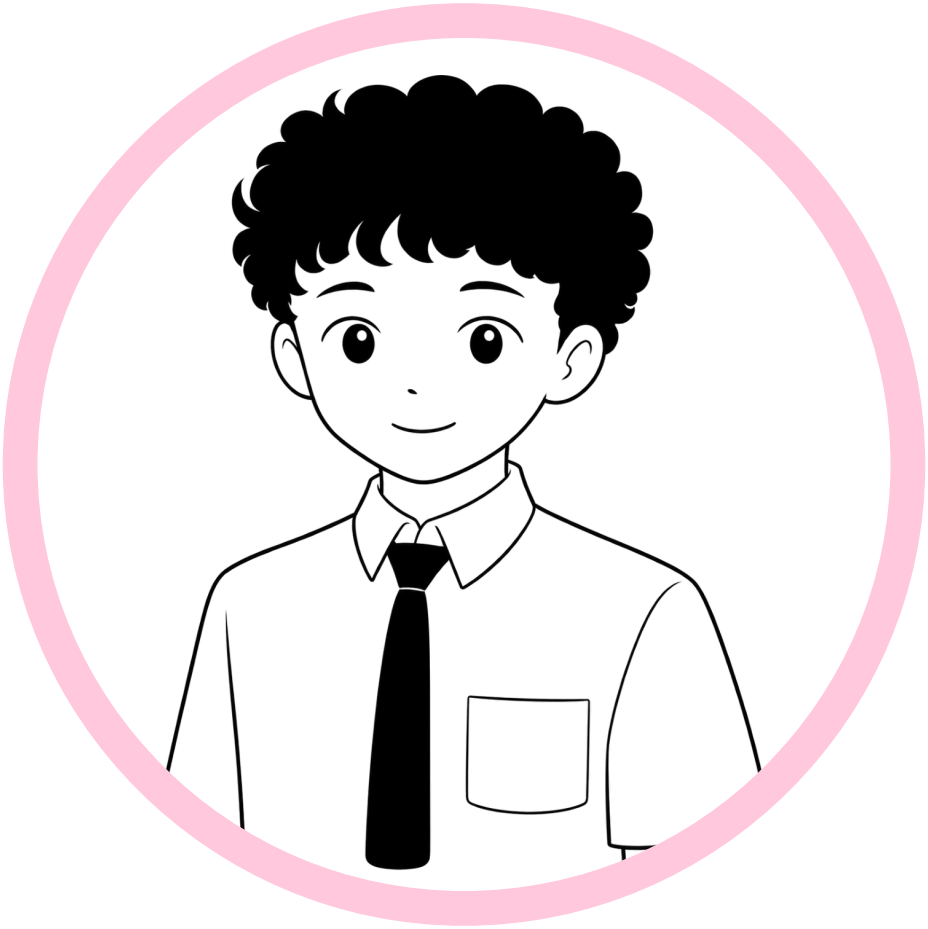
AQA A-LEVEL PSYCHOLOGY · 7182

PSYCHOLOGY



MARK SCHEME

3.2.3.1 AIMS AND HYPOTHESES

SCORE	PERCENTAGE
/289	%

STEMUP

These questions belong to AQA, not to us.

The examination questions, mark schemes and examiner reports in this pack were written and published by AQA. They are past assessment materials, reproduced here so that students can revise from them.

What we do not own

STEMUP is not affiliated with AQA and does not own, and makes no claim to own, any of the question, mark scheme or examiner report material in this pack. AQA has not endorsed, approved or been involved in the production of it. AQA and the qualification titles used here belong to their respective owners and appear for identification only.

Where to find the originals

AQA publishes its past papers, mark schemes and examiner reports free of charge. Always check the current specification at aqa.org.uk before you revise, because content and wording can change between series.

If you hold the rights to any material in this pack and would like it amended or removed, contact admin@stemup.co.uk and we will act on it.

Q1.

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

Level	Mark	Description
3	5-6	Consent form is well detailed and practical, showing sound understanding of the requirements of a good consent form. All three elements are appropriately represented. The answer is clear and coherent. Specialist terminology is used effectively. Minor detail and / or explanation sometimes lacking.
2	3-4	Consent form is mostly detailed and practical, showing some understanding of the requirements of a good consent form. The answer is mostly clear and well organised. Specialist terminology is mostly used effectively.
1	1-2	Consent form only includes basic details of what participants can expect to happen in the experiment or how they will be protected. Knowledge of the requirements of a good consent form is limited. The whole answer lacks clarity, has many inaccuracies and is poorly organised. Generic consent forms without link to the stem.
	0	No relevant content.

Possible content:

Outline of what the experiment entails:

- an explanation of the general purpose of the research
- they would need to watch an assigned horror film every night before going to bed for 7 nights and would need to watch an assigned romantic comedy every night before going to bed for 7 nights
- a daily requirement to truthfully respond to a text message asking whether they had experienced a nightmare
- the two-week duration of the experiment.

Ethical guidelines:

- no pressure to consent
- they can withdraw at any time
- they can withdraw their data from the experiment
- 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.

Format and style of consent form:

STEMUP.CO.UK

- will require the participant's agreement
- could be written as a form that participants need to sign
- could include space for the participant to sign the consent form
- could include a space for the participant to write the date
- could include a space for the participant to print their name.

Credit any other relevant information.

6

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

3 marks for an appropriate, clear and coherent directional operationalised hypothesis.

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

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

0 marks for expressions of aim / questions / correlational / null / non-directional hypotheses or statements with only one condition of the IV present OR if the IV is incorrectly described as group 1 and group 2.

Possible content:

- participants will report more nightmares after watching a horror film before bedtime than after watching a romantic comedy film before bedtime. Accept alternative wording
- participants will report fewer nightmares after watching a horror film before bedtime than after watching a romantic comedy film before bedtime. Accept alternative wording.

3

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

2 marks for a clear explanation that is linked to the experiment

1 mark for a limited or muddled explanation.

