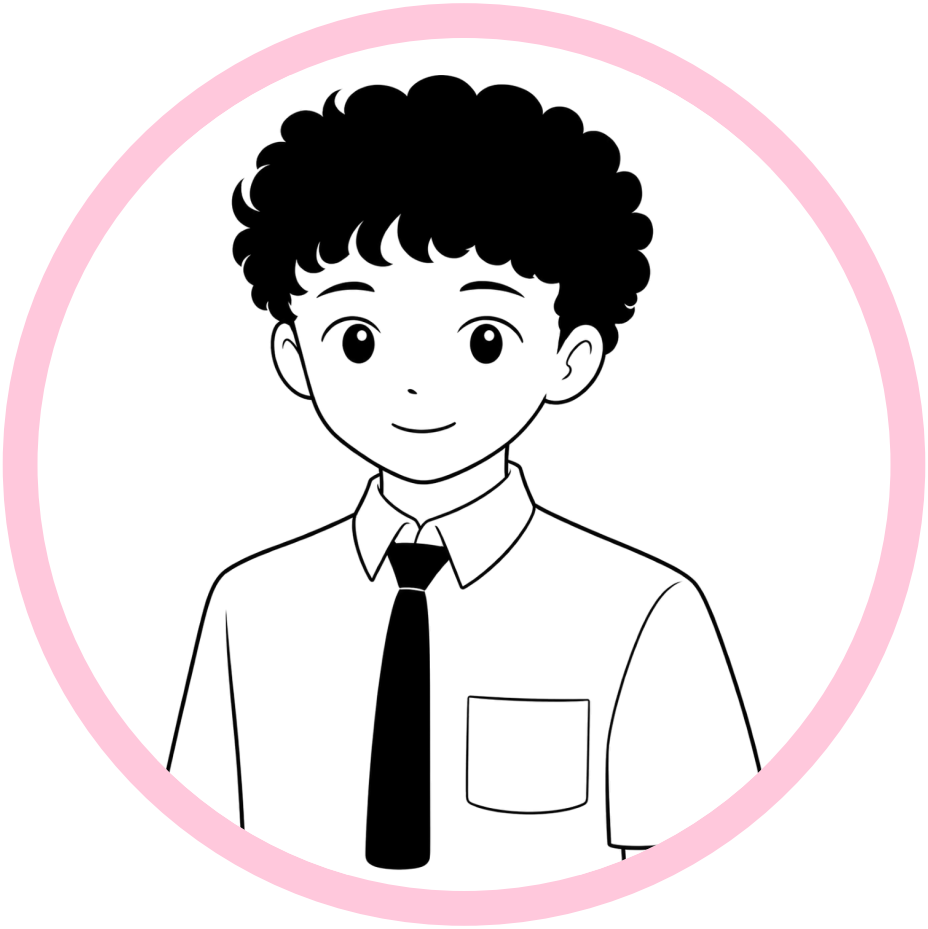
AQA A-LEVEL PSYCHOLOGY · 7182

PSYCHOLOGY



MARK SCHEME

3.2.3.1 REPORTING INVESTIGATIONS

SCORE	PERCENTAGE
/138	%

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.

[AO3 = 2]

2 marks for an accurate reference including essential elements in the accepted format as follows (punctuation can be ignored):

Duck, S. (1992) Human relationships. London. Sage.

1 mark for a reference that contains at least surname, date and title in reference format (surname must be first).

0 marks for a reference that does not meet the descriptor for 1 mark.

Q2.

[AO3 = 2]

2 marks for an accurate reference including essential elements in the accepted format as follows (punctuation can be ignored):

Bem, S. L. (1993) The lenses of gender: transforming the debate on sexual inequality. New Haven. Yale University Press.

1 mark for a reference that contains at least surname, date and title in reference format (surname must be first).

0 marks for a reference that does not meet the descriptor for 1 mark.

Q3.

[AO3 = 2]

2 marks for an accurate reference including essential elements in the accepted format as follows (punctuation can be ignored):

Selman, R. L. (1980) The growth of interpersonal understanding: developmental and clinical analysis. New York. Academic Press.

