools, and decides to use 100 children (50 in 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.

**Question 20**
AO2/3 = 3 marks

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 textbooks 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 i.e. 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' i.e. cannot see the actual slips of paper.

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

**Question 21**

AO2/3 = 3 marks

Candidates could focus on:
• Even if a sample is random, it may not be truly representative of the population e.g. might all come from the same school, or be all boys or all girls.
• Practical limitations e.g. 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.

**Question 22**

AO2/3 = 5 marks

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.

**Question 23**

AO2/3 = 2 marks
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).

**Question 24**

AO2/3 = 2 marks

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 (e.g. 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 (i.e. 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 e.g. “because it is difficult to match participants appropriately”.

One further mark for elaboration.

**Question 25**

AO2/3 = 2 marks

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.

**Question 26**
AO3 = 12 marks

**Question Stem**

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 behaviour.

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, co-operative 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.

Question 18
AO1 = 2 marks
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 e.g. ‘other psychologists look at the research report before it is published.’

One further mark for elaboration.

Question 19
AO2/3= 5 marks

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.

**Question 20**

**Question Stem**

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.

**AO1 = 1 mark**

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….?’

**Question 21**

**AO2/3 = 3 marks**

- 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 e.g. carbohydrate, protein; the state of the food e.g. 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 4 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 e.g. “choose a theme like size”.
Note: maximum 2 marks if no engagement with the stem.

**Question 22**

**Question Stem**
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 that there is sufficient information about the study for the participants to make an informed decision. 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.

**Question 23**

**Question Stem**
The psychologist was also interested in the effects of a restricted diet on memory functioning and he expected memory to become impaired. His hypothesis was as follows:
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 4 week period. He recorded the memory scores
on both tests and analysed them using the Wilcoxon test. He set his significance level at 5%.

His calculated value was $T = 53$

**AO2/3 1 mark**
1 mark for correct answer – directional (one-tailed is acceptable)

**Question 24**

**AO2/3 3 marks**
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 critical 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.

**Question 25**

**Question Stem**

A psychologist uses the observational method to look at verbal aggression in a group of schoolchildren 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>
<thead>
<tr>
<th>Table 2: Observation of one child - number of verbally aggressive acts in ten-minute time slots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time slots</td>
</tr>
<tr>
<td>Observer A</td>
</tr>
<tr>
<td>Observer B</td>
</tr>
</tbody>
</table>

**AO2/3 = 4 marks**
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 on page 21 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.
**Question 26**

AO2/3 = 4 marks

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 than 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.

**Question 27**

AO2/3 = 3 marks

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.
Question 28
AO2/3 = 4 marks
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 e.g. explain how the training might be improved or outline several improvements in less detail e.g. establish clearer criteria for categorising verbal aggression, filming the child so that the observers can practise the categorisation.