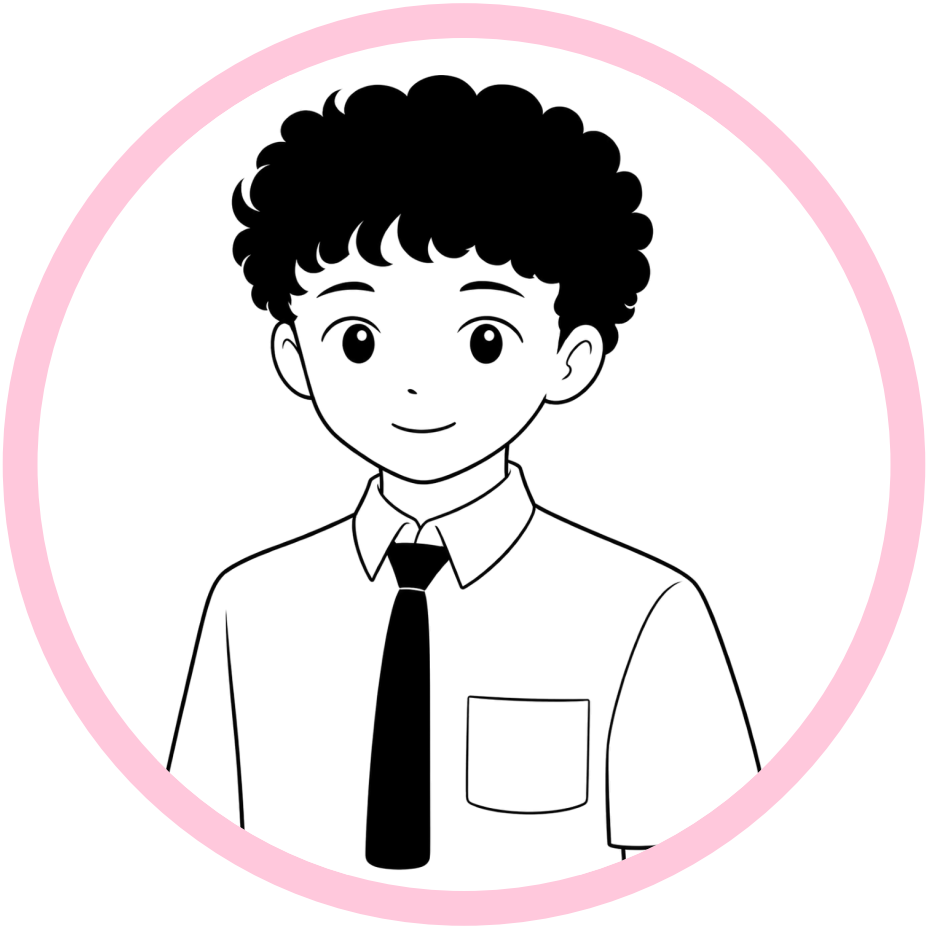
AQA A-LEVEL PSYCHOLOGY · 7182

PSYCHOLOGY



MARK SCHEME

3.2.3.1 QUESTIONNAIRES

SCORE	PERCENTAGE
/161	%

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 = 2]**

2 marks for a clear explanation of why it was important to match the students for this variable.

1 mark for a limited or muddled explanation, or for an answer without application.

Content:

- level of interest could affect their satisfaction scores so needs to be controlled; matching students' levels of interest in business studies keeps this participant variable constant.

2

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

1 mark for a suggestion of an appropriate variable – interest in using computers (e.g., for studying), past experience of using computers (e.g., for studying), disability issues, prior knowledge of this business studies topic.

Accept other appropriate suggestions.

1 mark for an explanation of how it might relate to satisfaction scores.

2

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

2 marks for a clear outline of one benefit of random allocation in this situation.

1 mark for a limited or muddled explanation, or for an answer without application.

Relevant points:

- removal of researcher bias in determining which participant has the most to least number of online hours
- increased validity as the researcher will not be biased in how students are allocated
- removal of participant bias as each participant has the same chance of having a particular number of online hours as any other participant.

2

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

3 marks for a clear and coherent explanation of how stratified sampling would improve the design.

2 marks for some explanation of how stratified sampling would improve the design.

1 mark for a muddled / limited explanation.

Content:

- stratified sampling ensures each identified sub group of the whole target

STEMUP.CO.UK

population of possible participants is represented in the sample so generalisation of the results to the target population is more valid than if the participants were just available at the time of asking
- stratified sampling is an improvement because people who are identified by opportunity are likely to be present at same time, doing similar things, possibly of similar personality types etc which may limit representativeness of sample.

Accept other appropriate explanation points.

3

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

1 mark for a question which would generate qualitative data: what did you like / dislike about the online / classroom lessons?

Accept any appropriate question or item that generates qualitative data.

1

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

2 marks for a clear and coherent explanation of a possible problem.

1 mark for a muddled / limited explanation.

Possible problems:

- analysis of qualitative data relies on interpretation by the researcher which could be inaccurate / biased / subjective
- qualitative data may not be easy to categorise / collate into a sensible number of answer types
- the researcher could be left with lots of individual responses that cannot be summarised.

Do not accept answers which simply identify a problem with no further appropriate elaboration e.g. ‘hard to analyse’, ‘time consuming’.

2

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

1 mark for an outline of an appropriate ethical issue that would be relevant when using a questionnaire: issues such as anonymity, protection from harm, treating with respect, parental consent.

Possible issues:

- the students should be treated with respect by explaining that their answers will be kept confidential
- the students must be assured that they can withdraw their answers if they want to
- the students must be told exactly what the researcher was trying to find out about the relationship between satisfaction and the way information was delivered.