1 mark for a reference that contains at least surname, date and title in reference format (surname must be first).

0 marks for a reference that does not meet the descriptor for 1 mark.

Q4.

(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)

STEMUP.CO.UK

(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'.

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

STEMUP.CO.UK

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 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

STEMUP.CO.UK

(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

STEMUP.CO.UK

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

Q5.

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

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

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

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

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

- This involved subjective judgement on the part of the teacher who rates the students' musical ability. Her judgement may not reflect real differences in musical ability and is likely to differ from other people's judgement and / or any absolute criteria for tunefulness.
- Lack of reliability in rating musical ability would compromise the validity of the measure.
- As the students can choose the song they will sing, the rating of ability could reflect the teacher liking / dislike of the song rather than the student's ability.
- The rating may be invalid as the students selected songs which varied in difficulty so the tunefulness reflected the difficulty of the song not the students' ability.
- Operationalising musical ability as tuneful singing is a very narrow measure. Someone can have musical ability such as playing an instrument which would not be reflected by this measure.

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

STEMUP.CO.UK

(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.

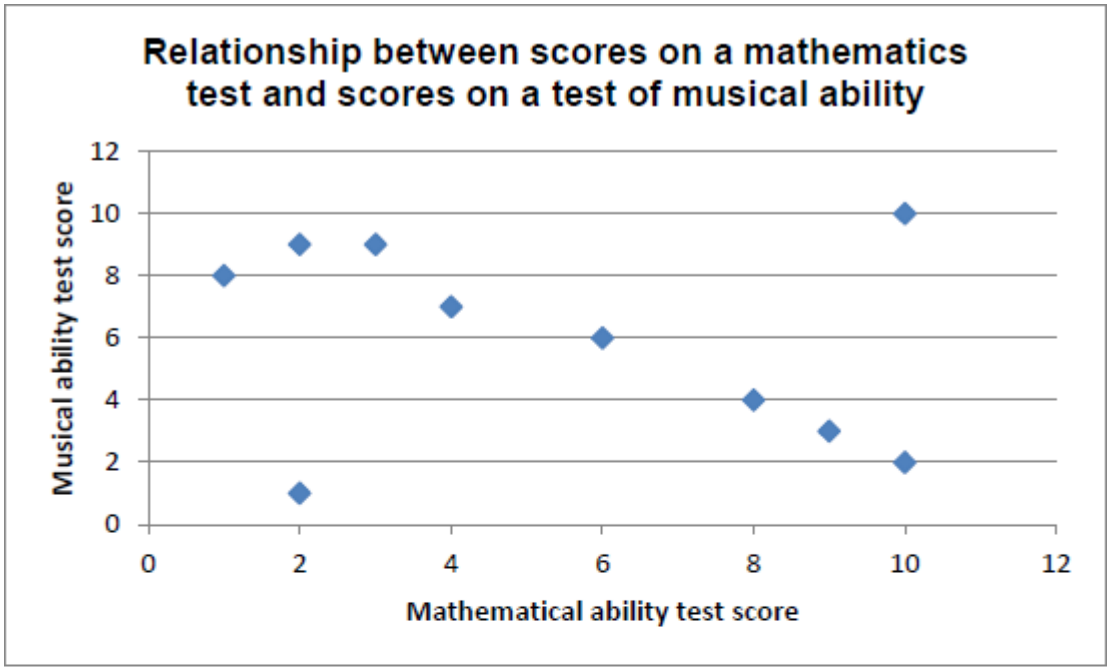
No credit for definition of a random sample.
1 mark for a brief or muddled reason (it is not biased).
2 marks for a reason that clearly points to an advantage of random sampling. This could be achieved through a comparison with another method (it is less likely to be biased than a volunteer sample).

