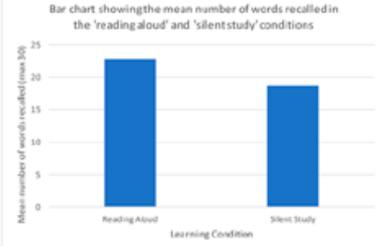


Mark scheme

Question		Answer/Indicative content	Marks	Guidance
1		A		range
		Total	1	
2	a	B		15
	b	A		5
		Total	2	
3		<p>Outline one way you could uphold the ethical consideration of respect in this study.</p> <div style="border: 1px solid black; padding: 5px;"> <p>Under BPS ethical considerations ...</p> <p>respect = informed consent, right to withdraw and privacy/confidentiality.</p> <p>For example, one way to do this would be to inform primary school children at the beginning of the study on concentration that they can withdraw at any time.</p> <p>One way to uphold ethical consideration of respect clearly presented in context.</p> <p>Attempt to outline one way to uphold ethical consideration of respect (whether in context or not)</p> <p>The candidate has not provided any creditworthy information</p> </div>	<p>Max 2</p> <p>2</p> <p>1</p> <p>0</p>	<p>Context = concentration, slippers, shoes, learning, primary school children etc.</p> <p>Mere identification of the ethical guideline gains no credit.</p> <p>Informed consent is only creditworthy if linked to parental/headteachers consent.</p> <p>NB: Only first response is marked.</p> <p><u>Examiner's Comments</u></p> <p>This question differentiated well as answers covered the full range of the mark scheme. The best responses identified relevant ethical guideline that comes under the principle of respect and explained how they would uphold it in the context of the scenario presented.</p> <p>A significant proportion of candidates got zero marks because they chose an ethical consideration that is not under respect such as debrief or they did not consider the implications of the scenario, i.e. they referred to gaining consent from the children themselves. Candidates need to be aware that informed consent can only be gained from a consenting adult - either the participant themselves or a parent/guardian.</p>

		Total	2						
4		<p>Explain two factors that could affect the external validity of this study.</p> <p>Likely answers: size/diversity/representativeness of sample; ecological validity of task set to assess concentration, ecological validity of the setting, etc.</p> <p>3 marks for each factor outlined ...</p> <p>Clear explanation of how external validity could be affected in context.</p> <table border="1"> <tr> <td rowspan="2">Clear explanation of how external validity could be affected, but not in context.</td> <td>OR</td> </tr> <tr> <td>attempted explanation of how external validity could be affected in context.</td> </tr> </table> <p>Identification of relevant factor/type of external validity or weak attempt to explain how external validity could be affected (whether in context or not).</p> <table border="1"> <tr> <td>The candidate has not provided any creditworthy information</td> <td></td> </tr> </table>	Clear explanation of how external validity could be affected, but not in context.	OR	attempted explanation of how external validity could be affected in context.	The candidate has not provided any creditworthy information		<p>Max 6 [3+3] 3 2 1 0</p>	<p>Context = concentration, slippers, shoes, learning, primary school children etc.</p> <p>NB. If candidates refer to types of external validity that are not listed on the specification these can be credited, e.g. temporal validity (teaching methods change over time)</p> <p><u>Examiner's Comments</u></p> <p>Good responses accurately identified factors which could impact external validity. The strongest referred to population and ecological validity as separate points with contextualisation. A notable number of candidates was unable to differentiate between internal and external validity, which meant low (or zero) marks.</p>
		Clear explanation of how external validity could be affected, but not in context.		OR					
			attempted explanation of how external validity could be affected in context.						
		The candidate has not provided any creditworthy information							
		Total	6						
5		<p>Outline one conclusion that can be made from the raw data presented in this table.</p> <p>Conclusions could include:</p> <ul style="list-style-type: none"> • Reading aloud seems to facilitate memory, perhaps because the act of reading enables the words to be practiced more and processed at a deeper level • There are some individual differences, so reading aloud does not improve memory for everyone, indicating cognitive processes work differently for different people. <p>Accept any other appropriate conclusions here.</p> <p>Clear, detailed conclusion in context (or supported by data).</p> <table border="1"> <tr> <td>Clear, detailed conclusion but not in context.</td> <td>OR attempt in context.</td> </tr> </table> <p>Brief and/or weak attempt (whether in context or not)</p>	Clear, detailed conclusion but not in context.	OR attempt in context.	<p>3 3 2 1 0</p>	<p>Context = reading aloud, silence, recall, memory, etc.</p> <p>A conclusion must be an interpretation/application of the findings / data (not simply a statement of the result(s) obtained).</p> <p>Max 1 mark for presentation of a finding (involving comparison of data) with no interpretation/explanation of it.</p> <p>Zero marks if just data is given.</p> <p>NB: Only first response is marked.</p> <p><u>Examiner's Comments</u></p> <p>Candidates really struggled to differentiate between findings and conclusions. Although almost all candidates included context in their</p>			
		Clear, detailed conclusion but not in context.	OR attempt in context.						
				Total	6				

		<p>The candidate has not provided any creditworthy information</p>		<p>response, the vast majority of those simply noted a finding and did not explain what the implications of the finding are for memory/recall. For those who attempted to explain the implications, often they made a general statement that memory is improved, rather than attempting to explain why reading aloud may improve memory.</p> <p> Assessment for learning</p> <p>Candidates need to practise forming a conclusion. Phrases such as: 'this is because...', 'this suggests that...' might be useful.</p>						
		<p>Total</p>	<p>3</p>							
<p>6</p>		<p>Draw a fully labelled bar chart showing the mean number of words recalled in each condition. Plot the means to two significant figures.</p> <div data-bbox="236 1272 746 1545" style="border: 1px solid black; padding: 5px; text-align: center;"> <p>Bar chart showing the mean number of words recalled in the 'reading aloud' and 'silent study' conditions</p>  <table border="1" style="margin: 0 auto; border-collapse: collapse;"> <thead> <tr> <th>Learning Condition</th> <th>Mean number of words recalled (max 30)</th> </tr> </thead> <tbody> <tr> <td>Reading Aloud</td> <td>23</td> </tr> <tr> <td>Silent Study</td> <td>19</td> </tr> </tbody> </table> </div> <div data-bbox="236 1556 746 1697" style="border: 1px solid black; padding: 2px;"> <p>1 mark is awarded for correctly calculating then presenting by value each bar representing the mean number of words recalled (to 2 significant figures) in the 'reading aloud' and 'silent' conditions</p> </div> <div data-bbox="236 1709 746 1747" style="border: 1px solid black; padding: 2px;"> <p>1 mark is awarded for clear labelling of the x axis</p> </div> <div data-bbox="236 1758 746 1830" style="border: 1px solid black; padding: 2px;"> <p>1 mark is awarded for clear labelling of the y axis including measurement (which must start at 0).</p> </div> <div data-bbox="236 1841 746 1877" style="border: 1px solid black; padding: 2px;"> <p>1 mark is awarded for a fully operationalised title.</p> </div>	Learning Condition	Mean number of words recalled (max 30)	Reading Aloud	23	Silent Study	19	<p style="text-align: center;">Max 4 [1+1+1+1] 1 1 1 1</p>	<p>Context = reading aloud, silence, recall, memory, etc.</p> <p>Mean values for each condition need to be calculated first.</p> <p>Mean for reading aloud condition = 23 to 2 sf.</p> <p>Mean for silent condition = 19 to 2 sf.</p> <p>Title must include both variables: (mean) number of words recalled and reading aloud/silent study conditions).</p> <p>Response must make it clear that this is the mean number of words (max 30) recalled in either title or y axis, if not max 3.</p> <p>Labels on axes must be clear. X axis - reading aloud/silent study Y axis - mean number of words recalled - and measurement must start at 0 (does not need to go up to 30 and can go beyond 30).</p> <p>If two bars are together (or if the bar touches y axis) do not award</p>
Learning Condition	Mean number of words recalled (max 30)									
Reading Aloud	23									
Silent Study	19									

