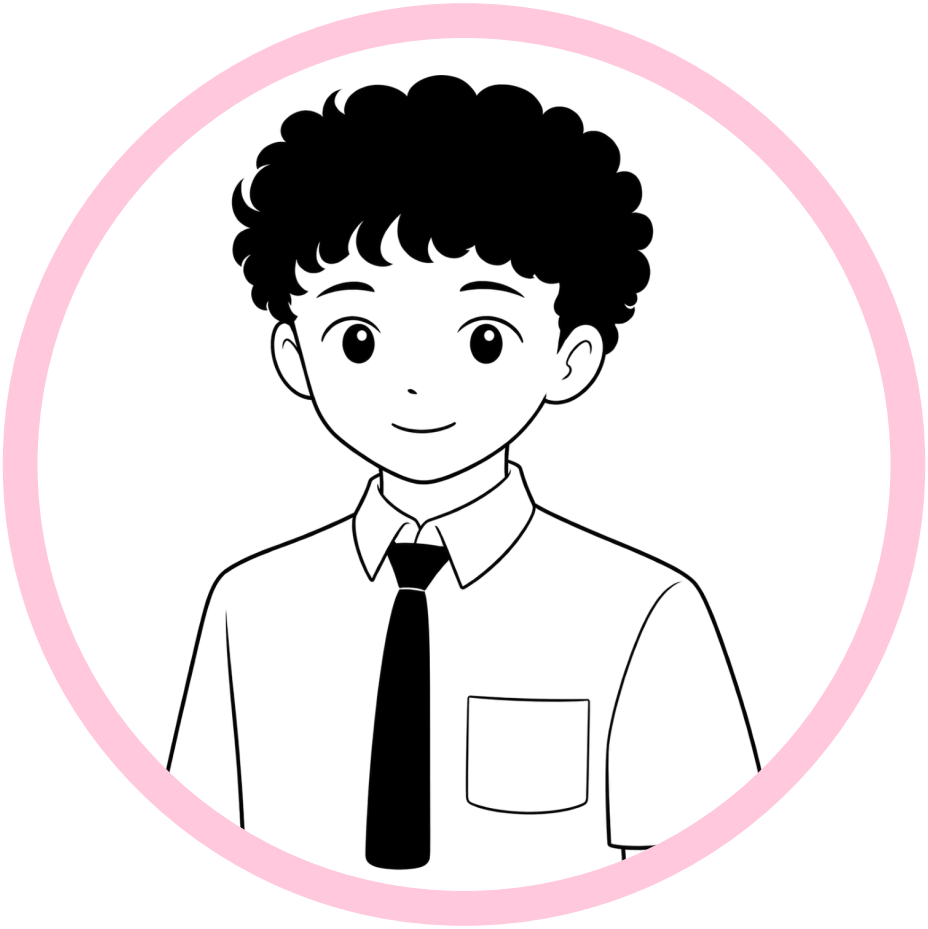
AQA A-LEVEL PSYCHOLOGY · 7182

PSYCHOLOGY



MARK SCHEME

3.2.3 CORRELATIONS

SCORE	PERCENTAGE
/172	%

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.