Answer must be an outline and not just a name or term.

1

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

Award **1 mark** for each of the following points.

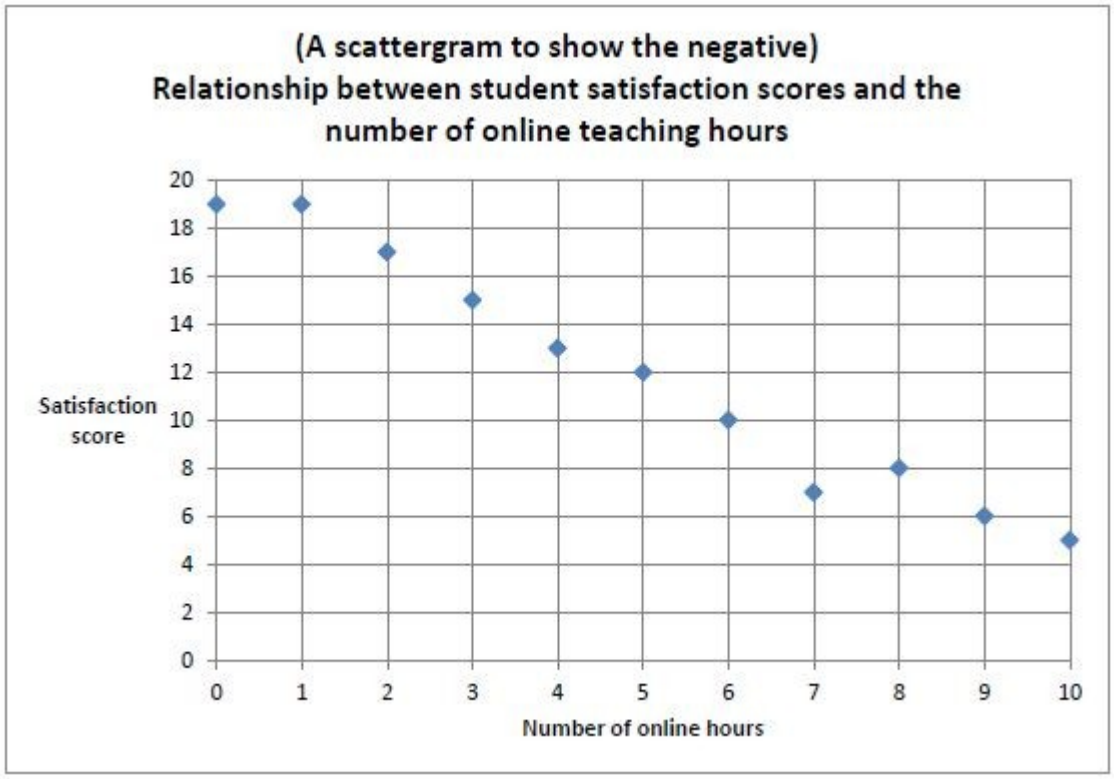
STEMUP.CO.UK

Content:

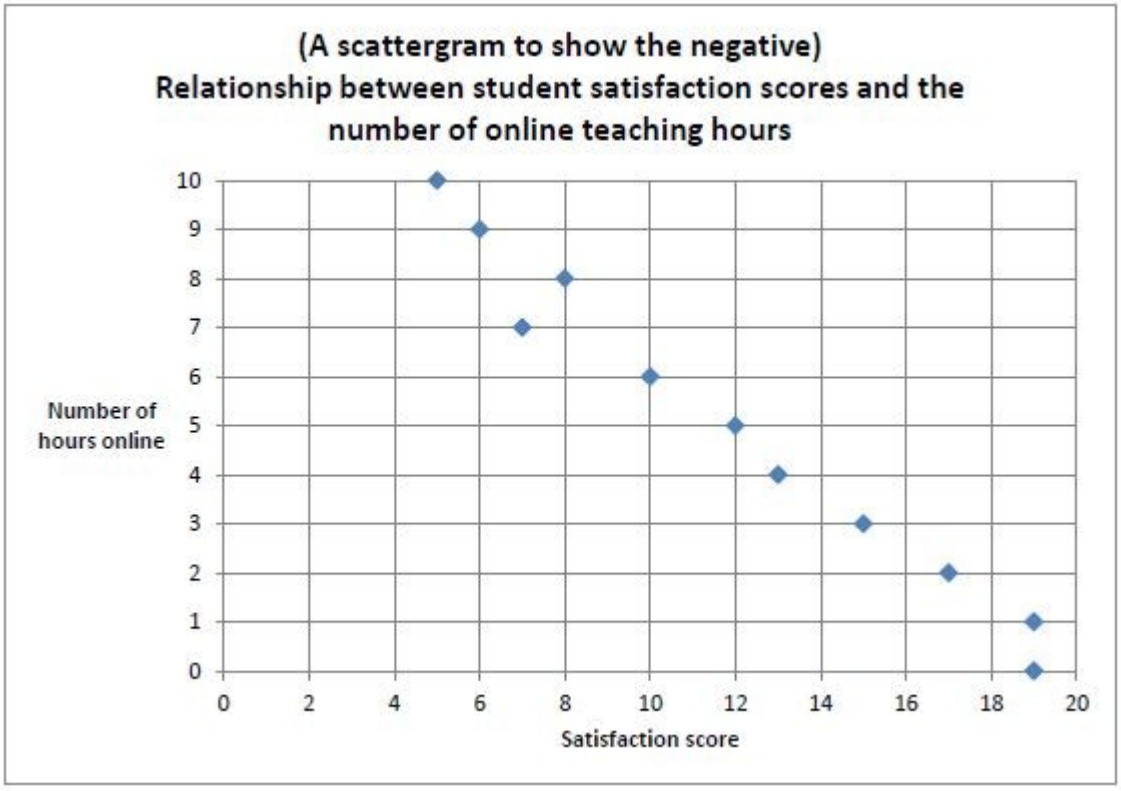
- appropriate title, including reference to (negative) relationship and the two variables
- axis 1 – number of online hours and suitable scale 0-10
- axis 2 – satisfaction score and suitable scale 0-20
- accurate plotting of 11 points on scattergram.