				mark for correct presentation of data. Examiner's Comments Almost all candidates knew how the bar chart is supposed to look. However, a significant proportion of candidates lost at least 1 mark. The most common mistakes included: not plotting the bar to two significant figures, not operationalising conditions in the title, incorrect labelling of the y axis by stating the number of words recalled rather than mean number. Compared to previous series, fewer candidates placed the bars together.									
			Total	4									
7		<p>Give one reason why the Mann Whitney U test is the appropriate inferential test to use to analyse the data from this study.</p> <table border="1"> <tr> <td colspan="2">Any one reason (in bold) in context from:</td> </tr> <tr> <td> <ul style="list-style-type: none"> It is a test for independent measures design, and this study had different participants in the reading aloud condition compared to the silent condition. It is a test that uses at least ordinal data and this study had ordinal data as it was scores out of 30 in a memory test. It is a test that assesses differences between conditions, and this study assessed differences in memory between the reading aloud and silent conditions. </td> <td></td> </tr> <tr> <td colspan="2">One appropriate reason in context</td> </tr> <tr> <td>One appropriate reason but not in context.</td> <td>OR attempt to give one appropriate reason in context.</td> </tr> <tr> <td colspan="2">The candidate has not provided any creditworthy information.</td> </tr> </table>	Any one reason (in bold) in context from:		<ul style="list-style-type: none"> It is a test for independent measures design, and this study had different participants in the reading aloud condition compared to the silent condition. It is a test that uses at least ordinal data and this study had ordinal data as it was scores out of 30 in a memory test. It is a test that assesses differences between conditions, and this study assessed differences in memory between the reading aloud and silent conditions. 		One appropriate reason in context		One appropriate reason but not in context.	OR attempt to give one appropriate reason in context.	The candidate has not provided any creditworthy information.		<p>Context = reading aloud, silence, <u>recall of words</u>, memory, etc.</p> <p>'6 participants' not enough for context.</p> <p>NB: Only first response is marked</p> <p>Examiner's Comments</p> <p>Candidates were able to identify at least one reason for the use of Mann-Whitney U test, however the use of context was inconsistent. The most common mistake was to give two reasons and not to include context, which resulted in only 1 mark.</p>
Any one reason (in bold) in context from:													
<ul style="list-style-type: none"> It is a test for independent measures design, and this study had different participants in the reading aloud condition compared to the silent condition. It is a test that uses at least ordinal data and this study had ordinal data as it was scores out of 30 in a memory test. It is a test that assesses differences between conditions, and this study assessed differences in memory between the reading aloud and silent conditions. 													
One appropriate reason in context													
One appropriate reason but not in context.	OR attempt to give one appropriate reason in context.												
The candidate has not provided any creditworthy information.													
			Total	2									
8		<p>Before using the formula for the Mann Whitney U test, the data obtained must be ranked. In the result two participants have the same</p>	<p>Max 3 3 2</p>	<p>Context = reading aloud, silence, recall, memory, etc.</p> <p>NB. Specific scores could be</p>									