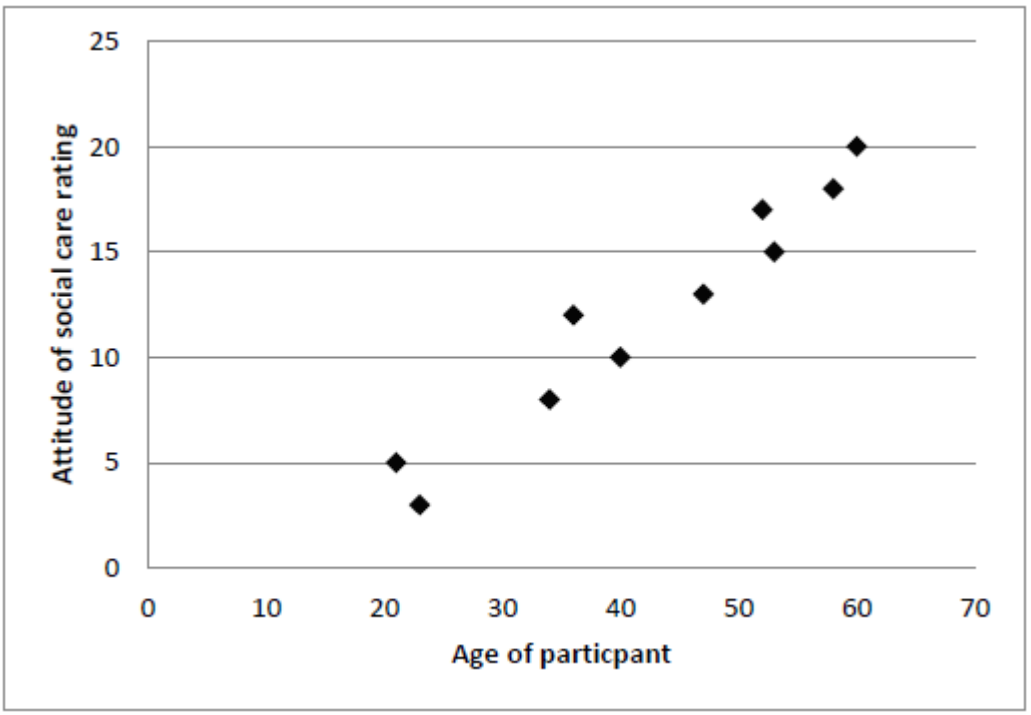
Where to find the originals

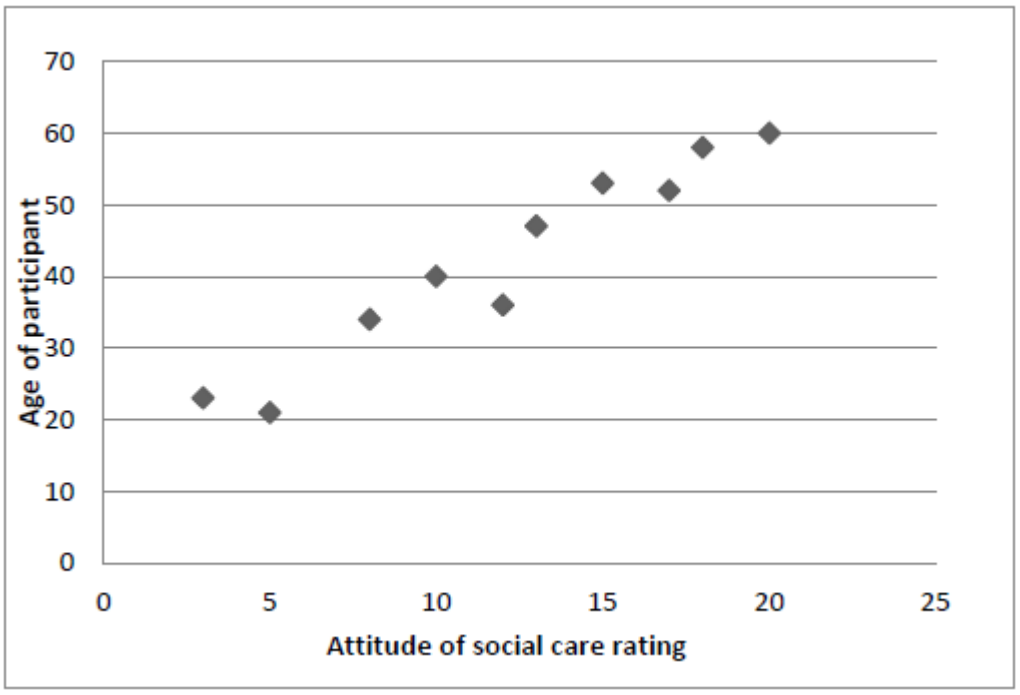
AQA publishes its past papers, mark schemes and examiner reports free of charge. Always check the current specification at aqa.org.uk before you revise, because content and wording can change between series.