No penalty for drawing a line of best fit.

If display is not a scattergram award max 1 mark if title is appropriate.



STEMUP.CO.UK



4

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

1 mark for saying there is a negative relationship / correlation (between the number of online hours and satisfaction scores).

1 mark for further detail e.g. the **more** hours they have online the **less** they are satisfied with the course (or vice versa), however this does not indicate a cause and effect relationship.

Accept other valid conclusions.

2

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

1 mark for the mode = 19

1 mark for the median = 12

2

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

2 marks for a clear and coherent explanation.

1 mark for a muddled or limited explanation.

- The median is more representative than the mode because the mode is at the top end of the set of scores, whereas more of the scores are numerically closer to the median.

Accept other explanations, for example, in terms suitability in relation to ordinal data vs nominal data.

2

STEMUP.CO.UK

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

1 mark for B

1

[24]

Q2.

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

1 mark for **each** of the following points:

- excitement is a broad construct / complex behaviour / has many aspects
- heart rate is a narrow, biological / physical component / unit / element / factor in overall excitement.

2

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

For **each** outline award marks as follows:

2 marks for a clear and coherent outline with some elaboration.

1 mark for a limited / muddled outline.

Possible content:

- questionnaire measuring attitude – questions assess how the participant feels about going on the roller coaster
- rating scale of own excitement – the participant could rate their perceived level of excitement on a scale of 1 – 10
- observation of non-verbal behaviour whilst waiting in the queue; e.g. tallying appropriate behavioural categories
- interviews before / after the participant has been on the ride asking questions about how they felt / whether they enjoy other exciting occasions / events / how they cope with excitement, etc.

Do not credit measures of biological function.
Just naming a method e.g. questionnaire / interview / observation is not sufficient.

4

[6]

Q3.

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

1 mark for an accurate, relevant disadvantage of quantitative data eg lacks detail (which may decrease validity of findings) etc.

Plus

1 mark for clear application regarding the negative impact of using quantitative data in this study (such as rating a patient's improvement).

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

1 mark for an appropriate question which produces quantitative data, eg 'On average how many hours of physiotherapy does the patient receive per week?'

STEMUP.CO.UK

Plus

1 mark for an appropriate question which produces qualitative data, eg 'How does physiotherapy help your patients?'

Note: can accept relevant items that are not phrased as questions, eg describe… or rate....

Q4.

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

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

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

Credit other relevant disadvantages.

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

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

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

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

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

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

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

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

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

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

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

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

Possible points:

- Both therapies seem to be effective.
- The pre-therapy scores for both groups suggest that the offenders in Group 1 were more angry on average than those of Group 2 at the start of study.

STEMUP.CO.UK

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

Credit other relevant points.

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

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

Possible points:

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

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

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

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

- In this study the participants are all in custody so despite the suggestion that they are volunteers they may have been coerced in some way to participate, this could relate to issues of consent or stress or right to withdraw.
- Distress/psychological harm such as: finding out they have high-anger scores; the questionnaire might trigger more anger; finding out there are two therapies and finding out one might be better; therapies might cause more harm than good; the participants might have some expectation that the programme would

STEMUP.CO.UK

be effective and this might not be the case.
- Reducing the anger of the participants may be more beneficial for staff than the participants, manipulation of their behaviour for this reason could be unethical.

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

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

Credit other relevant information.

(g) **AO2 = 1**

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

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

(h) **AO2 = 2**

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

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

Possible points

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

Credit other relevant information.

(i) **AO3 = 2**

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

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

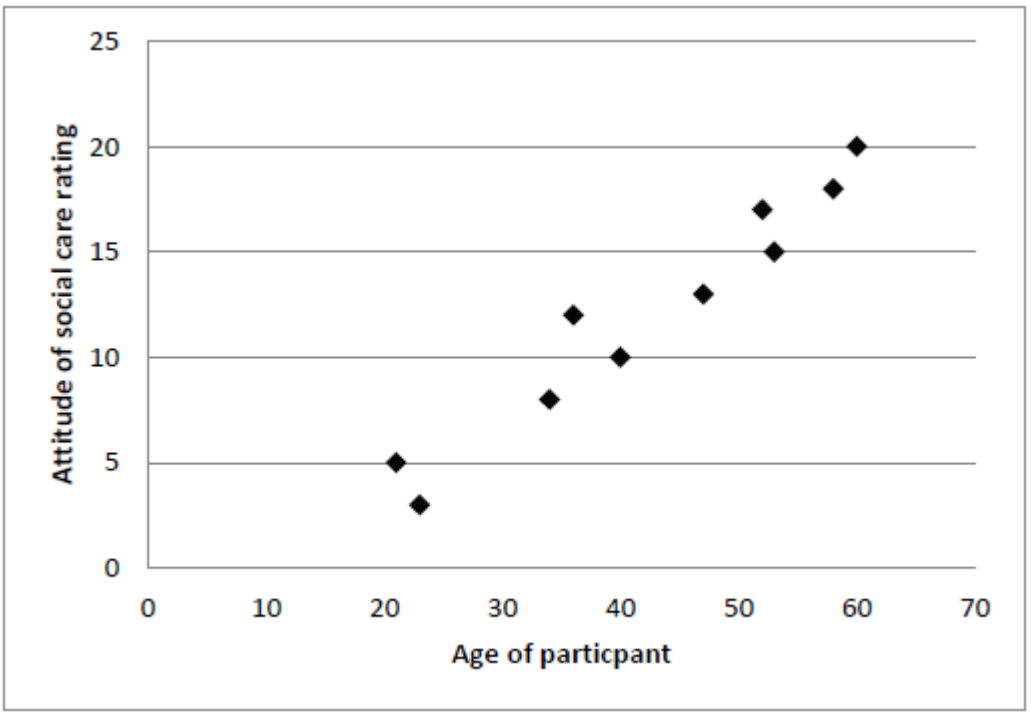
N.B If the strength discussed relates to a particular type of questionnaire, the type of questionnaire must be made clear for the second mark.