		<p>score of 24. Explain how this is dealt with when ranking the data.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td colspan="2" style="padding: 5px;">As the two scores of 24 were the same, so they receive the same rank of 9.5 as the ranks 9 and 10 have been shared, i.e. $9 + 10 / 2 = 9.5$</td> </tr> <tr> <td colspan="2" style="padding: 5px;">Clear explanation in context of how having the same scores is dealt with including references to the same rank and a way of finding the same rank.</td> </tr> <tr> <td style="width: 50%; padding: 5px;">Clear explanation of how having the same scores is dealt with including references to the same rank (or simply 9 and 10) and a way of finding the same rank (no context).</td> <td style="width: 50%; padding: 5px;">OR Explanation of why two ranks are the same in context.</td> </tr> <tr> <td colspan="2" style="padding: 5px;">Attempt to explain why two ranks are the same (whether in context or not).</td> </tr> <tr> <td colspan="2" style="padding: 5px;">The candidate has not provided any creditworthy information</td> </tr> </table>	As the two scores of 24 were the same, so they receive the same rank of 9.5 as the ranks 9 and 10 have been shared, i.e. $9 + 10 / 2 = 9.5$		Clear explanation in context of how having the same scores is dealt with including references to the same rank and a way of finding the same rank.		Clear explanation of how having the same scores is dealt with including references to the same rank (or simply 9 and 10) and a way of finding the same rank (no context).	OR Explanation of why two ranks are the same in context.	Attempt to explain why two ranks are the same (whether in context or not).		The candidate has not provided any creditworthy information		<p>1 0</p>	<p>credited as context.</p> <p><u>Examiner's Comments</u></p> <p>Most candidates were able to identify the issue of tied ranks, which was contextualised. Not all explained how to find the shared rank (by either calculating the mean of ranks or finding median/middle value).</p>
As the two scores of 24 were the same, so they receive the same rank of 9.5 as the ranks 9 and 10 have been shared, i.e. $9 + 10 / 2 = 9.5$														
Clear explanation in context of how having the same scores is dealt with including references to the same rank and a way of finding the same rank.														
Clear explanation of how having the same scores is dealt with including references to the same rank (or simply 9 and 10) and a way of finding the same rank (no context).	OR Explanation of why two ranks are the same in context.													
Attempt to explain why two ranks are the same (whether in context or not).														
The candidate has not provided any creditworthy information														
Total		3												
9	a	<p>Calculate the U value for the Mann Whitney U test for the data collected in this study. Show your workings. You may use the formula presented below.</p> <p style="font-size: small;">U = the smaller of U_1 and U_2</p> <p style="font-size: x-small;">Where U_1 is ... and U_2 is ...</p> $U_1 = R_1 - \frac{n_1(n_1 + 1)}{2}$ $U_2 = R_2 - \frac{n_2(n_2 + 1)}{2}$ <table border="1" style="width: 100%; border-collapse: collapse; margin-top: 10px;"> <tr> <td style="padding: 5px;">1 mark for each of the following correct / evident in answer ..</td> </tr> <tr> <td style="padding: 5px;">1 mark for sum of ranks for reading aloud condition calculated correctly (46.5)</td> </tr> <tr> <td style="padding: 5px;">1 mark for sum of ranks for silent study group calculated correctly (31.5).</td> </tr> <tr> <td style="padding: 5px;">1 mark for correct calculation of U_1 (25.5) value OR correct calculation of U_2 value (10.5).</td> </tr> <tr> <td style="padding: 5px;">1 mark for <u>all</u> workings of U_2 value shown.</td> </tr> <tr> <td style="padding: 5px;">1 mark for choosing the final U value.</td> </tr> <tr> <td style="padding: 5px;">The candidate has not provided any creditworthy information.</td> </tr> </table>	1 mark for each of the following correct / evident in answer ..	1 mark for sum of ranks for reading aloud condition calculated correctly (46.5)	1 mark for sum of ranks for silent study group calculated correctly (31.5).	1 mark for correct calculation of U_1 (25.5) value OR correct calculation of U_2 value (10.5).	1 mark for <u>all</u> workings of U_2 value shown.	1 mark for choosing the final U value.	The candidate has not provided any creditworthy information.	<p>Max 5 [1+1+1+1+1]</p> <p>1 1 1 1 0</p>	<table border="1" style="width: 100%; border-collapse: collapse; margin-bottom: 10px;"> <tr> <td style="width: 50%; padding: 5px;"> $U_1: = \frac{46.5 - \frac{6(6+1)}{2}}{2}$ $= \frac{46.5 - 6 \times 7}{2}$ $= \frac{46.5 - 42}{2}$ $= \frac{46.5 - 21}{2}$ $= 25.5$ </td> <td style="width: 50%; padding: 5px;"> $U_2: = \frac{31.5 - \frac{6(6+1)}{2}}{2}$ $= \frac{31.5 - 6 \times 7}{2}$ $= \frac{31.5 - 42}{2}$ $= \frac{31.5 - 21}{2}$ $= 10.5$ </td> </tr> </table> <p>U_2 is the smaller of the two scores so $U = 10.5$ (candidates may indicate this by circling the correct U value)</p> <p>NB. Accept alternative workings for U_1 and U_2 if the correct answer is calculated.</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates responded really well to this question showing full workings and calculating correct U value. Errors on this question typically stemmed from incorrectly adding the ranks for the conditions. A fair number of candidates lost a mark as they did not indicate the final U value.</p>	$U_1: = \frac{46.5 - \frac{6(6+1)}{2}}{2}$ $= \frac{46.5 - 6 \times 7}{2}$ $= \frac{46.5 - 42}{2}$ $= \frac{46.5 - 21}{2}$ $= 25.5$	$U_2: = \frac{31.5 - \frac{6(6+1)}{2}}{2}$ $= \frac{31.5 - 6 \times 7}{2}$ $= \frac{31.5 - 42}{2}$ $= \frac{31.5 - 21}{2}$ $= 10.5$	
1 mark for each of the following correct / evident in answer ..														
1 mark for sum of ranks for reading aloud condition calculated correctly (46.5)														
1 mark for sum of ranks for silent study group calculated correctly (31.5).														
1 mark for correct calculation of U_1 (25.5) value OR correct calculation of U_2 value (10.5).														
1 mark for <u>all</u> workings of U_2 value shown.														
1 mark for choosing the final U value.														
The candidate has not provided any creditworthy information.														
$U_1: = \frac{46.5 - \frac{6(6+1)}{2}}{2}$ $= \frac{46.5 - 6 \times 7}{2}$ $= \frac{46.5 - 42}{2}$ $= \frac{46.5 - 21}{2}$ $= 25.5$	$U_2: = \frac{31.5 - \frac{6(6+1)}{2}}{2}$ $= \frac{31.5 - 6 \times 7}{2}$ $= \frac{31.5 - 42}{2}$ $= \frac{31.5 - 21}{2}$ $= 10.5$													

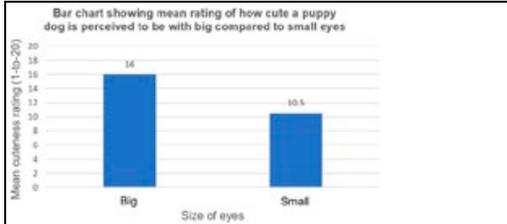
	b	<p>How is the critical value used to determine if the findings are statistically significant?</p> <p>For one mark this must refer to comparison with calculated value.</p> <p>It is compared to the appropriate/correct calculated/observed value.</p>	1	<p>Accept answers related directly to the Mann-Whitney test.</p> <p><u>Examiner's Comments</u></p> <p>Overall, the question was answered well. Some candidates lost marks as they referred to expected rather than observed/calculated value.</p>								
		Total	6									
10		<p>Explain what $p > 0.05$ would mean if it appeared as part of the significance statement when reporting the findings from this study.</p> <div style="border: 1px solid black; padding: 5px;"> <p>Answer could refer to the following features:</p> <ol style="list-style-type: none"> 1. Greater than 5% probability that the results are due to chance. Null hypothesis is accepted 2. AND/OR the alternative hypothesis is rejected. The results are not 3. significant/there is no significant difference. 4. Less than 95% confident that the results are significant. Observed value was greater than 5. the critical value, therefore results are not significant. </div> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td colspan="2">2 correct features in context.</td> </tr> <tr> <td style="width: 50%;">2 correct features, but not in context.</td> <td style="width: 50%; text-align: center;">OR 1 correct feature in context.</td> </tr> <tr> <td colspan="2">1 correct feature, not in context.</td> </tr> <tr> <td colspan="2">The candidate has not provided any creditworthy information</td> </tr> </table>	2 correct features in context.		2 correct features, but not in context.	OR 1 correct feature in context.	1 correct feature, not in context.		The candidate has not provided any creditworthy information		<p>Max 3</p> <p>3</p> <p>2</p> <p>1</p> <p>0</p>	<p>Context = reading aloud, silence, recall, memory, etc.</p> <p>Example 3-mark answer: The results are not significant, so the null hypothesis is accepted (i.e. there is no difference in memory when words are read aloud compared to when studied in silence).</p> <p>Example 2-mark answer: There is no significant difference between how many words are recalled when words are read aloud or studied in silence.</p> <p>Example 1-mark answer: The null hypothesis should be accepted.</p> <p><u>Examiner's Comments</u></p> <p>The vast majority of candidates misread the greater than sign, which meant that their interpretation of significance was incorrect. It was quite common for candidates to incorrectly state that the research is significant. Some candidates did not understand what a significance statement is and instead provided suggestions as to how to look up a critical value.</p> <p>Exemplar 3</p> <p><i>There is no significant difference as there is a more than 5% probability the difference between correct recall of words in a silent study or reading aloud condition are due to chance. Accept the null hypothesis and reject the alternate as the conditions of study do not have a significant effect on correct recall. /30 on a memory test. ps</i></p>
2 correct features in context.												
2 correct features, but not in context.	OR 1 correct feature in context.											
1 correct feature, not in context.												
The candidate has not provided any creditworthy information												