Possible explanations:

- necessary to avoid the effects of individual differences in frequency of nightmares
- film viewing habits, gender, hours of sleep, personality, etc, can have a big impact on the number of nightmares recalled.

Credit other relevant explanations.

2

(d) **[AO1 = 1 AO2 = 1]**

1 mark for each of the following points:

- order effects are likely to occur
- if you saw the horror films in the first week you may still be thinking about them / have higher levels of fear in the second week, resulting in more nightmares.

STEMUP.CO.UK

Accept alternative wording.

2

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

1 mark for each of the following points (up to 3 marks):

- all 50 participants' names / numbers are put into a hat / container / computer
- a name is drawn from the container or a random name is generated by the computer and is assigned to the first group
- a second name is selected as before but this time goes in to the second group; this process continues until there are 25 in each group.

OR

1 mark for each of the following points (up to 3 marks):

- use 50 slips of paper, label 25 'horror' and 25 'romantic comedy'
- put / mix the slips of paper in a hat / container
- each participant draws a group from the container

Note: Credit alternative methods of randomly splitting the sample

3

(f) **[AO2 = 2 AO3 = 2 maths]**

Mean:

1 mark – participants who watch horror films before going to bed report more nightmares then those who watch romantic comedies before bed. Accept alternative wording.

Plus

1 mark – mean number of nightmares reported is greater when horror films are watched than when romantic comedies are watched. Accept alternative wording.

Standard deviation:

1 mark – there is greater dispersion / variation in scores / spread of scores in the horror film condition than in the romantic comedy condition. Accept alternative wording.

Plus

1 mark – standard deviation is greater when horror films are watched before going to bed than when romantic comedies are watched before going to bed. Accept alternative wording.

Note - **0 marks** for just stating the data from the table.

Note – Justifications are not creditworthy in isolation

4

(g) **[AO2 = 2 maths]**

2 marks for a clear and appropriate explanation in the context of this experiment.

1 mark for a limited or muddled explanation.

STEMUP.CO.UK

This means that the difference in the number of nightmares reported after watching horror films compared to romantic comedies is significant at 0.05 level. This means there is less than 5% (1 in 20) likelihood (probability) that the difference was due to chance / due to something other than the IV.

Accept any other valid answer.

2

(h) **[AO2 = 1 AO3 = 2]**

1 mark for a modification to the design of the experiment that could improve validity.

Plus

2 marks for clear and coherent explanation for how the suggested modification might improve validity of this study.

1 mark for a limited/muddled explanation for how the suggested modification might improve validity of this study.

Possible content:

- include more than one question in the text message to the students. This would make the aim of the experiment less obvious to guess which would in turn reduce demand characteristics and improve the validity of the experiment
- make the conditions less obvious. Rather than having one film the students could be asked to watch an episode from a TV series. The episodes watched in one condition would be those which contained any scary concepts / themes whereas episodes watched in the other condition would not contain any scary themes / concepts
- guarantee anonymity so people will give honest answers and not feel embarrassed
- use a broader sample, not just students. Students may be more or less inclined to watch horror films anyway
- use a different sampling technique to avoid a self-selected sample and thus avoid volunteer bias as that may make them more susceptible to demand characteristics.

Accept alternative answers.

Note: Independent groups design would not be an appropriate modification in this study; however candidates can make the case for a matched pairs design, which could be credit worthy.

3

[25]

Q2.

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

2 marks for explanation that a non-directional hypothesis is suitable or ‘it should not be directional,’ (1) as there is no reference to evidence that allows the psychologist to predict the direction of the results (1).

1 mark for a muddled/limited explanation of why the hypothesis should be non-directional or 1 mark for stating non-directional/‘no’.

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

STEMUP.CO.UK

2 marks for correct answer 55(%)

1 mark for correct workings only 22/40 = 0.55 × 100

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

2 marks for a clear explanation of why stratified sampling might improve the study.

1 mark for a muddled/limited explanation

Relevant points:
- stratified sampling increases representation/generalisation by sampling many subsets of students
- stratified sampling reduces research bias as the subsets of students are selected randomly.

Credit other relevant answers, e.g. stratified sampling increases external validity/population validity by sampling many subsets of students, or justification through comparison with volunteer sampling.

Note: No marks for simply stating increases validity/reliability

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

2 marks for a clear explanation of primary data.

1 mark for a muddled or limited explanation (e.g. reference to 'first hand' or focus on researcher not data).

Content
- primary data is original data/first hand response of the participants/source in the research.
- collected specifically for the research being carried out.

Credit answers that refer to the data in the study described.

Note: do not credit an answer which simply states collected by the researcher

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

2 marks for identification and a clear explanation of one reason why the mean is the most appropriate measure of central tendency.

1 mark for identification of one reason.

Possible reasons:
- the mean can be said to be **representative of all the data** collected as it is calculated using all the individual values.
- the mean is the **most sensitive measure of central tendency** as it uses all the values in set of data.

Accept other valid reasons.

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

Task A

2 marks for correct answer 63 (secs)

STEMUP.CO.UK

1 mark for correct workings only 630/10

Task B

2 marks for correct answer 77.4 (secs)

1 mark for correct workings only 774/10

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

3 marks for a clear and coherent discussion of the purpose of counterbalancing.

2 marks for some relevant discussion.

1 mark for a muddled/limited discussion.

Possible content:
- it controls the impact of order effects (practice, fatigue or boredom)
- allows order effects to be distributed evenly across both conditions
- making each condition of the IV occur as the first task and the second task equally
- counterbalancing does not eliminate order effects which will be present because there are two separate tasks to be completed by each person.

Note: No marks for answers which only state decreasing/reducing/eliminating/preventing order effects without reference to any of the above points.

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

1 mark for identification of **one** appropriate extraneous variable.

Possible variables: that can be controlled and need to stay constant for all participants or from Task A to Task B to avoid affecting the dependent variable:
- level of difficulty in the two 'spot the difference tasks'
- wearing of headphones
- the music/tracks
- instructions for participants
- characteristics of the psychologist
- change in environmental conditions.