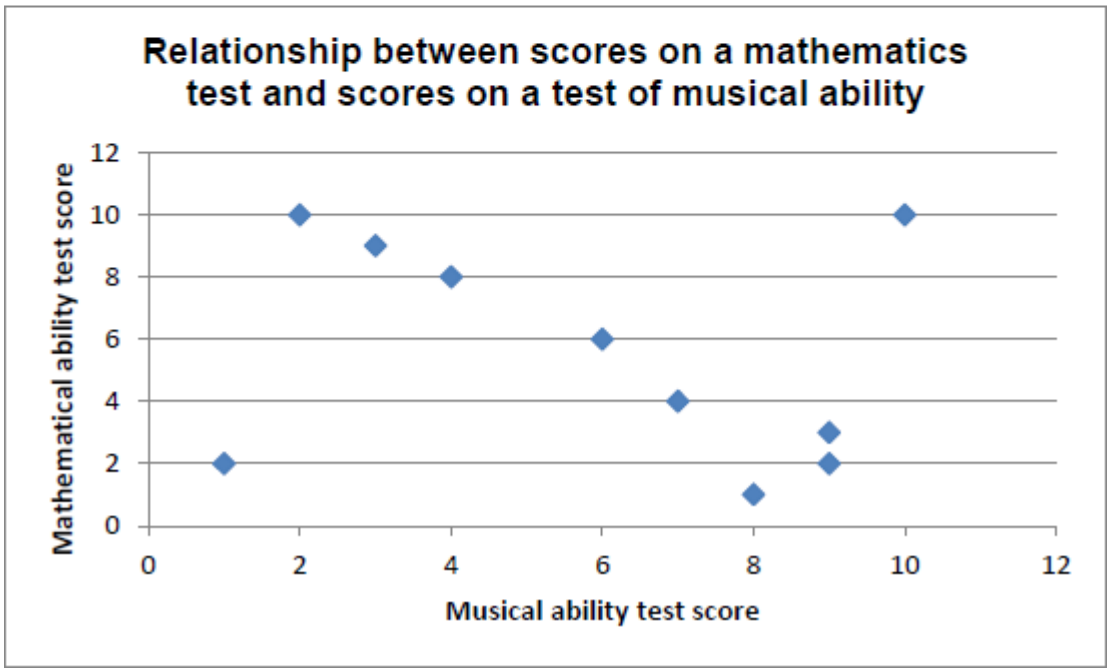
(e) **AO2 / AO3 = 3**

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.

STEMUP.CO.UK

- 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.

Q6.

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**

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

STEMUP.CO.UK

measures = 1 mark.
Answers such as ‘truth’ or ‘whether it is true’ ‘legitimacy’ or ‘accuracy’ = 0 marks.

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

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

- to help decide what colour to pick for an educational facility
- to help decide what colour enhances persuasion in a consumption context
- to help decided what colour enhances creativity in a new product design process.

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

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

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

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

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

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

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

The instructions should be clear and succinct. They must:

- explain the procedures of this study relevant to participants
- include a check of understanding of instructions.

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

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

AO3 Mark Bands

STEMUP.CO.UK

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.

Q7.