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 = 3]**





3 marks for the following points:

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

STEMUP.CO.UK

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

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

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

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

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

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

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

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

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

Credit other relevant procedures.

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

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

Possible content:
- Closed questions would present participants with options for their response so the researchers would be able to collate and display the

STEMUP.CO.UK

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

Credit other relevant procedures.

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

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

(g) **[AO1 = 2 AO2 = 2]**

AO2
1 mark: the responses to the open questions in the interview constitute qualitative data.
Plus
1 mark: the attitudes ratings **AND / OR** the collated responses to the closed questions in the interview constitute quantitative data.

AO1
1 mark for an explanation of how the responses to the open questions is qualitative data ie is non-numeric / descriptive / retains detail of actions / thoughts / feelings.
Plus
1 mark for an explanation of how the ratings / collated responses to closed questions is quantitative data ie numerical such as a score / behaviour is represented in the form of a score on a scale.

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

2 marks for each explanation of how the chosen ethical issue could be dealt with.

1 mark for a brief muddled explanation.
2 marks for a clear explanation.

Consent – to be part of what is in essence two studies. Participants should be forewarned – a briefing.

Protection from harm – at the end of participation all will have to be fully aware that they were rated for their social care interest and a low score might indicate they are ‘uncaring’. They may wish to withdraw their data.

STEMUP.CO.UK

Right to withdraw – being made aware that they can at any time stop participating and at the end of their participation they can withdraw detail of their behaviour in the research.

The explanation must demonstrate an appreciation that people should be dealt with, with respect and competence.

Credit other relevant ethical issues.

Q2.

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

1 mark – correlation

Plus:

1 mark – she is investigating the (numerical) relationship between two co-variables

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

1 mark – scattergram

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

1 mark – 10 children

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

1 mark – pattern shows a (strong) negative correlation

Plus:

1 mark – suggesting that the more years are spent in an institution, the lower the language ability (or the fewer years spent in an institution), the better the language ability

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

1 mark for correct answer 14 (or 15)

Plus:

1 mark for correct workings (18 − 4) or (18 − 4 + 1)

Q3.

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

Award **1 mark** for outline of a positive correlation / as one variable increases, so does the other.

and

1 mark – It would not be appropriate because correlation only shows a relationship between the two variables, not cause.

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

STEMUP.CO.UK

2 marks for a clear, coherent outline of an appropriate way of dealing with a relevant ethical issue.

1 mark for a vague / muddled or incomplete outline of an appropriate way of dealing with a relevant ethical issue.

0 marks if the ethical issue is irrelevant or the way is inappropriate.

Relevant issues would include:

- asking for consent
- preferably written and on more than one occasion
- offering the right to withdraw from the study
- maintaining confidentiality
- treating with respect.

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

Award **one mark** for each of the relevant points below:

- the same participants would complete the sleep questionnaire on more than one occasion
- each participants' scores from the first occasion should be correlated with his / her results from the later occasion to be shown on a scattergraph to describe the correlation, with scores from the first test plotted on one axis and the scores from the second test plotted on the other axis
- the strength of the correlation should then be assessed using either a Spearman's rho test (or a Pearson's r test)
- the degree of reliability is then determined by comparing the correlation with the statistical table to determine the extent of correlation – there should be a (strong) positive correlation between the two sets of scores.

Q4.

(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]**

STEMUP.CO.UK

1 mark for stating scattergraph or scattergram.

Plus

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

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

Possible content

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

Accept other relevant comments

(e) **[AO2 = 2 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

STEMUP.CO.UK

- 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]**

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.

Q5.

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.

STEMUP.CO.UK

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

One mark for answers either:

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

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

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

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

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

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

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

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

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

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

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

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

Up to three marks for a discussion of reasons for correlation studies rather than experiments when investigating behaviour.
Likely answers: unethical / impossible to manipulate these variables (reading

STEMUP.CO.UK

and empathy in children) to investigate cause and effect; impractical to sometimes do an experiment; may discover a link between two existing variables which might suggest future research ideas; interested in relationships **rather than** a causal explanation.
Accept comparison with the experimental approach.
For full marks, the answer must be coherent and applied to this study.
Maximum of two marks for general answers not applied to this study.