Accept other relevant variables
Do **not** accept participant variables

Possible explanations:

1 mark for elaboration of the effect of the EV.

1 mark for how it might have affected the DV (the effect on time).

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

Level	Marks	Description
2	3 – 4	There is a clear and detailed explanation of how the second study improves on the first study. The answer is generally coherent with effective use of appropriate

STEMUP.CO.UK

		terminology.
1	1 – 2	There is limited or partial explanation of how the second study improves on the first study. The answer lacks coherence and use of appropriate terminology.
	0	No relevant content.

Possible content:

- the second study is more objective. In the first study participants self-reported what they believed to be the case and this is refuted by the follow-up study
- the measurement of time allows for more detailed analysis of data
- use of experimental method involves manipulation of IV, establishing cause and effect in follow-up study which is not a feature of the self-report method used in the first study
- possibility of socially desirable answers as a feature of first study which might be unreliable and this is not present in follow-up study where measurements are more direct.

Credit other possible explanations.

Note: no marks for simply stating more valid/reliable.

Q3.

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

2 marks for a clearly outlined disadvantage that is explicitly related to the study – sample is biased, lacking representativeness, more confident/helpful/curious which could affect how they respond to therapy or represent other offenders.

1 mark for knowledge of a disadvantage of using volunteers which could be related to the study but this is not explicitly presented.

Credit other relevant disadvantages.

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

Award **1 mark** for each of the following points up to a maximum of 3 marks.

- All the volunteers are identified either by name or number.
- The 50 names/numbers are put in a container or computer.
- Assign alternate names/numbers drawn to Group 1 then Group 2 and so on until there are 25 in each group or alternative system **OR** set parameters for two groups of 25 to be randomly generated.

Credit alternative descriptions of a practical procedure which would result in the same outcome.

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

3 marks for a clearly stated and appropriate non-directional (or directional) operationalised hypothesis: there is a difference in the reduction in anger scores of offenders who complete Therapy A and those who complete Therapy B.

2 marks for a statement with both conditions of the IV and the DV that lacks the clarity of the 3 mark answer.

STEMUP.CO.UK

1 mark for a muddled statement with both the IV and DV present.

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

Full credit can be awarded for a null version of the hypothesis.

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

Award **1 mark** for any of the following points up to a maximum of 3 marks.

Possible points:

- Both therapies seem to be effective.
- The pre-therapy scores for both groups suggest that the offenders in Group 1 were more angry on average than those of Group 2 at the start of study.
- The post-therapy scores for both groups suggest that the offenders in Group 1 were more angry on average than those of Group 2 at the end of study.
- Anger scores are lower at the end of the study for both groups suggesting both therapies reduced feelings of anger.
- The improvement in Group 1 is 50-30, a decrease of 20 and in Group 2 is 40-15, a decrease of 25, which suggests that there was a greater reduction in feelings of anger in Group 2.

Credit other relevant points.

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

Level	Marks	Description
2	3 – 4	The explanation for how matched pairs would have improved the study is clear and the answer identifies the difference in the pre-therapy scores as the issue. The answer is generally coherent with effective use of terminology.
1	1 – 2	There is limited/partial explanation for how matched pairs would have improved the study. The answer may lack coherence. Use of terminology may be either absent or inappropriate. **OR** the answer only xplains matched pairs design in the context of this study.
	0	No relevant content.

Possible points:

- The researcher wants to be sure that the final anger scores are a result of the therapy and not participant differences between the groups.
- With independent groups participant variables might have caused the lower post-therapy scores for Group 2 rather than the therapy.
- Matching the anger scores of the participants at the pre-therapy stage would reduce this difference – seen in the Group 1 median of 50 which is ten points higher than the Group 2 median of 40.
- Other variables could also affect response to therapy e.g. gender, type of offence.

STEMUP.CO.UK

N.B. Answers that only explain a matched pairs design in the context of this study cannot be awarded above level 1.

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

2 marks for a clear outline of an ethical issue that might have occurred in the study.
1 mark for a brief or muddled outline of an ethical issue that might have occurred in the study.

- In this study the participants are all in custody so despite the suggestion that they are volunteers they may have been coerced in some way to participate, this could relate to issues of consent or stress or right to withdraw.
- Distress/psychological harm such as: finding out they have high-anger scores; the questionnaire might trigger more anger; finding out there are two therapies and finding out one might be better; therapies might cause more harm than good; the participants might have some expectation that the programme would be effective and this might not be the case.
- Reducing the anger of the participants may be more beneficial for staff than the participants, manipulation of their behaviour for this reason could be unethical.

Plus
2 marks for a clear explanation of how the issue described above could be dealt with. It is expected that withdrawal of data or debriefing would be appropriate answers but these should be clearly matched with the chosen ethical issue.

1 mark for a brief or muddled explanation of how the issue could be dealt with.

Credit other relevant information.

(g) **AO2 = 1**

1 mark for knowledge of an appropriate disadvantage of the median.

Possible answers
- Any outlier values/extreme values (such as the scores of 50) would be ignored/would not form part of the average measurement.
- Less sensitive than the mean.
- It does not represent all the findings.

(h) **AO2 = 2**

2 marks for a clear explanation of how demand characteristics might have occurred in the study. The explanation will make clear that participants alter their behaviour in some way because of a belief they have about their participation in the study.

1 mark for a limited or muddled explanation that is linked to the study.

Possible points

- Participants might respond to the questionnaire with answers giving a reduction in anger as they expect the therapy is to reduce anger/they know this is what the researcher expects.
- Participants might ‘respond’ with an increase/decrease in anger in order to ‘ruin’ results.
- When the researcher gives out the questionnaire for the second time, this sets up a demand characteristic that the second answers should be different to the first answers.

STEMUP.CO.UK

Credit other relevant information.

(i) **AO3 = 2**

2 marks for a clear discussion of one strength of using questionnaires in research which might include: the ability to collect large amounts of data relatively quickly and conveniently; discussion of how they might be easy to score/collate – when items are closed; reference to standardisation/replication of method.