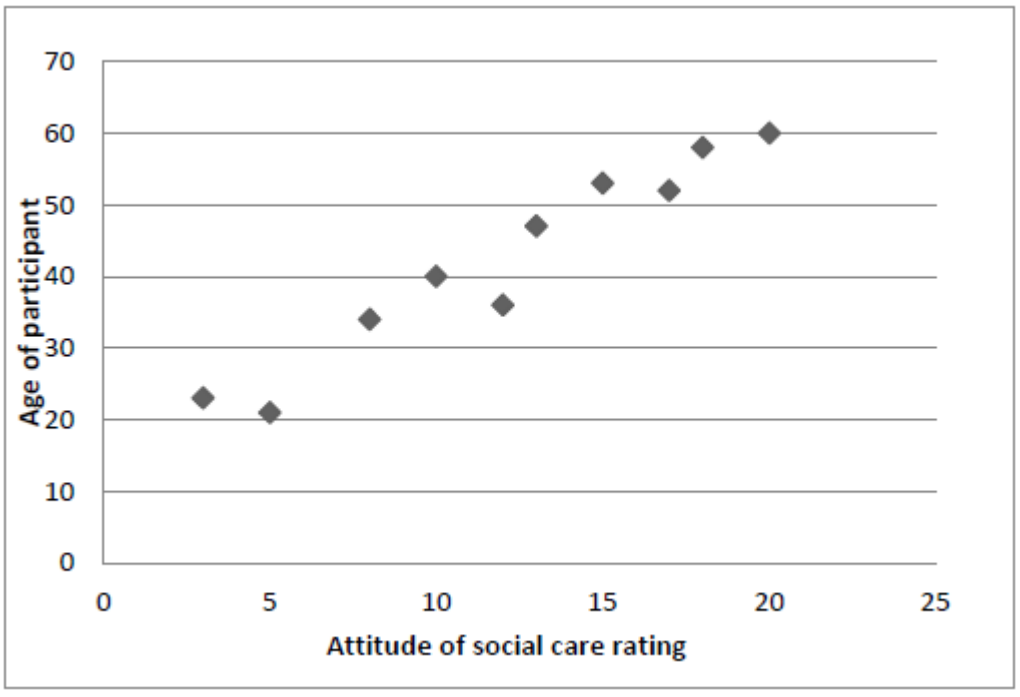
Credit other relevant strengths.

STEMUP.CO.UK

Q5.

(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.

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

2 marks for: there is a **positive relationship** between age and interest in

STEMUP.CO.UK

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 information collected easily.
- Closed questions make it easy to compare specific response to questions the researchers wanted answered – they can be sure there

STEMUP.CO.UK

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.

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

STEMUP.CO.UK

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.

Q6.

(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]**

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.

Q7.

Please note that the AOs for the new AQA Specification (Sept 2015 onwards) have

STEMUP.CO.UK

changed. Under the new Specification the following system of AOs applies:

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

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

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

(a) **AO3 = 4**

Suitable behavioural categories for investigating children's aggressive behaviour could be:–
pushing, hitting, biting, punching, swearing, etc.
Maximum 2 marks – 1 for each suitable behaviour category.
Candidates may suggest recording playground behaviour on CCTV for later analysis by ticking a box when a relevant behaviour is shown by the child. Alternatively the researcher could watch each child's behaviour in the playground and tick the box when each behaviour is shown. In this case where the researcher stands and whether the children know they are being observed would be relevant.
1 mark for a very brief or slightly muddled explanation e.g. use a tally chart
2nd mark for accurate elaboration

(b) **AO3 = 4**

There are no ethical issues named in the specification, so any potentially relevant issues in this research should be credited.
Although the psychologist would not be responsible for the behaviour of the children in the playground he might consider his responsibility if he saw that one of the children was being harmed.
Likely ethical issues include informed consent, right to withdraw, confidentiality or respect. Ways of dealing will depend on the issue selected.
There are different routes to achieving 4 marks depending on the ethical issue selected, but for full marks both the ethical issue and how the psychologist could have dealt with it should be clear.

AO3 Knowledge of research methods
4 marks Accurate and reasonably detailed Accurate and reasonably detailed answer that demonstrates sound understanding of one relevant ethical issue and how the psychologist could have dealt with this issue.
3 marks **Less detailed but generally accurate** answer that demonstrates relevant understanding of one relevant ethical issue and how the psychologist could have dealt with this issue. or **Accurate and reasonably detailed answer** that demonstrates sound understanding of one relevant ethical issue or how the psychologist could have

STEMUP.CO.UK

dealt with an issue.
2 marks Basic Basic answer that demonstrates some relevant understanding of one relevant ethical issue and / or how the psychologist could have dealt with an ethical issue, but lacks detail and may be muddled.
1 mark Very brief / flawed Very brief or flawed answer demonstrating very little understanding of a relevant ethical issue and / or how the psychologist could have dealt with an issue.
0 marks No creditworthy material.