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

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

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

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

Mark bands

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

STEMUP.CO.UK

Q6.

AO3 = 2

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

Q7.

AO3 = 6

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

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

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) **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

STEMUP.CO.UK

difference / link / association / connection.

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

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

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

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

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

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

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

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

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

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

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

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

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

STEMUP.CO.UK

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.

Q9.

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

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

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

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

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

- This involved subjective judgement on the part of the teacher who rates the students’ musical ability. Her judgement may not reflect real differences in musical ability and is likely to differ from other people’s judgement and / or any absolute criteria for tunefulness.
- Lack of reliability in rating musical ability would compromise the validity of the measure.
- As the students can choose the song they will sing, the rating of ability could reflect the teacher liking / dislike of the song rather than the student’s ability.

STEMUP.CO.UK

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

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

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

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

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

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

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

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

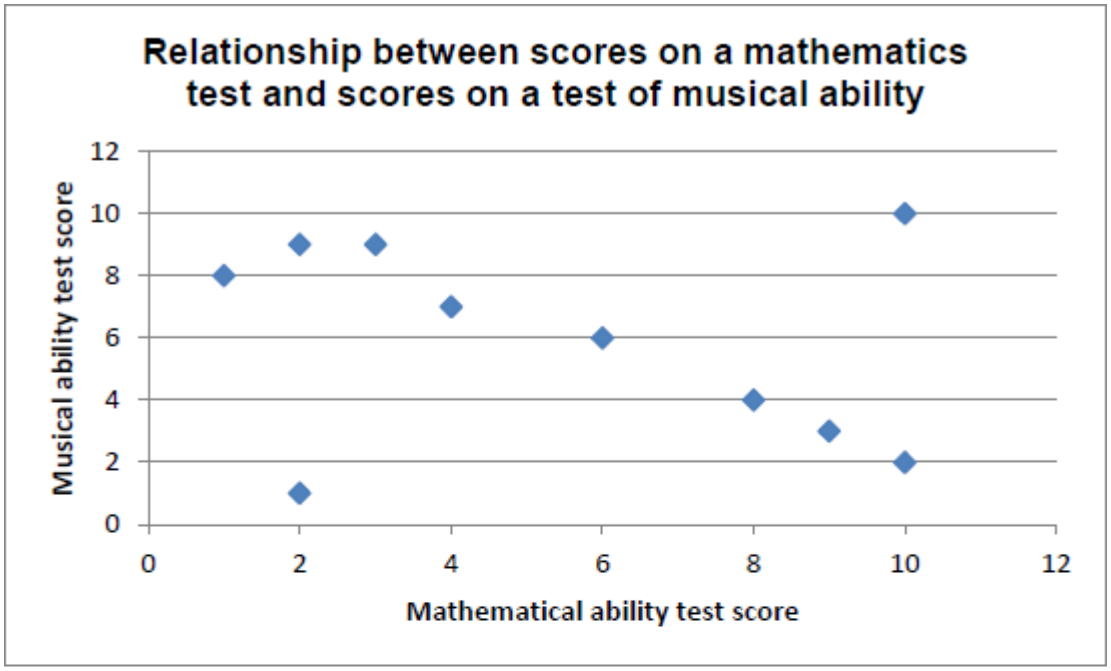
No credit for definition of a random sample.
1 mark for a brief or muddled reason (it is not biased).
2 marks for a reason that clearly points to an advantage of random sampling. This could be achieved through a comparison with another method (it is less likely to be biased than a volunteer sample).