1 mark for limited or muddled explanation of a possible strength.

N.B If the strength discussed relates to a particular type of questionnaire, the type of questionnaire must be made clear for the second mark.

Credit other relevant strengths.

Q4.

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

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

Plus

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

OR

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

Possible advantages:

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

Credit other relevant advantages.

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

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

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

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

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

1 mark for an appropriate and plausible suggestion.

Plus

STEMUP.CO.UK

1 mark for an appropriate justification.

Likely suggestions:

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

Credit other relevant suggestions.

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

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

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

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

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

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

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

1 mark: 3 (in each group)

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

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

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

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

or written as

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

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

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

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

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

STEMUP.CO.UK

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

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

Plus

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

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

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

Required content:

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

Optional content:

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

Q5.

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

2 marks for identification of dependent variable operationalised: number of verbal errors.

1 mark for dependent variable not operationalised: verbal errors or fluency or mistakes.

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

3 marks for an appropriate non-directional (or directional) operationalised hypothesis:

STEMUP.CO.UK

‘There is a difference in number of verbal errors made by participants who perceive / think / believe there are 5 listeners (there is a small audience) and by participants who perceive / think / believe there are 100 listeners (there is a large audience)’.

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

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

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

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

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

1 mark for identification of **one** appropriate extraneous variable.

Plus

2 marks for explanation of why the variable should have been controlled – for full marks this should include clear explanation of how it would have affected the DV. Award one mark only for muddled or incomplete explanations, eg unelaborated reference to ‘avoiding confounding’.

Appropriate variables: can be controlled and need to stay constant to avoid affecting the dependent variable, eg same article / conditions / instructions for each participant.

Do not credit gender (this is controlled) or time to complete task (cannot be controlled).

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

2 marks for clear and coherent explanation of one advantage of using a stratified sample in this study.

1 mark for a muddled answer with a relevant advantage and some explanation in relation to the study.

Possible advantage: ensures that this sample is truly representative because different types of people (males / females) working in this company are represented in the sample in the correct proportions.

Accept other relevant advantages.

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

1 mark for each point as follows:

Manual method:

- put all 60 male names in a hat (or similar)
- determine the proportion of males needed to mirror the number of males in the target population as follows: 60%
- calculate 60% of 20 = 12 and draw out 12 names.

STEMUP.CO.UK

Random number table or computer method:

- assign each of the 60 men a number between 1 and 60
- determine the proportion of males needed to mirror the number of males in the target population as follows: 60%
- calculate 60% of 20 = 12 and moving horizontally or vertically through random number tables find 12 numbers between 1 and 60 for the sample **OR** generate 12 numbers between 1 and 60 using random number generation function on computer.

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

Marks for a clear description of a practical way of randomly allocating the 12 men and 8 women to the two conditions as follows:

- give each man a number 1 – 12 (1 mark)
- put 12 numbers in a hat (1 mark)
- assign first six numbers drawn to Condition A with the remainder for Condition B (1 mark)
- repeat process for women – eight numbers in the hat and draw four for Condition A and remaining four go to Condition B (1 mark).

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

Q6.

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

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

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

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

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

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

1 mark for stating scattergraph or scattergram.

Plus

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

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

STEMUP.CO.UK

Possible content

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

Accept other relevant comments

(e) **[AO2 = 2 AO3 = 4]**

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

Possible problems:

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

Possible modifications:

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

Credit other relevant information.

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

Possible content

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

Credit other relevant information

(g) **[AO2 = 2 AO3 = 2]**

STEMUP.CO.UK

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

Possible content:

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

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

Q7.

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

1 mark – the hypothesis should be directional

Plus

1 mark – because there is past research indicating the likely direction of the effect (or similar)

(b) **[AO1 = 2 AO3 = 2]**

2 marks for a clear and coherent explanation of operationalisation

1 mark for a brief or muddled explanation of operationalisation

Content: operationalisation involves clearly specifying/defining observable behaviours that represent the more general construct under investigation/to enable the behaviour under investigation to be measured

Plus

1 mark each for two observable behaviours that could represent ‘riding a bike with care’.

STEMUP.CO.UK

Examples: use of cycle lanes/tracks, passing pedestrians at a distance of at least 1 metre, using bicycle bell.

Credit any relevant observable behaviour.

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

1 mark for knowledge of the term face validity – where a behaviour appears at first sight (on the face of it) to represent what is being measured

Plus

2 marks for clear and coherent application of the concept of face validity to the context

1 mark for brief or muddled application of the concept of face validity to the context

Application: Having a dog on a lead appears at first glance to be measuring considerate behaviour because if a dog is on a lead it is less able/likely to upset other people by coming close, frightening, chasing, biting, growling etc.

Credit other relevant applications.

(d) **[AO1 = 4]**

1 mark for each of two types of validity identified

Plus

1 mark each for a brief outline of each type of validity identified

Content:

- Concurrent – where performance on one measure correlates highly with performance on another measure of the same variable
- Ecological – where a measure of a behaviour accurately reflects the way in which the behaviour would occur in normal circumstances
- Temporal – where findings from research that took place at a certain point in time accurately reflect the way that behaviour would occur at a different point in time

Credit also other types of validity eg criterion, content, construct, population, predictive.