					Exemplar 3 shows correct interpretation of the significance statement <u>in context</u> .	
			Total	3		
11	a	<p>Explain one strength of using quantitative data in this study.</p> <p>Strengths could include:</p> <ul style="list-style-type: none"> Easier data analysis and comparison across conditions (reading aloud vs silent) Easier to record data Easier to interpret. Objective Easier to check for consistency etc. <p>Clear explanation of strength in context.</p> <table border="1"> <tr> <td>Clear explanation of strength but not in context.</td> <td>OR attempted explanation of strength in context.</td> </tr> </table> <p>Brief and/or weak attempt to explain strength (whether in context or not).</p> <p>The candidate has not provided any creditworthy information</p>	Clear explanation of strength but not in context.	OR attempted explanation of strength in context.	<p>Max 3</p> <p>3</p> <p>2</p> <p>1</p> <p>0</p>	<p>Context = reading aloud, silence, recall, memory, etc.</p> <p>NB: Only first response is marked.</p>
Clear explanation of strength but not in context.	OR attempted explanation of strength in context.					
	b	<p>Explain one weakness of using quantitative data in this study.</p> <p>Weaknesses could include:</p> <ul style="list-style-type: none"> Doesn't inform us about reasons why there is a difference or not in memory between reading aloud and studying words in silence, because it lacks detail. Construct validity issues (simplifying complex behaviours to a score) etc. <p>Clear explanation of weakness in context.</p> <table border="1"> <tr> <td>Clear explanation of weakness but not in context.</td> <td>OR attempted explanation of weakness in context.</td> </tr> </table> <p>Brief and/or weak attempt to explain weakness (whether in context or not).</p> <p>The candidate has not provided any creditworthy information</p>	Clear explanation of weakness but not in context.	OR attempted explanation of weakness in context.	<p>Max 3</p> <p>3</p> <p>2</p> <p>1</p> <p>0</p>	<p>Context = reading aloud, silence, recall, memory, etc.</p> <p>NB: Only first response is marked.</p> <p><u>Examiner's Comments</u></p> <p>As with previous strengths/weaknesses questions, the best responses followed a PEC structure, although sometimes the comment did not link back to the point identified and therefore did not get the final mark. Some simply stated 'therefore more/less valid and more reliable' rather than being specific.</p>
Clear explanation of weakness but not in context.	OR attempted explanation of weakness in context.					
		Total	6			

12		<p>Outline two ways that an independent measures design could affect the validity of this study.</p> <div style="border: 1px solid black; padding: 5px;"> <p>Answers could include:</p> <p>No order effects from having</p> <ul style="list-style-type: none"> • already studied words in previous condition, so increased validity. <p>Individual differences in memory may have influenced the findings</p> <ul style="list-style-type: none"> • irrespective of learning condition (reading aloud vs silent), so decreased validity. <p>Can keep the learning material (words) the same in each condition,</p> <ul style="list-style-type: none"> • so words in one condition are not easier or harder to remember than others, so increased internal validity. </div> <p>Accept any other appropriate strengths or weaknesses.</p> <p>Identified strength/weakness explained in relation to impact on validity and in context.</p> <p>Some understanding of impact on validity, not in context.</p> <p>The candidate has not provided any creditworthy information.</p>	<p style="text-align: center;">4 [2 + 2] 2 1 0</p>	<p>Context = reading aloud, silence, recall, memory, etc.</p> <p>Evaluation points are likely to be based on those in part (a) but other points can be used. However, do not credit a point that is just a repetition of a strength/weakness given in Q32a, unless it is specifically related to the effect on validity.</p> <p>Accept both positive and negative impacts on validity.</p> <p>Do not credit 'eradicates/eliminates/rules out demand characteristics'.</p> <p>Do not accept population validity.</p> <p>Do not credit issues that relate to reliability (e.g. therefore increases consistency).</p> <p><u>Examiner's Comments</u></p> <p>Common responses to this question included reference to demand characteristics, order effects and participant variables. These were clearly explained within the context of the scenario. Weaker responses did not make a judgement as to how validity would be affected, i.e. increased or decreased. A number of candidates gave answers focusing on the use of a lab experiment rather than the independent measures design (specifically making references to ecological validity), which was not creditworthy.</p>
		Total	4	
13		A		inter-rater
		Total	1	
14		C		$p < 0.025$
		Total	1	

15			C		Median <u>Examiner's Comments</u> A lot of candidates responded well to this question, however some candidates opted for D incorrectly.
			Total	1	
16			D		81 <u>Examiner's Comments</u> A lot of candidates responded well to this question, however some candidates opted for A incorrectly
			Total	1	
17			B		it is based on the data collected
			Total	1	
18			D		58
			Total	1	
19			D		an observation study which records whether customers say 'thank you' or not to the checkout assistant in a shop
			Total	1	
20			B	1	Peer-review <u>Examiner's Comments</u> Many answered correctly but there was still a large number of incorrect answers.
			Total	1	
21			B	1	Social desirability <u>Examiner's Comments</u> The majority answered correctly.
			Total	1	
22			D	1	Secondary <u>Examiner's Comments</u> Most answered correctly.

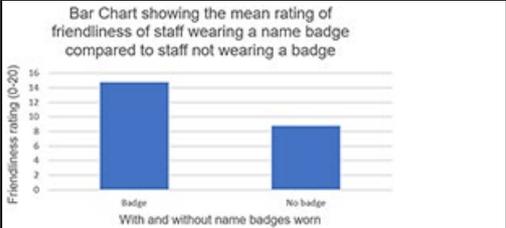
			Total	1	
23			D	1	29 Examiner's Comments Most answered correctly.
			Total	1	
24			D	1	Right to withdraw Examiner's Comments Nearly all answered correctly.
			Total	1	
25			<p>Identify <u>one</u> reason why the study might have low population validity.</p> <div style="border: 1px solid black; padding: 5px;"> <p>Answers could include:</p> <ul style="list-style-type: none"> • Low response rate. • People may be embarrassed and so not want to take part. • People may not complete the self-report. • Only people with certain characteristics might take part. • A small sample size might be collected. </div> <div style="border: 1px solid black; padding: 2px; margin-top: 5px;"> <p>One correct reason identified</p> </div> <div style="border: 1px solid black; padding: 2px; margin-top: 5px;"> <p>The candidate has not provided any creditworthy information</p> </div>	<p>Max 1 1 0</p>	<p>Context not needed here, but if used as part of an example provided by the candidate it is ok (context = mental health, grateful, feelings, etc.)</p> <p>Examiner's Comments</p> <p>Many candidates did not get this question correct; they did not identify an issue from the study in general and focused too heavily on the statistics given in the stem of '1 in 4 people experience mental health issues'. Candidates need to understand that this section on design and response is about a response to the stem, not a requirement for them to evaluate the stem.</p>
			Total	1	
26			<p>Outline how you would address <u>two</u> ethical considerations in the design of this study.</p> <div style="border: 1px solid black; padding: 5px;"> <p>Answers could include:</p> <ul style="list-style-type: none"> • Respect (consent, withdrawal and confidentiality) • Competence (having a researcher trained in mental health conduct the interview) • Responsibility (protection from harm and debrief) </div>	<p>Max 6 [3+3] 3 2 1 0</p>	<p>Context = mental health, grateful, feelings, etc.</p> <p>Candidates do not need to use the headings Respect, Competence, Responsibility, Integrity.</p> <p>Candidates must outline two separate ethical considerations.</p> <p>Examiner's Comments</p> <p>Candidates were often able to give two separate ethical considerations in context however often they didn't</p>