(c) **AO3 = 4**

There are different routes to full marks in this question. Candidates explain one advantage in reasonable detail or more advantages in less detail. Advantages of using an interview rather than a questionnaire could include it would allow the interviewer to clarify questions and answers; it might be easier to see if participants were answering honestly because their reactions could be observed; it is easier to collect detailed qualitative data.

AO3 **Knowledge of research methods**
4 marks Accurate and reasonably detailed Accurate and reasonably detailed answer that demonstrates sound understanding of one or more advantages of using interviews rather than questionnaires in this situation.
3 marks Less detailed but generally accurate Less detailed but generally accurate answer that demonstrates relevant understanding of one or more advantages of using interviews rather than questionnaires in this situation.
2 marks Basic Basic answer that demonstrates some relevant understanding of one or more advantages of using interviews rather than questionnaires in this situation, but lacks detail and may be muddled.
1 mark Very brief / flawed Very brief or flawed answer demonstrating very little understanding of one or more advantages of using interviews rather than questionnaires in this situation.
0 marks No creditworthy material.

Q8.

AO3 = 4

Two strengths of using questionnaires could include:

- Compared to interview they are easy to use (1 mark). The researcher doesn't

STEMUP.CO.UK

need any special training to hand out the questionnaires (2nd mark for elaboration).

- People may be happier to disclose personal information on a questionnaire (1 mark) compared to a face-to-face situation (2nd mark for elaboration).
- Participants can answer the questions without the need for the researcher to be present (1 mark) so reducing experimenter bias (2nd mark for elaboration).
- If the questionnaire used closed questions which generate quantitative data, this is easier to analyse (1 mark) than open questions which generate qualitative data which is difficult to analyse (2nd mark for elaboration).
- Can be given to a large group of people (1 mark).

For each strength, 1 mark for identifying the strength and a further mark for explaining why it is a strength. The final bullet point is an example of a 1-mark answer as there is no explanation of why it is a strength.

Examiners should be aware that this question asks about the strengths of the method, not of the type of data collected. Answers that refer to data should not receive credit unless they are explicitly related to the type of question used (as illustrated in the bullet point above).

Q9.

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

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

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

One mark for either B or C.

One mark for an appropriate advantage of using open questions.
Likely points: open questions provide depth / detail / greater diversity of responses / more meaningful information in the response; they avoid participant frustration associated with fixed choice responses.

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

One mark for an appropriate conclusion that might be drawn, eg: the majority of people **regard themselves** as kind and helpful people.(Accept alternatives such as ‘see themselves, believe or think they are / say they would’)

One mark for justification of the answer with reference to the data given, eg: the number of people who reported they would help the person is much higher than any other response given (about 75% said they would help the person).

Accept other valid conclusions with an appropriate matching justification.

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

Up to 2 marks for an appropriate experimental hypothesis. For full credit the

STEMUP.CO.UK

hypothesis must be a testable statement and contain both the IV and DV.

Possible answers for 2 marks:
Non-directional: There is a difference in the number of participants who go to help / help someone when the participant waits alone and when the participant waits with another person.

Directional: More participants who wait alone go to help / help someone than participants who wait with another person. (Accept 'Fewer'.)

Accept null version of the hypothesis.

Possible answers for 1 mark:

There will be a difference in the number of participants who go to help / help in Condition 1 and Condition 2

People who wait alone are more likely to go to help / help than people than people who wait with someone else.

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

One mark for identification of a possible extraneous variable.
Likely answers: the behaviour of the interviewer who 'falls'; the behaviour of the confederate in the waiting room. Accept EVs based on participant variables eg gender and appropriate condition variables such as 'noise.'

One mark for explaining why the EV should be controlled.

One mark for explaining how it could be controlled.

Possible answers:
The behaviour of the interviewer who falls must be the same – the same sounds and cries so that each participant has the same incident to react to. This could be controlled by using a taped recording of the falling and crying out.
The behaviour of the confederate must be the same so that each participant has the same environment in the waiting room. This could be controlled by using the same person as a confederate who has a script he / she follows for each participant.

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

One mark for identification of the experimental design as independent groups / measures.

Up to 2 marks for explanation of why this is a suitable design for this study.
Likely points: the participants can only be exposed to the person 'falling' once (1) as they will then have some understanding of what the study is trying to find out and their behaviour will be affected by this knowledge (lack of naivety) (1).

Maximum of 1 mark for generic explanations not linked explicitly to the study

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

Up to 2 marks for an outline of the procedure of random sampling:

STEMUP.CO.UK

Possible answer:
Put the name of every first year student at the university into a hat (number every first year student)(1).
Draw out 40 names or numbers for the sample (use a random number table / computer program to generate a set of 40 numbers – this represents the sample) (1).

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

One mark for an appropriate suggestion.
Likely answer: Bar chart / bar graph, frequency graph. Accept pie chart.

One mark for justification of the suggestion.
Likely point: the display clearly demonstrates the numerical difference between the two conditions. Credit discrete data / categorical data.

If more than one graphical display is listed – mark the first answer.

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

For each of the TWO points, allow one mark for identification of the point and one further mark for discussion of why that point should be raised when the participants are debriefed. Max 2 marks for each point.

For full marks at least one of these points must focus on imparting the aim / purpose of the study or detail of the two conditions.

One further mark for discussion of the chosen point.

Maximum 2 marks if only ethical issue(s) discussed. These 2 marks can only be given for **one** ethical issue (1) that is appropriately discussed (1).

Likely points: explanation of the aim of the study; explanation of the use of independent groups; ethical issues, (these include deception, protection from harm / treating participants with respect; right to withdraw data from the study.)