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

1 mark – C Event sampling

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

1 mark for each of the following:

- The student pair should discuss and agree beforehand their interpretation of the behavioural categories
- Each student should then observe the same people/space/target at the same

STEMUP.CO.UK

time but record/tally independently

- Their respective recordings/tallies should be correlated using an appropriate statistical test ascertain the level of agreement

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

1 mark for the correct ratio: 3:2

Plus

2 marks for full workings: 23 + 12 + 19 = 54 and 10 + 17 + 9 = 36

54:36 both divisible by 9 (or 18)

54 ÷ 9 = 6 and 36 ÷ 9 = 4

6:4 can be simplified to 3:2

1 mark for partial workings eg first 2 of the above stages

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

1 mark for each of the following:

- Data is categorical/nominal/frequency
- The students are looking for a difference or an association between two variables
- Design is independent/unrelated or categories are exclusive (observations cannot appear in more than one cell)

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

1 mark for the correct answer: df = 2

Plus

1 mark for correctly substituting values into the formula as follows:

(2 − 1) x (3 − 1) = 1 x 2 = 2

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

1 mark for each of the following:

- Yes, the result is significant
- Because the calculated value of Chi-square is more than the critical table value at 0.05, either 4.60 or 5.99, depending on whether student uses one or two-tailed values
- One-tailed is consistent with the hypothesis and Q13, but there is an argument that the Chi-square test should always be two-tailed so either can be credited
- Where df equals 2

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

Level	Marks	Description
2	3 – 4	The paragraph is clear and coherent, showing sound understanding of the concept of levels of significance and effective application to the context. There is effective use of terminology.
1	1 – 2	The paragraph shows some understanding of the concept of levels of significance and/or some relevant application to the context. The answer lacks clarity and coherence. Terminology is either absent or inappropriately used.
	0	No relevant content.

Possible content:

- Explanation of levels of significance as an indication of the measure of the influence/effect of chance/random factors on the findings.
- With the present results there is a 95% confidence in accepting the research hypothesis/confidence that any difference/effect is due to the variables under investigation, in this case the location of the public spaces.
- There is a 5% possibility that the same frequencies would occur if there was no real difference between the two towns.
- The calculated value in this case well exceeds the critical value at 0.05 but does not meet the more stringent level of significance of 0.01.
- Possibility of type one error.

Credit other relevant material.

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

1 mark for explaining that content analysis is suitable because the students are analysing recordings which are a form of media.

Plus

2 marks for a clear, coherent account of how the content could be analysed

1 mark for a brief or muddled account of how the content could be analysed

Content:

Students could identify specific ideas/concepts that occur in the recordings

They could then set up a system of categories and tally the ideas/concepts

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

1 mark for any relevant inconsiderate behaviour eg leaving rubbish, leaving dirty mugs/plates, playing music loudly, throwing books, shouting

STEMUP.CO.UK

Q8.

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

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

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

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

(a) **AO3 = 4**

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

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

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

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

AO3 Application of knowledge of research methods
4 marks Accurate and reasonably detailed Accurate and reasonably detailed explanation that demonstrates sound knowledge and understanding of why a pilot study would be appropriate, including at least one detail from the experiment.
3 marks Less detailed but generally accurate Less detailed but generally accurate answer that demonstrates sound knowledge and understanding of why a pilot study would be appropriate, including at least one detail from the experiment.
2 marks Basic

STEMUP.CO.UK

Basic answer that demonstrates some understanding of why a pilot study would be appropriate in this study, but lacks detail and may be muddled.
1 mark Very brief / flawed Very brief or flawed answer demonstrating very little understanding of why a pilot study would be appropriate in this study.
0 marks No creditworthy material.

(b) **AO3 = 2**

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

(c) **AO3 = 4**

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

(d) **AO3 = 2**

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

Q9.

STEMUP.CO.UK

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Students should state that the obtained value of + 0.42 exceeds the critical

STEMUP.CO.UK

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

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

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

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

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

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

Q10.

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:

STEMUP.CO.UK

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

(a) **AO3 = 2**

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

(b) **AO3 = 2**

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

(c) **AO3 = 4**

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

(d) **AO3 = 4**

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

Q11.

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:

STEMUP.CO.UK

- 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.

Q12.

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

STEMUP.CO.UK

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

STEMUP.CO.UK

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.

Q13.

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

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

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

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

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

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**

STEMUP.CO.UK

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.’

STEMUP.CO.UK

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.

Q14.

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 / 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**

STEMUP.CO.UK

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.

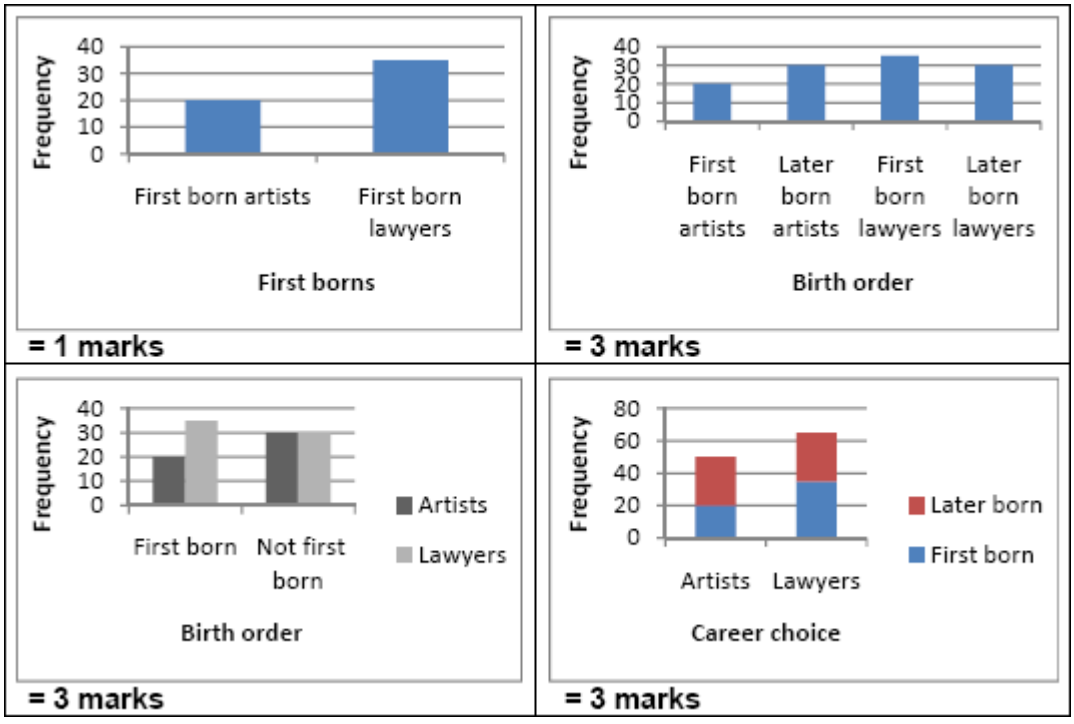
Table: Table to show the career choices of first born and non-first born children

	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:

STEMUP.CO.UK

- 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.