		<ul style="list-style-type: none"> Integrity (maintain high standards and no use of deception). 		explain 'how' these considerations would be employed.
		3 marks for each way to address an ethical consideration:		
		Clear outline in context		
		Attempted outline in context	OR Clear outline but not in context	
		Brief and/or weak outline (whether in context or not)		
		The candidate has not provided any creditworthy information		
		Total	6	
27		<p>Draw a fully labelled bar chart displaying the mean rating for how cute the puppy was perceived to be in each condition for this study.</p>  <p>1 mark is awarded for correctly calculating then presenting by value each bar representing the mean rating of how cute the puppy dog is perceived to be in each condition</p> <p>1 mark is awarded for named categories/clear labelling of the x axis</p> <p>1 mark is awarded for clear labelling of the y axis (must be contextualised to 'cuteness')</p> <p>1 mark is awarded for a clearly contextualised title</p>	<p>Max 4 [1+1+1+1]</p> <p>1 1 1 1</p>	<p>If a clearly contextualised title has been included, this could make up for any shortfall of labelling on an x and y axis.</p> <p>Candidates must make clear in the title that this bar chart is showing the 'mean'.</p> <p>Examiner's Comments</p> <p>Many candidates had a good understanding of how to draw a bar chart. Candidates often lost marks due to not fully operationalising their title or the label on their y axis. There was a significant number of candidates who presented all raw data on their graph rather than their mean.</p>
		Total	4	
28		<p>Outline <u>two</u> conclusions that can be made from the data collected in this study.</p> <p>Conclusions could include:</p> <ul style="list-style-type: none"> A puppy with big eyes is regarded as cuter than a puppy with smaller eyes perhaps because the larger eyes make us focus on the dog for longer which makes it more familiar. A puppy with big eyes is regarded as cuter than a puppy with smaller eyes as the larger 	<p>Max 6 [3+3]</p> <p>3 2 1 0</p>	<p>Context = puppy, dog, cute, etc.</p> <p>Clear (explicit) interpretation of findings (not simply stating a finding) is required to achieve beyond 1 mark.</p> <p>Examiner's Comments</p> <p>This question gained a variety of responses with few candidates accessing the full range of marks. Most candidates were restricted in their marks as they only identified</p>

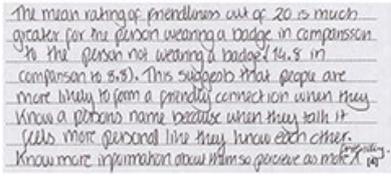
		<p>eyes are more distinctive and therefore perceived as more human like.</p> <p>Accept any other appropriate conclusions here.</p> <p>3 marks for each conclusion</p> <p>Clear, detailed conclusion in context</p> <table border="1" data-bbox="240 383 746 465"> <tr> <td>Clear, detailed conclusion but not in context</td> <td>OR Attempt at conclusion in context</td> </tr> </table> <table border="1" data-bbox="240 465 746 548"> <tr> <td>Brief and/or weak attempt (whether in context or not)</td> <td>OR Simply stating a finding</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	Clear, detailed conclusion but not in context	OR Attempt at conclusion in context	Brief and/or weak attempt (whether in context or not)	OR Simply stating a finding		<p>findings rather than giving conclusions.</p> <p> Assessment for learning</p> <p>Candidates would benefit from being given opportunities to examine raw data tables and practise extrapolating conclusions that could be drawn from them by looking at the nature of the data and any patterns that can be inferred. These skills can also be illustrated through data collected through practical investigations conducted in preparation for this exam.</p> <p>The difference in findings and conclusions from the core studies component could be highlighted to assist in demonstrating to candidates how to differentiate between findings and conclusions.</p>
Clear, detailed conclusion but not in context	OR Attempt at conclusion in context							
Brief and/or weak attempt (whether in context or not)	OR Simply stating a finding							
		<p>Total</p>	<p>6</p>					
<p>29</p>		<p>Outline <u>one</u> strength and <u>one</u> weakness of the use of quantitative data in this study.</p> <div style="border: 1px solid black; padding: 5px;"> <p><u>Strengths could include:</u></p> <ul style="list-style-type: none"> • Allows participants to be more precise about level of cuteness as the scale (1 – 20) is quite broad and has verbal labels at each end (not very cute and extremely cute). • Easy to analyse and compare findings on ratings of cuteness of puppy. • Easy to present findings and plot visually (e.g., in a bar chart displaying mean cuteness rating across each condition) therefore easier to compare. <p><u>Weaknesses could include:</u></p> <ul style="list-style-type: none"> • Numbers in the ratings lack detail therefore don't indicate why participants thought the puppy was cute or not. </div>	<p>Max 6 [3+3] 3 2 1 0</p>	<p>Context = puppy, dog, cute, etc.</p> <p><u>Examiner's Comments</u></p> <p>Candidates were able to access a wide range of marks on this question, with most accessing the highest band. Responses that couldn't access top band usually lacked context or elaboration on their evaluation point.</p>				

		<ul style="list-style-type: none"> Some participants may interpret the scale differently to others (e.g. one person's idea of what 7 on the scale is, might be different to another person's idea of what 7 indicates in terms of cuteness). 																
		<p>For each strength/weakness:</p> <p>Clear outline of strength/weakness in context</p> <table border="1"> <tr> <td>Attempted outline of strength/weakness in context</td> <td>OR Clear outline of strength/ weakness, but not in context</td> </tr> </table> <p>Brief and/or weak attempt to outline strength/weakness (whether in context or not)</p> <p>The candidate has not provided any creditworthy information</p>	Attempted outline of strength/weakness in context	OR Clear outline of strength/ weakness, but not in context														
Attempted outline of strength/weakness in context	OR Clear outline of strength/ weakness, but not in context																	
		Total	6															
30	a	<p>The inferential statistical test used to analyse the data collected in this study is the Wilcoxon Signed Ranks test. Explain how data is ranked for this test in this study.</p> <table border="1"> <tr> <td colspan="2">Ranking involves assigning data values that denote their position in an ordered sequence (i.e. putting the values in order from lowest to highest). For the data in this study, this involves ranking the difference between the ratings of cuteness for each condition (i.e. not ranking the scores/ratings themselves).</td> </tr> <tr> <td colspan="2">Clear explanation in context of what ranking involves and how it is done using the Wilcoxon test for the data in context.</td> </tr> <tr> <td colspan="2">Explanation of what ranking in general involves and explanation of how ranking is done using the Wilcoxon test, not in context.</td> </tr> <tr> <td colspan="2">OR Explanation of what ranking in general involves in context.</td> </tr> <tr> <td colspan="2">OR Explanation of how ranking is done using the Wilcoxon test in context</td> </tr> <tr> <td>Explanation of what ranking in general involves</td> <td>OR Explanation of how ranking is done using the Wilcoxon test</td> </tr> <tr> <td colspan="2">The candidate has not provided any creditworthy information</td> </tr> </table>	Ranking involves assigning data values that denote their position in an ordered sequence (i.e. putting the values in order from lowest to highest). For the data in this study, this involves ranking the difference between the ratings of cuteness for each condition (i.e. not ranking the scores/ratings themselves).		Clear explanation in context of what ranking involves and how it is done using the Wilcoxon test for the data in context.		Explanation of what ranking in general involves and explanation of how ranking is done using the Wilcoxon test, not in context.		OR Explanation of what ranking in general involves in context.		OR Explanation of how ranking is done using the Wilcoxon test in context		Explanation of what ranking in general involves	OR Explanation of how ranking is done using the Wilcoxon test	The candidate has not provided any creditworthy information		<p>Max 3</p> <p>3</p> <p>2</p> <p>1</p> <p>0</p>	<p>Context = puppy, dog, cute, etc.</p> <p>Convention is that the lowest score is assigned rank one and then proceed on this basis (but can be done vice versa (highest score is assigned rank 1)).</p> <p><u>Examiner's Comments</u></p> <p>Many candidates misunderstood this question and outlined why the Wilcoxon test would be used. A significant number of candidates were however able to answer this question effectively but often missed out context for the final mark.</p>
Ranking involves assigning data values that denote their position in an ordered sequence (i.e. putting the values in order from lowest to highest). For the data in this study, this involves ranking the difference between the ratings of cuteness for each condition (i.e. not ranking the scores/ratings themselves).																		
Clear explanation in context of what ranking involves and how it is done using the Wilcoxon test for the data in context.																		
Explanation of what ranking in general involves and explanation of how ranking is done using the Wilcoxon test, not in context.																		
OR Explanation of what ranking in general involves in context.																		
OR Explanation of how ranking is done using the Wilcoxon test in context																		
Explanation of what ranking in general involves	OR Explanation of how ranking is done using the Wilcoxon test																	
The candidate has not provided any creditworthy information																		
	b	<p>Using the Wilcoxon Signed Ranks test, calculate the value of T for the data collected in this study by completing the table below.</p>	<p>Max 3</p> <p>1+1+1</p> <p>1</p> <p>1</p>	<p>If candidate has the correct answer but not shown any workings, award 1 mark.</p>														