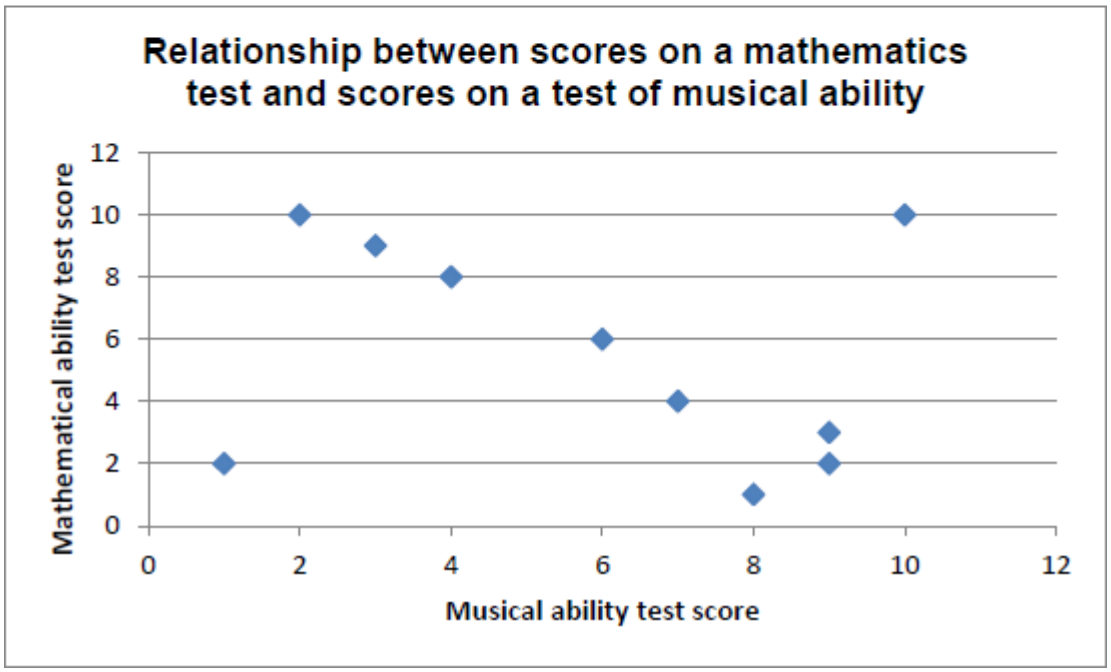
(e) **AO2 / AO3 = 3**

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

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

STEMUP.CO.UK





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

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

- The graph seems to show a negative correlation between mathematical and musical ability.
- This means that high scorers in mathematical ability tend to achieve low scores on musical ability and vice versa.
- The presence of two strong outliers, means that the actual correlation is very weak and closer to zero.
- Comment on the small sample size which limits the conclusions that could be drawn.

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

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

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

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

Design – 1 mark

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

Sampling – 2 marks

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

Procedure and assessment of musical ability – 4 marks

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

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

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

Debrief – 3 marks

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

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

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

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

Musical ability scores:

Participant number	Left handed	Right handed
1		

STEMUP.CO.UK

2		
3		

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

Q10.

AO3 = 4

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

1 mark for each of the following:

- Strength (it is a moderately strong / strong correlation)
- Direction (negative)
- Description of the relationship (the longer the treatment the lower the depression score)
- Indication of plateau / change in direction.

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:

- 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

STEMUP.CO.UK

- 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.

STEMUP.CO.UK

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 = 1**

10 children.

(b) **AO3 = 2**

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

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

(c) **AO3 = 3**

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

1 mark eg correlation doesn't prove cause or simply stating that aggression may be a result of other factors.
2 further marks for accurate elaboration as above.
Credit any explanation that focuses on causal inference.

STEMUP.CO.UK

An alternative approach is to argue that some research contradicts this claim. 1 mark for there is contradictory evidence and further marks for accurate reference to relevant research.
0 marks for reference to media exaggeration or individual children.

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.

AO3 = 2

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

Q14.

AO3 = 4

Strength: can study relationships between variables that occur naturally, eg stress from exams and students getting ill. Can measure things that cannot be measured experimentally. Can suggest trends that can lead to experiments.

Weakness: it is not possible to say that one thing causes another. Just because there is a correlation between stress and the immune system, it does not mean that stress directly caused the immune system to become less effective: there may be another variable connecting the two.

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

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 = 1**

Identification of a positive correlation, +, +ve.

(b) **AO3 = 4**

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

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

STEMUP.CO.UK

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

Q5.