Q15.

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 directional hypothesis.
1 mark if not operationalised, eg “Age affects memory.” “There will be a difference between the two conditions.”
2 marks for eg “There will be a difference in how many numbers are correctly recalled by children and adults.” “Children and adults have different short-term memory spans.” Or “The capacity of short-term memory is different for adults and children.

Candidates may write a hypothesis where the IV is how many numbers are in the list and the DV is the number of participants who can recall that digit span.
Eg As numbers in the list increase, recall changes. 1 mark.
As the number of random numbers in the list increases, the number of participants recalling the list correctly, changes. 2 marks.

(b) **AO3 = 2**

The experiment uses adults in one condition and children in the other so it would be impossible to use a repeated design unless the researchers waited for the children

STEMUP.CO.UK

to grow into adults.
Given the nature of this experiment, demand characteristics and order effects are inappropriate.
1 mark for a brief explanation. A further mark for elaboration. Eg Can compare the two different groups to see who is better. 0 marks (because this relates to all experimental designs).
They needed to have different people in each condition. 1 mark.
They needed to have different people in each condition based on age. 2 marks.
They needed to have children in one group and adults in the other. 2 marks.

(c) **AO3 = 2**

Children 6
Adults 7
1 mark for each correct answer.

(d) **AO3 = 3**

The frequency distribution shows that there is a difference in results between the two age groups.
Adults recalled more digits than children. However, the difference is small and some children recalled more digits than some adults. Candidates might refer to the modal scores being different while the range is the same.
Any credit-worthy material should be credited.
1 mark for a very brief answer eg identifying there is a difference between adults and children and / or adults score more than children. Further marks for more detail as above.

(e) **AO2 = 2**

Other research has suggested the capacity of short-term memory is 7 + / – 2. The results do support this as the range is from 5-9.
1 mark for a brief or muddled explanation eg capacity is 5-9 / other research has similar findings.
2nd mark for elaboration as above.
Candidates will be credited for reference to research such as Jacobs which found STM increases with age. However, reference to such research is not a requirement.

Q16.

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.

STEMUP.CO.UK

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

Q1.

Part (a) was generally quite poorly answered. Although many students could gain some credit for providing detail from the study and / or giving ethical issues, few were able to demonstrate sound understanding of the requirements of a practical consent form with many not referring to consent at all. Ethical issues were generally covered reasonably well and were the strongest of the three elements, whilst format / style was the weakest and was often inappropriate (eg 'You will', or 'You must', or 'You have taken part'). Although there were some excellent responses, generally it appeared that the students had not really thought about the nature of giving informed consent, with sparse experimental detail and the incorrect tone used, often reading more like a brief / debrief from or a legal disclaimer. This is perhaps an area that schools / colleges could look at in more detail.

In part (b) despite many students showing an understanding of directional hypothesis, just less than half failed to achieve any marks, mainly due to a focus on horror films leading to more nightmares with no reference to the romantic comedy condition. Additionally, many students failed to fully operationalise the variables and some struggled to write an appropriate hypothesis for a repeated measures design, employing a writing frame for an independent measures design.

In part (c) most students identified individual differences as a reason to use a repeated measures design but only about half of these students managed to clearly link their reason to the experiment. Surprisingly, there was some confusion over the term 'repeated measures', with some students focussing their answers on reliability and others on order effects. Responses to this question again suggest that students need to gain more practical experience of designing and conducting experiments as part of their course, rather than focussing on theoretical learning of these concepts.

Part (d) offered a great opportunity to distinguish genuine understanding of reasoning behind the use of counterbalancing as opposed to whether students had merely rote learned what it is and why it is used. Overall, there appeared very little understanding with approximately half of students failing to achieve any marks and approximately one tenth of students accessing both marks. Most students achieving marks identified the issue of order effects but application to the experiment was limited, with many rote learned, inappropriate responses such as boredom / tiredness given rather than engaging with this study.

Part (e) was generally well answered with over half of the students accessing all 3 marks. Generally, answers referring to a hat / container were more successful than those involving a random number generator / computer, as these responses often failed to fully describe the process.

Overall, students were able to answer part (f) well, although understanding of mean scores was generally better than standard deviation. It appeared that students often understood the data but failed to appropriately justify their suggestions, often simply restating information given in the table. There were also a few costly examples of justifications given without suggestions.

Part (g) was a good differentiator, with some excellent responses by students who clearly explained what 'significant at $p<0.05$' meant and successfully managed to do so in the context of the experiment. Weaker students generally gave a rote learned reason for why psychologists generally use the 5% level or failed to contextualise their response. Unfortunately, many students failed to explain that the '<0.05' sign meant less than 5%. Teachers should ensure that students are familiar with the mathematical symbols required

by the new specification.

Part (h) was generally not well answered. A huge range of modifications to the design of the experiment were given but unfortunately a lot of these were inappropriate, such as 'use a control group', 'use a blind procedure', 'use a repeated measures design' or 'use an independent measures design', suggesting either a lack of understanding of the reasoning for experimental design or a failure to recall the experimental protocol. Some students suggested using a matched pairs design and went on to explain effectively how this could improve validity by reducing the likelihood of participants guessing the aim of the study which would reduce demand characteristics. More commonly students justified the modified design by inappropriately suggesting that it would avoid individual differences. Changing the sample was a common answer but it was often not clear how this would improve validity, with few being able to give more than a generic explanation. Stronger responses generally focussed on the problem of using a text message to ascertain whether a nightmare had been experienced, suggesting a questionnaire/interview to reduce the chances of participants lying or to help distinguish nightmares from unpleasant dreams.

Q2.