	<p>Formula / process: T = the sum of the ranks of the differences with the least frequent sign</p> <table border="1" data-bbox="240 315 746 528"> <thead> <tr> <th>Participant</th> <th>Condition (a) with big eyes</th> <th>Condition (b) with small eyes</th> <th>Difference (a) - (b)</th> <th>Ranks of differences</th> <th>T value (sum of the ranks of the differences with the least frequent sign)</th> </tr> </thead> <tbody> <tr> <td>a</td> <td>18</td> <td>7</td> <td>11</td> <td>5.5</td> <td rowspan="6">1 + 2 = 3 T = 3</td> </tr> <tr> <td>b</td> <td>14</td> <td>3</td> <td>11</td> <td>5.5</td> </tr> <tr> <td>c</td> <td>12</td> <td>14</td> <td>-2</td> <td>1</td> </tr> <tr> <td>d</td> <td>17</td> <td>20</td> <td>-3</td> <td>2</td> </tr> <tr> <td>e</td> <td>19</td> <td>9</td> <td>10</td> <td>4</td> </tr> <tr> <td>f</td> <td>16</td> <td>10</td> <td>6</td> <td>3</td> </tr> </tbody> </table> <p>1 mark for the correct answer (T=3)</p> <p>1 mark for Ranks of differences</p> <p>1 mark for working out the Difference (a) - (b)</p> <p>The candidate has not provided any creditworthy information</p>	Participant	Condition (a) with big eyes	Condition (b) with small eyes	Difference (a) - (b)	Ranks of differences	T value (sum of the ranks of the differences with the least frequent sign)	a	18	7	11	5.5	1 + 2 = 3 T = 3	b	14	3	11	5.5	c	12	14	-2	1	d	17	20	-3	2	e	19	9	10	4	f	16	10	6	3	<p>1 0</p>	<p>Examiner's Comments</p> <p>Some candidates were able to achieve full marks on this question. Many candidates were able to complete the difference column but nothing further. Those achieving 2 marks were normally able to complete the difference and get the final T value correct - however their ranking was often incorrect due to ranking minus 3 as 1 and minus 2 as 2.</p> <p> Assessment for learning</p> <p>Candidates would benefit from having opportunities to calculate practice inferential statistical tests. It would also be useful to remind them that in a Wilcoxon Signed Ranks test they are to ignore the minus signs when ranking the data.</p>
Participant	Condition (a) with big eyes	Condition (b) with small eyes	Difference (a) - (b)	Ranks of differences	T value (sum of the ranks of the differences with the least frequent sign)																																			
a	18	7	11	5.5	1 + 2 = 3 T = 3																																			
b	14	3	11	5.5																																				
c	12	14	-2	1																																				
d	17	20	-3	2																																				
e	19	9	10	4																																				
f	16	10	6	3																																				
	<p>Total</p>	<p>6</p>																																						
<p>31</p>	<p>Outline one reason why the reliability of the data collected in this study could be low.</p> <p><u>Answers could include:</u></p> <ul style="list-style-type: none"> Some participants interpreting the rating scale for cuteness differently to others. If a different type/breed of dog was used in the photo with big eyes compared to the photo with small eyes. Type of people used in the sample (if dog owners or not) because they might interpret the photos differently. <p>Clear outline in context</p> <table border="1" data-bbox="240 1771 746 1850"> <tr> <td>Clear outline, but not in context</td> <td>An attempt in context</td> </tr> </table> <p>Brief and/or weak outline, whether in context or not</p> <p>The candidate has not provided any creditworthy information</p>	Clear outline, but not in context	An attempt in context	<p>Max 3 3 2 1 0</p>	<p>Context = puppy, dog, cute, etc.</p> <p>To score above 1 mark, the candidate must show an understanding of the concept of reliability (not just giving a generic example that could also affect the validity).</p> <p>No credit for reference to population validity or generalisability.</p> <p>Examiner's Comments</p> <p>Many candidates did not achieve marks on this question, often confusing reliability with validity. The more common incorrect responses focused on social desirability bias/population validity.</p> <p>Exemplar 3</p>																																			
Clear outline, but not in context	An attempt in context																																							

					<p>Reliability refers to how replicable a study is. In this study on whether eye size affects behaviour, it could be low because it's not conducted in a controlled setting, so other future research may not imitate the same setting which reduces reliability.</p> <p>Exemplar 3 demonstrates one possible 3 mark answer to this question. This candidate has identified the control of a setting and how if it is not kept the same then this would reduce the reliability of the data collection. Their response is also in context.</p>
			Total	3	
32			C	1	<p>Whether the measurement of a variable can predict outcomes for similar behaviour in the future</p> <p>Examiner's Comments</p> <p>Most answered this correctly.</p>
			Total	1	
33			D	1	<p>22%</p> <p>Examiner's Comments</p> <p>Nearly all answered this correctly</p>
			Total	1	
34			D	1	<p>Parametric statistical tests</p> <p>Examiner's Comments</p> <p>A number of candidates were unable to give a correct response.</p>
			Total	1	
35	a		<p>Bar Chart showing the mean rating of friendliness of staff wearing a name badge compared to staff not wearing a badge</p>  <p>1 mark is awarded for correct presentation of data 1 mark is awarded for clear labelling of the x axis 1 mark is awarded for clear labelling of the y axis including measurement must start at 0 (does not need to go up to 20) 1 mark is awarded for fully operationalised title</p>	Max 4	<p>Context = friendliness, friend, name, badge etc</p> <p>Title must include both variables ([mean] rating of friendliness and name badges/no name badges). Response must make it clear that this is the mean rating of friendliness in either title or y axis, if not max 3</p> <p>Labels on axes must be clear. X axis – badge/not wearing badge (or similar wording) Y axis – rating of friendliness and</p>