Verbatim answers are likely to be credited with a maximum of two marks as there would be no discussion / explanation.

Q10.

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

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

(a) **AO3 = 3**

Advantages of using a questionnaire in this study could include that data from the hundred adults could be collected relatively quickly because the researcher would not need to be present when the questionnaires were completed; participants might be more willing to answer honestly because they would feel more anonymous; there might be a reduction in investigator effects because the researcher's reactions would not be visible. The advantage must be one that could be applied to this study.

STEMUP.CO.UK

1 mark for a slightly muddled or very brief outline of an advantage. Further marks for accurate elaboration.

(b) **AO3 = 2**

Qualitative is non-numerical and uses words to give a full description of what people think or feel.
1 mark for a very brief or slightly muddled answer eg qualitative data uses words.
2nd mark for accurate elaboration eg by comparison or by using an example.

(c) **AO3 = 2**

One mark for a question which would produce qualitative data but is not appropriate eg "How are you feeling?"
Two marks for an appropriate question eg "Tell me what it was like in the institution"
(Full marks can be awarded if it is not in the form of a question)
0 marks for a question that would not produce qualitative data.

(d) **AO3 = 1 + 1 + 3**

There are no ethical issues named in the specification, so any potentially relevant issues should be credited.
Likely ethical issues include informed consent, right to withdraw, protection from harm, confidentiality, respect or the need for debriefing in this particular case.
Other issues such as deception (deliberate or by omission) can be credited as they could apply in this research.
One mark each for identification of a relevant ethical issue.
One mark for a brief mention of how the issue could be dealt with.
Two further marks for elaboration appropriate to this research.
There is a depth / breadth trade-off. Candidates may explain one way of dealing with the issue in some depth, or mention several ways (of dealing with one issue) more briefly. Ethical issue one eg, right to withdraw (1 mark); ethical issue two eg confidentiality
(1 mark); Don't identify the participants (1 mark). Don't use photographs or names in published research. Names of people and / or places should be changed (2 further marks).

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) **[AO3 = 2]**

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

STEMUP.CO.UK

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

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

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

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

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

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

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

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

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

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

Accept ‘scientific’ and ‘validity’ only if justified.

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

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

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

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

STEMUP.CO.UK

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

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

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

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

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

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

Q12.

AO3 = 4

Strengths of questionnaires:

- Can be given to a large sample of people.
- Participants can answer the questionnaire without the need for the researcher to be present, so reducing experimenter bias.
- Compared with interviews they are easy to use, the researcher doesn't need any special training to use them.

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

Q1.

Part (a) was generally answered very well and most students were able to explain why matching the students was important. However, some failed to apply their response to this specific situation as thus were limited to 1 mark.

In part (b) many students were able to identify an appropriate variable to match the students on and then go onto explain how this might affect satisfaction scores. There were many possible accurate responses with the most common being how difficult they may find the topic (often linked to IQ) and prior knowledge of business studies. A significant minority of students identified an appropriate variable but then failed to apply it to this situation and explain how it may have affected satisfaction scores. There was also a significant number of students who failed to engage with the question and simply suggested one of a number of generic variables such as IQ, gender or age. Unless these variables were clearly linked to an appropriate variable (such as linking age to interest/experience of using computers) they would not be appropriate variables to match students on in this study and thus would not receive credit. It is important to remind students to engage with the question posed and tailor their answer to that specific question not to simply produce a response of non-specific rote-learned material.

In part (c) it was clear from student responses that the vast majority of students appeared to have some understanding of the process of random allocation. However, fewer students seem to clearly understand the reasons why random allocation is used and thus failed to answer the question posed. Of those students who did answer the question correctly some failed to apply it to this situation and thus were limited to 1 mark. It is important to note that simply referring to removal of bias without any outline or application is not sufficient to gain any credit. Reference to bias needed to be clearly focused on either removal of researcher bias or removal of participant bias and should be outlined and applied for full credit.

In part (d) most students clearly understood the strengths of stratified sampling and / or the weaknesses of opportunity sampling. However, many failed to tailor this knowledge to the actual question posed and did not explain why stratified sampling might be better than opportunity sampling in this study. Most students produced limited responses where they simply explained the strengths of stratified sampling or the weaknesses of opportunity sampling. Some students did explain why stratified sampling was an improvement in general but failed to apply their reasoning to this specific design. Equally there were many responses which had implicit explanation of improvements which was not clear and so not sufficient for full marks. A significant minority of students either produced responses which had no relevant content or did not attempt the question suggesting that some students did not have the knowledge to enable them to answer this question. There was also some confusion with systematic sampling. This is particularly disappointing considering there was a similar question on the previous examination paper in the June 2017 series which also asked about how stratified sampling could be an improvement (although in comparison to volunteer sampling). It is important to remind students that just because a topic appears on an exam in one series does not mean that the same topic cannot appear on the exam in the following series. It is also important to use previous examination questions with students to ensure they are familiar with the style and type of questions which they may be presented with in their actual examination.

In part (e) the overwhelming majority of students answered correctly and there was a wide range of appropriate questions which would generate qualitative data seen in student responses. Unfortunately, a very small minority of students wrote Likert scale questions which would have generated quantitative data.

STEMUP.CO.UK

In part (f) there were some excellent responses with a range of different appropriate problems with analysing qualitative data seen. Most students who identified an appropriate problem went on to clearly explain it and thus gain full credit. However, some students failed to address the actual question posed and explained problems either with collecting qualitative data (such as ‘it takes time for participants to complete’) or with the validity of the qualitative data (such as ‘participants might lie in their responses’). Responses must be based on analysis to gain credit and the most common error seen was where students were focusing on subjectivity of the participants when answering the questions to produce the qualitative data rather than subjectivity of the researcher when analysing the qualitative data.