(a) The majority of students gained the mark for this question by outlining the relationship. A small number of students simply stated ‘positive correlation’ which did not gain the mark as reference to the variables was required for an outline.

(b) Most students were able to name a correct test (Spearman’s or Pearson’s) although not all could justify the choice of test with reference to the levels of measurement. Some answers simply stated ordinal or interval, which was not enough as it was not clear which variable was being referred to. One could argue that the empathy scale was not equal interval and therefore should be treated as ordinal data. However, the hours spent reading fiction per week could be treated as interval. Some suggested an incorrect test e.g. Wilcoxon or Chi-square.

(c) Most students gained at least one mark on this question as they could outline a type of validity. However, some just named a type which was not enough for even one mark. Where students gained both marks, they generally referred to either face or concurrent validity (although occasionally to predictive or criterion validity). For full marks, answers had to explain how the type of validity would have been implemented. Many failed to do this part of the question and thus gained only half marks. Some students confused validity and reliability, often referring to split-half or test-retest, and a noticeable number tried to answer the question with reference to pilot studies.

(d) Most students were able to identify a limitation of the study and answered the question well. The majority of the problems identified referred to the sample. A number of answers referred to ethical issues and therefore gained no credit.

(e) Answers which simply stated that correlation looks for a relationship and experiments investigate differences (or similar) gained only one mark. Students that did access further marks were able to explain issues around ethics and manipulation of variables in such a study.

(f) This question was answered well by those students who had clearly had practice at designing and implementing their own investigations. It was pleasing to note that some students gained full marks on this question and it was evident that some schools and colleges had prepared students well. Students who did not gain high marks on this question usually:

- wrote very brief answers or ran out of time
- did not address all 5 bullet points. This was a shame because a number of the points were often addressed comprehensively but then students failed to address certain sections, in particular materials. The best answers addressed each point in turn in a structured and comprehensive manner
- provided little or no justification. The question specifically asked for justification of design decisions, but some answers gave little or no justification, e.g. for design, sampling techniques etc. Without justification the question was not properly addressed and this led to lower marks.

Q6.

Most students were able to state correctly that the correlation was a positive one; as scores on a stress questionnaire increased, so did the number of days off work through illness. Credit was also given to any reference to strength (moderate or strong) as well as

to the flattening out of the graph.

While most students were able to consider the main weakness of correlations, few could correctly outline a strength. In fact many contradicted themselves by saying that a strength was the ability to show cause and effect on a graph and the weakness was that a causal relationship can never be assumed. Better answers referred to the strength as being the ability to study the relationship between variables that occur naturally; or to measure things that cannot be manipulated experimentally.

Q7.

Most students were able to state correctly that the correlation was a positive one; as scores on a stress questionnaire increased, so did the number of days off work through illness. Credit was also given to any reference to strength (moderate or strong) as well as to the flattening out of the graph.

While most students were able to consider the main weakness of correlations, few could correctly outline a strength. In fact many contradicted themselves by saying that a strength was the ability to show cause and effect on a graph and the weakness was that a causal relationship can never be assumed. Better answers referred to the strength as being the ability to study the relationship between variables that occur naturally; or to measure things that cannot be manipulated experimentally.

Q8.

(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.

Q9.

(a) Hypothesis writing continues to be problematic for many students, despite the requirement to do this at AS level. Around 40% of students achieved zero marks on this question , having mistakenly written a directional hypothesis or one which predicted a difference between mathematical ability and musical ability as opposed to a relationship. Many responses lacked clarity or failed to operationalise the variables sufficiently. The best answers were concisely and clearly worded such as “There is a correlation (relationship) between pupils scores on a test of mathematical ability and their scores on a test of musical ability”, which achieved the full 3 marks.

(b) This question was answered well, with most students scoring two or all three marks. Weaker students were able to spot the test was based on a subjective judgement and some also made the point that singing was a poor measure of all round musical ability. Stronger students identified the lack of control (different choices of song) and were able to link this appropriately to investigator bias. Some students also made the point that the test lacked validity as it had not been standardised.