		<table border="1"> <tr> <td>All 4 features from above - 4</td> </tr> <tr> <td>Any 3 features from above - 3</td> </tr> <tr> <td>Any 2 features from above - 2</td> </tr> <tr> <td>Any 1 feature from above - 1</td> </tr> <tr> <td>The candidate has not provided any creditworthy information - 0</td> </tr> </table>	All 4 features from above - 4	Any 3 features from above - 3	Any 2 features from above - 2	Any 1 feature from above - 1	The candidate has not provided any creditworthy information - 0	<p>measurement must start at 0 (does not need to go up to 20 and can go beyond 20).</p> <p>If two bars are together do not award mark for correct presentation of data.</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates plotted the data accurately and labelled the X axis correctly. Some responses did not get the mark for the title as either it was missing or did not include both conditions of the independent variable in the title or not mentioning mean rating of friendliness. Some responses did not get the mark for plotting the data as it was either incorrect or the bars of the bar chart were placed together. The majority of candidates achieved the mark for the Y axis although a few did just state 'mean' or 'mean rating' which did not achieve the mark for the Y axis.</p>					
All 4 features from above - 4													
Any 3 features from above - 3													
Any 2 features from above - 2													
Any 1 feature from above - 1													
The candidate has not provided any creditworthy information - 0													
b		<p>For example: Staff wearing a name badge were perceived as more friendly compared to those not wearing a name badge. This is perhaps because it was more personal and made customers feel that they knew the person better and could interact and relate to them more.</p> <table border="1"> <tr> <td colspan="2">Clear conclusion in context that refers to the findings. - 4</td> </tr> <tr> <td>Attempted conclusion in context that refers to the findings. - 3</td> <td>Clear conclusion in context and attempt to refer to findings. - 3</td> </tr> <tr> <td>Attempted conclusion in context. - 2</td> <td>Clear conclusion not in context - 2</td> </tr> <tr> <td colspan="2">Unclear and/or brief conclusion (whether in context or not) - 1</td> </tr> <tr> <td colspan="2">The candidate has not provided any creditworthy information - 0</td> </tr> </table>	Clear conclusion in context that refers to the findings. - 4		Attempted conclusion in context that refers to the findings. - 3	Clear conclusion in context and attempt to refer to findings. - 3	Attempted conclusion in context. - 2	Clear conclusion not in context - 2	Unclear and/or brief conclusion (whether in context or not) - 1		The candidate has not provided any creditworthy information - 0		<p>Context = friendliness, friend, name, badge etc</p> <p>Note: a conclusion must be an interpretation/application of the findings / data (not simply a statement of the result(s) obtained)</p> <p>No credit for stating just the mean scores (e.g. the mean friendliness score for wearing a badge is $14.8=0$)</p> <p>Max 4</p> <p>Max 2 marks if only findings presented. For example - Staff wearing a name badge were perceived as more friendly (1) compared to staff not wearing a name badge(1).</p> <p>Needs to state the direction of the findings. If not, max 1 if just findings presented e.g. there is a difference in perceived friendliness rating if</p>
Clear conclusion in context that refers to the findings. - 4													
Attempted conclusion in context that refers to the findings. - 3	Clear conclusion in context and attempt to refer to findings. - 3												
Attempted conclusion in context. - 2	Clear conclusion not in context - 2												
Unclear and/or brief conclusion (whether in context or not) - 1													
The candidate has not provided any creditworthy information - 0													

				<p>the staff wore a badge or not.</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates were able to make a brief statement about what the results show (i.e. that shop assistants wearing a name badge are perceived as more friendly than those who don't) but many struggled to elaborate beyond the findings. Better responses tended to say that knowing someone's name means you feel you know them more personally or you feel more comfortable to speak with the shop assistant.</p> <p> Assessment for learning</p> <p>It would be helpful to candidates to practise differentiating between a finding and a conclusion. This could be practised through the core studies where the results of the study (the findings) can be given and then the conclusions outlined (what can be inferred from the results).</p> <p>Exemplar 3</p> <p></p> <p>Exemplar 3 shows a response that received full marks. The response outlines the difference in the findings but then makes inferences about these differences and outlines why participants may have found a shop assistant wearing a name badge to be more friendly.</p>
			Total	8

36		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td colspan="2" style="padding: 5px;">Because the data contains outliers (e.g. in the wearing badge condition the rating of 5, which is much lower than any other rating) and the median is less sensitive measure of central tendency when there are outliers.</td> </tr> <tr> <td colspan="2" style="padding: 5px;">Clear explanation in context - 2</td> </tr> <tr> <td style="padding: 5px; width: 50%;">Attempted explanation in context - 1</td> <td style="padding: 5px; width: 50%;">Clear explanation no context - 1</td> </tr> <tr> <td colspan="2" style="padding: 5px;">The candidate has not provided any creditworthy information - 0</td> </tr> </table>	Because the data contains outliers (e.g. in the wearing badge condition the rating of 5, which is much lower than any other rating) and the median is less sensitive measure of central tendency when there are outliers.		Clear explanation in context - 2		Attempted explanation in context - 1	Clear explanation no context - 1	The candidate has not provided any creditworthy information - 0		Max 2	<p>Context = friendliness, friend, name, badge etc OR 5 AND/OR 18</p> <p>1 mark response – identifying the anomaly/extreme value 2 marks – example from the data (5 or 18) and reference to the median not being affected by anomalies (and the mean is).</p> <p>Award credit to - Median is the most appropriate with ordinal data. (1) and reason why- This is because ordinal data is a less precise measurement than interval data.(1) OR Median is less sensitive and is therefore more appropriate for ordinal data.(1)</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates were able to achieve 1 mark by correctly identifying that outlier/extreme scores would affect the mean but not the median. Better responses then identified the extreme score in the data (5 or 18). However, many responses did not achieve this second mark for not giving examples of these extreme scores within the data set (5 or 18). It was common for responses to identify an incorrect extreme score such a one of the rankings of friendliness rather than the ratings of friendliness. A common error in responses that did not gain marks was to state that the median is better due to the large range of data.</p>
Because the data contains outliers (e.g. in the wearing badge condition the rating of 5, which is much lower than any other rating) and the median is less sensitive measure of central tendency when there are outliers.												
Clear explanation in context - 2												
Attempted explanation in context - 1	Clear explanation no context - 1											
The candidate has not provided any creditworthy information - 0												
Total			2									
37	a	$U1 = 135 - 110/2 = 80$ $U2 = 75 - 110/2 = 20 \text{ (U1 and U2 are inter-changeable in terms of order candidates choose to do them)}$ $U = 20$	Max 5	<p>= 20 is creditworthy (rather than U=20) or circling the correct U value.</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates responded to this question well and identified the correct U value and</p>								