In part (g) most students answered correctly and a wide range of outlined ethical issues was seen in responses. Unfortunately, a few students simply named an ethical issue without outlining it.

In part (h) most students were able to draw an accurate scattergram with two appropriately labelled axes on suitable scales. A small minority of these students also produced an accurate and appropriate title which referred to the relationship between the two co-variables. However, many students failed to either provide a title for their graph or produced an inaccurate or vague title which was not sufficient for credit. There were also errors in plotting points in some responses and some students produced very small graphs with inappropriate scaling. Some students failed to read the question correctly and did not draw a scattergram. A variety of incorrect types of graph were seen but the most common was a line graph where students had joined up the plotted points ethical issue without outlining it.

The vast majority of students scored one mark in part (i) as few responses had both components required for two marks. Most responses only included a description of the relationship between online hours and satisfaction but did not identify the relationship as a negative correlation. The description of the relationship was expressed in several different accurate ways but the most common was more online hours the less satisfied they were. A small number of students produced confused conclusions based on causal relationships which were incorrect.

Part (j) was answered very well with the vast majority of students correctly calculating both the mode and median satisfaction scores and thus being awarded full marks. A very small minority made a mistake in their calculations leading them to produce an incorrect median and/or mode. Unfortunately, a very small number of students failed to fully answer the actual question posed and calculated a mean value instead of either the median or mode.

In part (k) there was evidently much confusion amongst students as to why the median would be more representative than the mode in this data set. Most responses simply consisted of statements of how to calculate the median and the mode rather than explaining a reason why the median might be more suitable to represent the average score in this study. Some students were able to explain why the mode was less representative of the average satisfaction scores and thus produced a limited explanation but very few were able to explain why the median was more representative. Only a very small minority of students were able to explain that the median was more representative as it is numerically closer to more of the other scores in the data set compared to the mode and thus few students gained 2 marks. Some students simply repeated the question stating that the median is more representative than the mode without any explanation.

There were a lot of correct answers to part (l) and it was apparent that many students had the knowledge to be able to identify the negative skew in this data. However, when compared to the other MCQ on this paper, more students did not attempt this question and the number of correct responses was lower. This suggests that their knowledge of negative skew was poorer than their knowledge of introspection and overt observations.

Q2.

Performance on part (a) was surprisingly poor because many students focussed on causes of behaviour rather than measurement. Answers had to be contextualised in order to gain credit.

In part (b) most students gained some credit here and there were some excellent straightforward suggestions. High scoring answers often suggested the use of a Likert scale with scale and gave example statements, or observations of behaviour in the queue, with examples of behavioural categories and details of event or time sampling. There were some ingenious suggestions, such as analysing facial expressions from the photographs taken by auto cameras at the top of the roller coaster. Other ideas were just too impractical or implausible to credit. Vague suggestions, such as 'interview them in the queue and ask how they feel' gained limited credit, but a few responses were just too brief, for example, simply stating 'interview' or 'questionnaire'. A few biological suggestions were seen.

Q3.

(a) Most students were able to give a disadvantage of quantitative data, but some failed to provide appropriate application. Some common incorrect answers included disadvantages of qualitative data, advantages of quantitative data and advantages/disadvantages of questionnaires.

(b) Nearly all students managed to gain at least one mark and many gained full credit. The quantitative question suggested was generally effective, but the qualitative question suggested could have produced a quantitative response, so did not gain credit. Some students got qualitative and quantitative the wrong way round and consequently gained no marks. The most popular method of answering the quantitative part was either to give a rating scale or to pose a question followed by the words *"Yes/No"*. Although giving a closed answer response that is not officially quantitative, it was credited for this series. However, schools and colleges should ensure students are able to distinguish between quantitative and closed questions for future series.

Q4.

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

(b) There were many different suggestions given for random allocation and all suggestions that were practical and would result in the required groups were credited. However, many responses never went beyond the idea of putting names in a hat or into a random number generator, meaning that all the researcher would have is names in a hat or in a generator. Students should be able to describe accurately what they would do to reach the final stage of having the required number of people in the required groups. It was very clear that many students had no idea of what they would have to instruct the computer or random number generator to do in order to reach this stage.

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

STEMUP.CO.UK

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

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

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

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

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

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

Q7.

(a) Some students clearly understood what behavioural categories are while others had no idea. A number of students were unable to suggest operationalised behaviour categories. Kicking and swearing would be suitable, physical and verbal aggression would be unsuitable. Given the material in the stem, just suggesting children would be observed in the playground, with no further elaboration did not attract credit. A few students did not take account of the stem material and suggested how children could be observed in the home at the age of two.

(b) Most students could identify an ethical issue but some chose issues that were difficult to apply to this scenario. Many ethical issues could be made relevant; the need for informed consent and the need for confidentiality were apt. The effectiveness of the elaboration distinguished between students' marks. Students who chose to write about protection from harm needed to make an appropriate case for this.

(c) Students who responded to this question in the context of the stem material were able to contextualise the comparison between interviews and questionnaires and that helped in producing an appropriate answer. A number of students produced good answers referring to the interviewer's opportunity

STEMUP.CO.UK

to clarify questions and answers and the benefits that might accrue from being face to face with the parents. Students needed to explain why parents might be more likely to lie in a questionnaire or in an interview, since a case could be made for either. Some students who did not read the question carefully wrote about the advantages of questionnaires. A few students thought the children, rather than the parents were being interviewed.

Q8.