(c) There was a broad range of answers to this question and about 40% of students achieved no marks at all. Some confused reliability with validity, suggesting various methods such as comparing the scores with another measure of maths ability. Few contextualised this by identifying alternate forms which would have been creditworthy. Others made reference to running a pilot study which received no marks.

The remaining 60% had some idea of ways of assessing reliability of the maths test, the most common methods being test-retest and split-half. Some used inter-rater reliability appropriately suggesting that two separate markers could be used for the maths test: others became sidetracked into assuming that the study was observational. Stronger students were able to explain two or three methods of checking reliability in reasonable detail.

(d) This straightforward question on a random sample caught out quite a few students. Most were able to achieve 1 mark by referring to the method as being likely to yield a more representative sample. The weakest students simply defined random sample and went no further.

(e) This question required students to draw a scatter graph to display the data. About half achieved all three marks here. Many students failed to gain full marks by inaccurate or missing labels or title. About one third of students drew an incorrect graph, the most common error being to draw a bar chart.

(f) Most students were able to make some commentary on the generally negative correlation shown in the graph and table. Better students noted the presence of two outliers which weakened the overall strength of the relationship and some commented on the impact of outliers in a small sample. A small number of students made a rough calculation of Rs which was impressive but unnecessary to gain full marks.

STEMUP.CO.UK

(g) This question had a range of answers from students that covered marks from 0-10. The mark scheme allowed students to argue for different ways of designing the experiment (independent measures or matched pairs) and of generating a sample (volunteer or random selection from the two groups) provided these were workable and justified. Some common errors included:

- suggesting an inappropriate design (repeated measures) which did not take account of the information relating to left and right handers
- suggesting a sampling method but not explaining how it would yield an appropriate sample of left and right handers
- assuming that a maths test also needed to be completed (ie incorrect IV)
- failing to provide any procedural information
- producing a debrief which was not suitable to be read out to participants
- providing standardised instructions and claiming they were a debrief.

Some schools and colleges had clearly prepared their students well and many showed an impressive understanding of experimental design. Others struggled with the question and/or, failed to read the instructions and therefore gained very few marks. Once again, advice to teachers is: to do practical work. It was clear that some students were very familiar with designing experiments and they had a strong advantage here.

(h) This question required students to follow through their design from (g) and give some indication of how the data would be recorded and analysed. Most managed to sketch an appropriate table to record data, although a few misread the question and produced a summary table. Some students were able to follow through their design / data type with an appropriate test which could have been Mann Whitney (independent design) or Wilxocon (matched pairs). Students who had collected nominal data or recorded data in nominal form were credited if they suggested chi square to analyse it.

Q10.

Students seem to struggle with the interpretation of graphs and many do not seem to understand the difference between what data show ie findings and what can be concluded from the data. The graph shows a negative correlation, which is moderate-strong; this shows that the more weeks of treatment the lower the depression score. Students could also comment on the apparent “plateau” or change in direction in the middle of treatment. Very few students considered the strength of the relationship and weaker students just read the data off the graph.

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 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.

(b) Candidates were required to identify and elaborate an appropriate methodology for measuring aggression. Many were able to do this.

STEMUP.CO.UK

Unethical answers such as ‘provoke the children and observe their response’, did not gain credit.

(c) By identifying the lack of causal inference in correlation and suggesting other relevant factors which may affect children’s aggression, some candidates obtained full marks. Another appropriate route was to provide evidence from research to create a counterclaim. Trying to answer by using the scattergram was not a successful strategy as the question asks about causal inference and graph shows only the relationship between factors.

Q13.

Most candidates offered accurate and detailed answers for this question.

Q14.

Most candidates could provide an accurate outline of the weakness of using correlations. They were rarely able to give a strength, instead giving a definition.

Q15.

(a) Most candidates were able to identify this as a positive correlation. There were very few incorrect answers.

(b) Candidates found it harder to identify a strength than a weakness. Good answers often explained the advantage when it would be unethical to manipulate variables. For the weakness most candidates focused on causality issues.