(a) There was evidently much confusion amongst students as to when a directional or non-directional hypothesis should be used and many failed to get any marks on this question. Those who did correctly state the hypothesis should be non-directional portrayed this is many ways, ranging from 'the hypothesis should be non-directional' or 'the hypothesis should not be directional', to simply 'no' in response to the question stated. However, only around half of these correct responses went on to accurately explain the reason for this in relation to the fact that there was no previous research. Many students stated that 'the hypothesis should be directional as it was clearly seen from the data that participants…'. It appears that many have not had sufficient practical experience of selecting an appropriate hypothesis and thus did not understand how to apply their knowledge in order to accurately answer this question.

(b) The overwhelming majority of students clearly knew how to calculate a percentage and most were awarded two marks. Virtually all also provided their workings to support their answer. Those students who gave an incorrect answer but included correct workings were awarded one mark.

(c) Most students clearly understood that stratified sampling was more representative or reduces research bias but many did not refer to subsets of students in the correct proportions. A minority confused stratified sampling with systematic sampling. There were also a number of students that didn't attempt the question, suggesting that they did not have the knowledge.

(d) The vast majority of students scored one mark for this question with their understanding that primary data is first hand data. Very few referred to it as being 'original' data. Even fewer students referred to primary data as data collected specifically for the research being carried out and thus few were awarded two marks. There was significant confusion in responses illustrating a lack of understanding surrounding primary data as many students were incorrectly focusing on who was collecting the data.

(e) Some good responses were provided by the minority of students who understood why the mean was the most appropriate value and could refer to it as being more representative. Of these correct responses, only around half went on to clearly explain why it was the most appropriate and thus few students achieved full marks. It was evident that the vast majority did not understand why the mean would be the

most appropriate measure and stated incorrect reasons, such as ‘the mean is the most appropriate as it shows the overall average’ and so scored no marks.

(f) This question was answered very well with the vast majority of students producing correct mean values for both tasks and thus being awarded full marks. A small minority were awarded partial credit for stating an incorrect answer but with correct workings.

(g) It was very clear from responses to this question that most students did not understand the purpose of counterbalancing and so scored zero marks. Most responses incorrectly stated that the purpose was to ‘reduce, eliminate or prevent’ order effects. Students seemed to be unaware that the purpose of counterbalancing is to control or equally distribute the impact of order effects. Other responses were tailored towards an explanation of how you would do counterbalancing rather than discussing the purpose. There were also a significant number of blank spaces suggesting that some students did not have the knowledge to enable them to attempt this question.

(h) There were some excellent answers to this question with appropriate extraneous variables (EVs) identified and elaborated. However, many students failed to achieve the third mark as they did not explicitly state what effect the EV would have on the time taken to complete the spot the difference task (DV). Many correct responses referred to the type or volume of music being played, or the difficulty of the spot the difference tasks. However, a minority of students failed to achieve any marks as their responses referred to participant variables, such as poor eyesight or IQ, which would have been accounted for through the use of a repeated measures design. Additionally, some answers were vague and referred to ‘environmental conditions’ as an EV without explicitly referring to the change in these conditions between the tasks and thus limited the marks awarded.

(i) Most students produced some limited explanation of how the follow up study improved upon the first study and thus achieved level 1. However, few went on to give a detailed explanation of this improvement, often only focussing on either why the follow up study was positive or why the original study was negative without comparison. Those students who achieved level 2 tended to focus on the follow-up study being more objective, a more scientific measurement or allowing cause and effect to be established, and then compared this with the relevant opposing weakness of the original study (more subjective, less scientific and only opinion so does not establish cause and effect). A minority of students mistakenly assumed that the first study generated qualitative data as it was self-report and so focused on the limitation of that versus the quantitative element of the follow-up study, and so scored no marks.

Q3.

(a) For full credit students needed to apply their knowledge of a problem associated with volunteers in this particular study. Many only referred to bias or representativeness and did not relate their responses to type of offender/ possible effect on therapy or on anger scores.

(b) There were many different suggestions given for random allocation and all suggestions that were practical and would result in the required groups were credited. However, many responses never went beyond the idea of putting names in a hat or into a random number generator, meaning that all the researcher would have is names in a hat or in a generator. Students should be able to describe accurately what they would do to reach the final stage of having the required number of people in the required groups. It was very clear that many students had

STEMUP.CO.UK

no idea of what they would have to instruct the computer or random number generator to do in order to reach this stage.

(c) Full marks were rarely achieved as most responses failed to show understanding of what was being measured in this study (the DV) which was a **reduction** in anger scores not anger scores or anger. Some students thought the study was looking for a difference between the start and end of therapy rather than between two different therapies.

(d) Many students failed to write enough about the data. Instead, they seemed to be providing answers to a question that required justification of one chosen suggestion, often just referring to the greater reduction in Group 2/Therapy B indicating that it was more effective. There were many points that could have been extrapolated from the data in the figure.

(e) Many students overlooked the most significant improvement to be derived from using a matched pairs design, ie the matching on pre-therapy anger scores. These students resorted to explaining matched pairs in a generic way without reference to the criteria which could be used for matching. Instead they often suggested IQ, age and sex as the crucial points, rather than pre-therapy anger scores. While credit was available for demonstrating an understanding of matched pairs design these responses did have to be contextualised and unfortunately many answers were generic.

(f) It was surprising that a number of students missed the second part of this question. An appropriate ethical issue was outlined (often fully) but then there was a failure to explain how this issue could have been dealt with. The importance of keeping the whole question in mind is highlighted in this case. Some responses presented a muddle of possible issues and markers had to try to match an appropriate issue with its appropriate solution. More worrying were the responses where one issue was described and a different issue was dealt with.

(g) Answers were often disappointing as students seemed unaware that median values are a form of average and stated that a problem was the median is not an average. Others made vague statements about how the median doesn't use all the values, but it would be difficult to find a median without the other values.

(h) This was generally answered well by the majority of students. However, there were quite a number of responses where knowledge was evident but marks were not obtained because the answer was not contextualised.

(i) It was again difficult to give full credit to responses in some cases as they were often vague or undeveloped so the information about the strength of using questionnaires as a technique for data collection often lacked clarity.