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**

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

STEMUP.CO.UK

- An appropriately labelled table = 2 marks

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

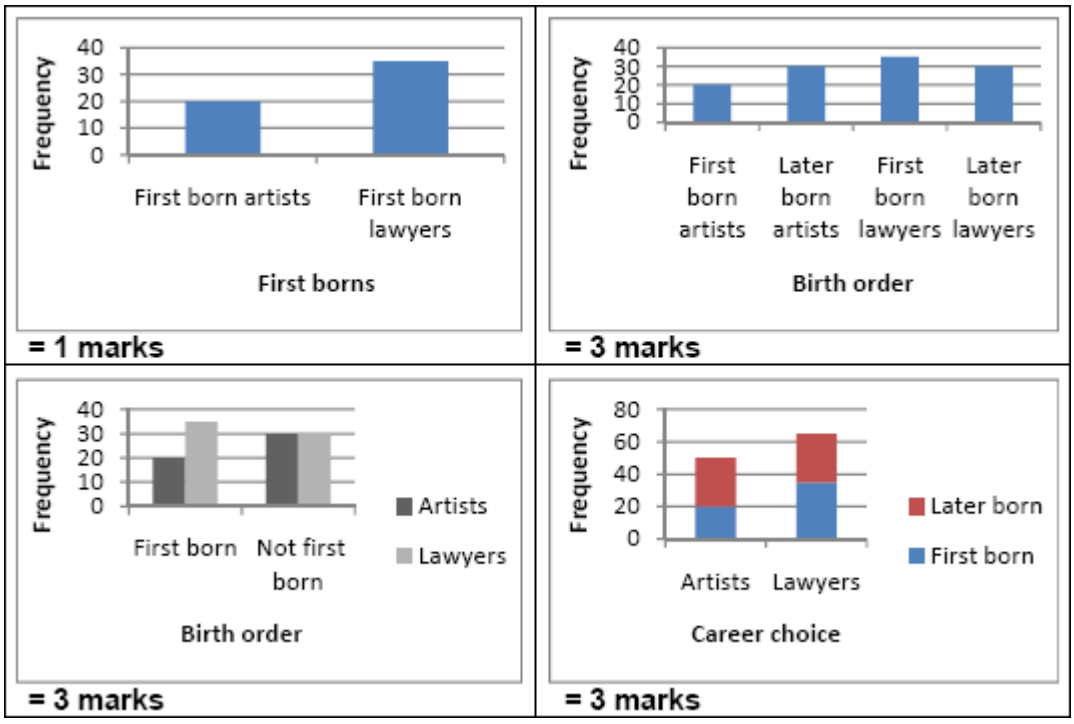
Table: Table to show the career choices of first born and non-first born children

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

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

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

For 2 marks, candidates display data as above but labels are missing or lack clarity.
For 1 mark, candidates graph the data supplied in the question relating to first born career choices only.



NB Labelled axes but no bars = 0 marks.

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

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

- data are independent
- level of measurement is nominal
- test of association / difference is required.

STEMUP.CO.UK

- 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.

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 / 3 = 2**

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

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

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

One mark for identification of a suitable test and 3 further marks for an appropriate

STEMUP.CO.UK

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

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

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

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

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

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

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

Possible reasons include:

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

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

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

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

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

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

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

Marks can be awarded for any appropriate advantages / disadvantages.

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

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

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

It should show awareness of ethical considerations, eg

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

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

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

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

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

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

STEMUP.CO.UK

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

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

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

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

There should be reasonable detail with regard to:

- design
- participants
- materials
- procedures.

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

Q1.

This question appeared to be unexpected despite ‘referencing’ appearing on the specification. Very rarely did responses receive full credit for the correct information in the correct order. The mark scheme was relaxed to enable a small proportion of responses to be awarded 1 mark for including the surname, date and title, with the surname first. Referencing is a crucial skill which students will find invaluable as they continue their studies beyond A-level and, whilst it was disappointing to see so few students gaining credit on this question, all students were equally affected because the same question appeared across all three options.

Q2.

This question appeared to be unexpected despite ‘referencing’ appearing on the specification. Very rarely did responses receive full credit for the correct information in the correct order. The mark scheme was relaxed to enable a small proportion of responses to be awarded 1 mark for including the surname, date and title, with the surname first. Referencing is a crucial skill which students will find invaluable as they continue their studies beyond A-level and, whilst it was disappointing to see so few students gaining credit on this question, all students were equally affected because the same question appeared across all three options.

Q3.

This question appeared to be unexpected despite ‘referencing’ appearing on the specification. Very rarely did responses receive full credit for the correct information in the correct order. The mark scheme was relaxed to enable a small proportion of responses to be awarded 1 mark for including the surname, date and title, with the surname first. Referencing is a crucial skill which students will find invaluable as they continue their studies beyond A-level and, whilst it was disappointing to see so few students gaining credit on this question, all students were equally affected because the same question appeared across all three options.

Q5.

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

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

(c) There was a broad range of answers to this question and about 40% of students

achieved no marks at all. Some confused reliability with validity, suggesting various methods such as comparing the scores with another measure of maths ability. Few contextualised this by identifying alternate forms which would have been creditworthy. Others made reference to running a pilot study which received no marks.

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

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

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

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

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

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

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

(h) This question required students to follow through their design from (g) and give

STEMUP.CO.UK

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.