Answers need to explain the strength rather than merely state it. So to state that questionnaires are “quick” gains no credit because there is no explanation as to why they are quick. Better answers compared questionnaires to other methods such as interviews. So to say that a questionnaire is quick in comparison to an interview becomes a creditworthy comment. Since the question asked about the strength of the methods, answers that referred only to the type of data did not receive credit, unless they were explicitly linked to the type of question that would generate such data. Some students gave generic answers that could apply to either questionnaires or interviews, and these received minimal credit.

Q9.

(a) This was generally answered well.

(b) Students failed to notice that the appropriate conclusion the researchers could draw was that the participants **believed / said** they would help. The actual behaviour was not measured by the questionnaire.

(c) Students are still poor at writing clear hypotheses, with both conditions of the IV present and a measurable DV. Many answers were aims and the expression ‘more likely’ was often used, as was reference to Condition 1 and Condition 2.

(d) It was evident from many answers that students could not spot obvious variables on which researchers would focus their attention. Instead, answers included reference to the ‘heating’, with the idea that the room was likely to become so hot / cold that participants would lose the ability to hear what was going on.

(e) Many students lost a mark for this question because they produced a generic explanation for the suitability of the experimental design and did not relate their answers to the study described.

(f) This was not answered well. Many explained what random sampling is or gave advantages or limitations of the method. It was rare to see an answer that identified what must go into the ‘hat / random-number generator’ and what would then happen. There was a great deal of confusion with random allocation.

(g) Most students suggested a bar chart and could produce a sensible reason for choosing such a display.

(h) Unfortunately, many students did not recognise that, if asked about debriefing participants, then the focus is on telling people everything about the study in which they have just taken part. This is especially important in studies using independent groups design, as participants are not only unaware of the aim of the investigation, but also that other people performed under different conditions. Too many answers merely concentrated on ethical points.

Q10.

(a) Students who scored well often focussed on the anonymity of questionnaires, the lack of investigator effects or the time advantage where questionnaires could be simultaneously completed. Whether students gained full marks depended on how effectively they were able to explain the advantage they had identified. Better answers compared questionnaires to interviews, or referred to the relatively large number of adults in this study. Some students referred to the advantages of analysing data from questionnaires which was not the focus of the question.

(b) Most responses explained the term qualitative data appropriately. A few students described quantitative data; given that the word quantitative can be so easily aligned with number, it is surprising that students get muddled about these terms.

(c) Most responses were appropriate, although a number of questions provided would have produced numerical data (eg how long ?x2, how many ?x2) or categorical answers (usually yes / no responses).

(d) Although most students had no difficulty in identifying two ethical issues, many students were less successful in providing suitable suggestions for how one of these issues could be dealt with. Some students just re-stated the ethical issue. Other students filled up the answer space by explaining how both ethical issues could be dealt with, leaving the examiner to decide which was the more credit-worthy answer.

Q11.

(a) Students often struggled with this question. Very few understood why measures of dispersion are used in addition to measures of central tendency and a number used the term ‘dispersed’. Applications to the stem lacked the necessary detail to attract a mark.

(b) Many students were able to make correct use of the table and draw an appropriate conclusion about the statistical significance of the T value. Stronger students were able to present this information well, with some even correctly stating that the results were not significant at the 0.02 level and explaining why. A few less successful students were confused about the critical and calculated values of T.

(c) It was heartening to see so many students being able to explain why this test was used, with many students scoring the full 3 marks. A few students, however, stated that the data was nominal or interval.

(d) Most students were able to score at least one mark on this question even if they scored poorly on preceding and subsequent questions. Students had to think carefully about this answer and many were able to suggest one or two sensible reasons for the use of a diary by each offender. Elaboration and explanation of each reason proved more of a challenge, sometimes resulting in overlap and repetition across the two reasons offered. Some students simply repeated the stem.

(e) This question produced some answers that showed a lack of understanding of the use of a control group with quite a few students suggesting that a control group would consist of ‘a group of non-offenders’ or ‘normal people / people with no anger issues from a normal population’. More informed students were able to explain why a group of people who would not have the anger management programme would have improved the study although some simply said that ‘it would make it more scientific’ without explaining why.

(f) Answers to this question were most disappointing with almost two fifths of students failing to score a single mark. What should have been a straightforward question proved challenging for many students who seemed unfamiliar with how to establish

STEMUP.CO.UK

the reliability and validity of the questionnaire. Some students confused reliability with validity, others simply provided definitions of each and, where an attempt was made to apply reliability and validity to the stem, students very often referred to pilot studies and peer review. Reliability was sometimes mistaken for replicability, the 'split-half method' was frequently explained as dividing the total score in half to see if the two halves correlated and test- retest was sometimes explained as testing one group of offenders and retesting another group of offenders. Face validity was rarely applied to the questionnaire ie anger scores. Even students who correctly addressed the issues of reliability and validity, failed to explain how statistical tests of correlation would be used in this context. The value of carrying out practical activities to enable students to 'think like a psychologist' and apply their knowledge of practical activities in answer to questions such as this one cannot be overstated. It is clear that where students had been presented with such opportunities, they were able to write confidently and in an informed manner.

Q12.

This question asked candidates to explain the strengths, not merely describe a characteristic of a questionnaire. There were a wide range of answers, such as if the questionnaire is anonymous, respondents may be more honest; if they are filled in private demand characteristics are reduced. However, answers that simply said that they are 'cheap' or 'quick', gained no credit unless there was some elaboration. For example, questionnaires are quicker than interviews. Describing the type of data as a strength was in itself not credit-worthy, since such data could often be generated by other methods. However, if reference was made to the type of question and the data it would generate, this could be credit-worthy. For example, open-ended questions produce qualitative data that can be rich and varied.

Many candidates misread the question and provided one strength and one limitation.