Q8.

(a) There were some excellent answers referring to checking the procedure of the words / words and picture study and making changes if necessary. There appeared to be three main areas of misunderstanding; that a pilot study should be used to see if the hypothesis was supported, to see whether a different experimental design should be used or to check participants. In relation to the final point there seemed to be an incorrect assumption by some students that the same participants would be used in the pilot study as the experiment.

(b) Students were required to operationalise the IV and the DV in order to score full marks. "There will be a difference in the number of words correctly recalled

when words are presented with pictures and without pictures" is operationalised. "There will be a difference in words recalled in condition one and condition two" is not operationalised and would therefore attract only one of the two marks available. Many answers referred to a correlation, using the term 'relationship' or 'link' when they meant a difference. A few students produced a null hypothesis in response to this question.

(c) The comparison of experimental designs proved difficult for some students. Students who considered the designs in relation to the stem material were more likely to produce an appropriate answer. Those who just repeated rote learnt problems of repeated measures designs were less successful. For example, in the context of this study, (learning 20 words / 20 words with pictures) fatigue was unlikely, given the minimal requirements of the tasks.

(d) Students were required to show some understanding of the outcome of the experiment. This could be achieved by reference to the median and range in such a way that it was apparent that the students understood the terms, eg referring to the average or spread of the scores. Students could also show understanding by drawing an inference from the figures. Simply repeating the contents of the table showed no understanding.

Q9.

(a) Hypothesis writing continues to be a problematic for many students, despite the requirement to do this at AS level. Less than a third students achieved the full 2 marks available and a further third scored no marks at all, having mistakenly written a directional hypothesis or one which predicted a difference between happiness scores and intelligence scores. Many responses were lacking in clarity or failed to operationalise the variables sufficiently. The best answers were concisely and clearly worded such as "There is a correlation (relationship) between pupils' scores on a standardised intelligence test and their scores on a questionnaire measuring happiness", which achieved the full 2 marks.

(b) Most students were able to identify an appropriate alternative method to collect data about happiness, the most popular choices being interviews, observations and diary studies. Some were able to provide a clear explanation of why this would be better, but weaker students became side-tracked into describing the possible method in detail (i.e. observational categories that might be used) and lost focus on the question of comparison. Better students were able to refer to precise advantages of their chosen method over questionnaires which were contextualised in relation to measuring happiness.

(c) This question required students to give two reasons why Spearman's rho was used to analyse the data. Almost all students were able to accurately identify one reason, which is encouraging!

(d) Weaker students still struggle to interpret critical and obtained values appropriately. About a third of students managed to pick up the full 3 marks. Most of the remainder were able to say the result was significant, gaining one mark but went on to select the incorrect critical value from the table. A small number erroneously compared the correlation coefficient (0.42) with 0.5 demonstrating misunderstanding throughout the entire reasoning.

(e) This question required students to interpret a further correlation coefficient (which was statistically insignificant) and put both pieces of information together to draw an overall conclusion to the reported study. This proved challenging and less than 5 % of answers achieved the full 4 marks here. Many students were able to identify that

the scores demonstrated different kinds of relationships (positive and negative) but were unable to take this further and think about possible explanations. The better answers focused on the inability to establish cause in correlational research and the role played by other variables, in this case age. Some made use of their knowledge about reliability which was creditworthy.

Q10.

(a) Students who had carefully read the stem material noted the reference to the age of the participants in the study being estimated by the researcher. They were thus more likely to recognise that one aim was to investigate the effect of age on eyewitness testimony.

(b) Students usually responded by referring to the benefits in terms of validity or fewer demand characteristics. Not all students elaborated their point in order to achieve a second mark.

(c) This question was poorly answered because many students could not accurately identify opportunity sampling and they confused this technique with random sampling. It is important for students to grasp that random refers to where everyone in the sampling frame has an equal chance of being selected and this does not apply in opportunity sampling. The word random has a different meaning in psychology compared with everyday usage.

(d) This was generally answered well with many answers showing a good understanding of extraneous variables. Students were very inventive referring to the effect of both situational (weather, noise levels) and participant (in a rush, alcohol consumption, mental illness) variables.

Q11.

(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

STEMUP.CO.UK

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.

Q12.

(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

STEMUP.CO.UK

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.

Q13.

(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.

Q14.

(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

STEMUP.CO.UK

(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).

Q15.

(a) It was disappointing to see such poor understanding of how to write a non-directional hypothesis for an experiment. Common pitfalls were to write a directional hypothesis, a null hypothesis or a correlational hypothesis. Some candidates did not appear to understand the conventions of hypothesis writing at all and simply wrote a statement such as: 'They wanted to test memory span' or 'The capacity of STM is 7+ or – 2'.

(b) Some candidates clearly understood why the researchers had chosen an independent groups design and were able to offer a full answer worth 2 marks. However, this question was not answered well by many candidates. Candidates needed to apply their knowledge to answer this question but, unfortunately, many candidates simply gave a rote–learned advantage of an independent design without thinking about the context of this particular experiment where the problem of order effects or demand characteristics did not apply.

(c) Another straightforward question generally answered correctly.

(d) There were some full-mark answers where candidates demonstrated good understanding and were able to comment on several aspects of the graph. However, many candidates were only able to make a single point ie that adults generally had better recall than children and then 'padded out' their answer with repetition of the same point. A common pitfall was for candidates to state that 6 was the maximum number of digits recalled by children (or 7 by adults) instead of realising that this was the modal score. This was surprising as they had often successfully identified

STEMUP.CO.UK

the modes in (c). Few candidates seemed to understand how to measure the range by looking at the information on the graph. It appears that candidates could benefit from more practical experience of interpreting graphs

(e) Most candidates answered this well and were able to refer to earlier research by Miller or Jacobs which had demonstrated the limited capacity of STM (ie 7+/– 2 items or a range between 5 and 9 items). A minority confused capacity with duration and tried to answer in terms of the Peterson and Peterson study.

Q16.

(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.

STEMUP.CO.UK