Q6.

(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.

Q7.

(a) Hypothesis writing is still a problematic area for many candidates – despite the requirement to do this at AS level. Many candidates achieved zero marks on part (a), having mistakenly written a directional or a null hypothesis. Many responses

STEMUP.CO.UK

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).

Q8.

Perhaps surprisingly, many candidates achieved higher marks for this section than for their other two questions. Most candidates attempted all parts of the question although a significant minority did not complete part (h) suggesting that they might have run out of time. This was a pity since this part carried 10 marks.

(a) This was a straightforward question and many candidates accessed full marks but a surprising number were confused by median and range and some did not understand what the range indicated about the data.

(b) Many candidates were very well prepared and got full marks here but some wrote very confused answers showing little understanding eg ‘Spearman’s Rho because it was nominal data and repeated measures’.

(c) There was a centre effect here. Some candidates had a good understanding of Type 1 error while others had clearly never heard of it. Some understood what is meant by the term and offered a definition but were not able to apply their knowledge to answer the question.

(d) Candidates offered a wide range of answers, although some were a little bit too brief or poorly explained to get both marks. There was some confusion about what is meant by a placebo and some candidates offered two explanations which were essentially the same as one another. Other candidates offered factors which could apply equally to the treatment and non-treatment group. It is important in this kind of question to read the stem carefully. Some candidates said that the therapists might have been biased in favour of the treatment group, but the stem clearly states that

STEMUP.CO.UK

the therapist did not know who had been in which group.

(e) A lot of candidates missed the point that the advantage / disadvantage needed to be in comparison to interviews. Many candidates gave advantages/disadvantages that could apply equally well to both self-reports and interviews. This was acceptable only if the candidate made it clear ie '*People are less honest in a questionnaire*' was not credit-worthy because it could apply to both interviews and questionnaires. However, '*People are less honest in a questionnaire because they are anonymous and feel they can lie about themselves without being found out*. *In an interview where they are face-to-face with the interviewer, they might find it more difficult to lie*'.

(f) Some candidates wrote excellent consent forms containing both ethical and procedural information and expressed them in appropriate language. Some candidates had a very vague understanding of what needed to be included here, either only focusing on all the ethical issues (*you will have the right to withdraw your data, yourself etc*) with no mention of the procedures, or vice versa. Many adopted a rather inappropriate tone eg '*Once you have signed this form, you are committed to being in the study*' or '*You have to subject yourself to an interview*'. It was surprising to see that a few candidates seemed to think a consent form acted as some kind of legal disclaimer – '*you may suffer harm but if you sign this you can't sue us.*'

It was notable on this question that candidates who were able to express themselves clearly and succinctly were much more likely to access full marks. Many answers were so poorly constructed that the content was difficult to understand. Many switched confusingly between pronouns eg *They will have to have an interview. You can withdraw at any time. I agree to be part of this study*.

(g) There were some very muddled answers to this question. Candidates often didn't read the question carefully, and wrote something like '*Reliability means if you do the study again you will get similar results*' for their definition and then didn't know what to write for the next part of the question. Those candidates who explained it in terms of inter-rater reliability generally gained full marks. Some candidates did not read the question carefully and did not relate their answer to checking the scores in this particular study. Many candidates thought incorrectly that test-retest involved using different participants. Some candidates suggested split-half methods indicating a lack of thought about the question. Some candidates confused reliability with validity.

(h) Many candidates showed limited awareness of a conventional reporting style. While it was not necessary to divide the method section into sub-sections, this strategy might have helped candidates to include all the relevant details. Weaker answers made no mention of gender or eating disorders and simply repeated details from the stimulus material. Many candidates completely lost sight of the fact that gender differences were being investigated and suggested randomly allocating participants to groups. A lot of time was wasted in including aims / hypotheses and statistical analyses which do not form part of a method section. Better answers included appropriate detail of IV, DV, design, sampling method, materials / equipment and procedure which would have enabled replication to take place. As in (f), poor expression and grammatical errors often obscured meaning.

STEMUP.CO.UK