		<p>1 mark for each of the following correct / evident in answer .</p> <ul style="list-style-type: none"> • 135 • 75 • correct calculation of U2 value obtained OR correct calculation of U1 value • All workings shown for U2 • Correct U value (20) <table border="1"> <tr> <td>All 5 features from above - 5</td> </tr> <tr> <td>4 features from above - 4</td> </tr> <tr> <td>3 features from above - 3</td> </tr> <tr> <td>2 features from above - 2</td> </tr> <tr> <td>1 feature from above - 1</td> </tr> <tr> <td>The candidate has not provided any creditworthy information - 0</td> </tr> </table>	All 5 features from above - 5	4 features from above - 4	3 features from above - 3	2 features from above - 2	1 feature from above - 1	The candidate has not provided any creditworthy information - 0		<p>showed their workings fully. Errors on this question typically stemmed from incorrectly adding the ranks for the conditions and therefore coming out with the incorrect U value.</p> <p>Other errors in terms of the analysis came from using the ratings, not the rankings, to find the R1 and R2 values.</p> <p>A few responses carried out the calculation but did not identify U=20 (as the smallest of the two calculated values) so were unable to achieve full marks for this question. There were some who did not attempt this question.</p>				
All 5 features from above - 5														
4 features from above - 4														
3 features from above - 3														
2 features from above - 2														
1 feature from above - 1														
The candidate has not provided any creditworthy information - 0														
	b	<table border="1"> <tr> <td colspan="2">Found by using the sample size in each condition</td> </tr> <tr> <td colspan="2">Table critical value = 23</td> </tr> <tr> <td colspan="2">Explanation provided and correct critical value stated - 2</td> </tr> <tr> <td>Explanation only provided - 1</td> <td>OR critical value only provided - 1</td> </tr> <tr> <td colspan="2">The candidate has not provided any creditworthy information - 0</td> </tr> </table>	Found by using the sample size in each condition		Table critical value = 23		Explanation provided and correct critical value stated - 2		Explanation only provided - 1	OR critical value only provided - 1	The candidate has not provided any creditworthy information - 0		Max 2	<p>Credit N1/N2 or Na/Nb</p> <p>Credit Na = 10; Nb = 10 as the explanation</p> <p>1 mark for indicating the critical value in the table and not in the answer space</p> <p><u>Examiner's Comments</u></p> <p>Candidates demonstrated good understanding of critical value tables and were able to identify 23 as the critical value and explain how this was found (by using the number of participants, 10, in each condition). A common error was reading from 5 (for the 5% level of probability) and then either 20 (total participants) or 10 (participants in each condition) instead of 10 and 10.</p>
Found by using the sample size in each condition														
Table critical value = 23														
Explanation provided and correct critical value stated - 2														
Explanation only provided - 1	OR critical value only provided - 1													
The candidate has not provided any creditworthy information - 0														
	c	<table border="1"> <tr> <td> <p>U = 20, n1 = 10, n2 = 10, p<0.05 OR U=20, critical value = 23, 20<23, p<0.05</p> <p>OR In words ... e.g. the alternative hypothesis (that there would be/is a significant difference in perceived friendliness between</p> </td> </tr> </table>	<p>U = 20, n1 = 10, n2 = 10, p<0.05 OR U=20, critical value = 23, 20<23, p<0.05</p> <p>OR In words ... e.g. the alternative hypothesis (that there would be/is a significant difference in perceived friendliness between</p>	Max 2	<p>Credit N1/N2 or Na/Nb</p> <p>Cannot credit any p value other than .05 or 5%</p> <p>For worded statements 1 mark for each of the following:</p>									
<p>U = 20, n1 = 10, n2 = 10, p<0.05 OR U=20, critical value = 23, 20<23, p<0.05</p> <p>OR In words ... e.g. the alternative hypothesis (that there would be/is a significant difference in perceived friendliness between</p>														

			<p>shop assistants wearing a badge and those who don't) is supported. This is because the calculated value (20) is less than the critical value (23).</p> <table border="1"> <tr> <td>Correctly written significance statement - 2</td> <td>OR written in words rather than a formal statement - 2</td> </tr> <tr> <td>Just stating $p < 0.05$ - 1</td> <td>OR weak and/or brief written response - 1</td> </tr> </table> <p>The candidate has not provided any creditworthy information - 0</p>	Correctly written significance statement - 2	OR written in words rather than a formal statement - 2	Just stating $p < 0.05$ - 1	OR weak and/or brief written response - 1	<ul style="list-style-type: none"> • Comparison between calculated value and critical value • Significance level ($p < 0.05$)/'it is significant'/alternative hypothesis is supported/null hypothesis is rejected <p>Max 1 mark if candidate contradicts themselves about whether it is significant but does state the correct comparison of calculated value and critical value.</p> <p><u>Examiner's Comments</u></p> <p>The vast majority of candidates presented a written statement rather than a mathematical one and correctly compared the critical and calculated value and stated there was a significant difference. Some candidates did not make a comparison between the calculated and critical value so could not achieve full marks for this question. The most common errors were from earlier mistakes being carried forward e.g. incorrect U value, and incorrect critical value that led to an incorrect significance statement in the response. A few candidates did have the correct response in Question 30 (a) and Question 30 (b) but then concluded that the results were not significant which was incorrect.</p>
Correctly written significance statement - 2	OR written in words rather than a formal statement - 2							
Just stating $p < 0.05$ - 1	OR weak and/or brief written response - 1							
		Total	9					
38		<p>Likely answers: lacks depth so we don't know reasons why wearing a name badge or not influences perceived friendliness; doesn't allow for consideration of other (extraneous) influences (e.g. physical features of person etc); May lack ecological validity as not how we would consider friendliness i.e. in number format</p> <table border="1"> <tr> <td colspan="2">Clear and precise explanation in context - 3</td> </tr> <tr> <td>Clear explanation but not in context - 2</td> <td>OR attempted explanation in context - 2</td> </tr> </table>	Clear and precise explanation in context - 3		Clear explanation but not in context - 2	OR attempted explanation in context - 2	Max 3	<p>Context = friendliness, friend, name, badge etc</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates focused on the difficulty of not being able to find out why shop assistants with badges were rated more friendly. Better responses explained why this was a problem such as lack of detail/depth when only quantitative data is collected and the effect this</p>
Clear and precise explanation in context - 3								
Clear explanation but not in context - 2	OR attempted explanation in context - 2							

		<p>Brief and/or weak attempt (whether in context or not) - 1</p> <p>The candidate has not provided any creditworthy information - 0</p>		<p>could have on the validity or usefulness of the results. Some candidates focused on the issue of not being able to rule out other extraneous variables, such as whether the shop assistant was smiling or their general behaviour towards participants and this was often done well. Weaker responses often did not explain why this extra insight is a benefit of qualitative data or a weakness of excluding it in this study.</p>
		Total	3	
39		<p>Credit answers which explain how reliability was maintained, increased or reduced.</p> <p>Likely answers: Consistency may be reduced by: Participants variables may produce an inconsistent/subjective view of the shop assistants who may be perceived as more friendly than others by different participants in different conditions.</p> <p>Different physical features of the shop assistants may be perceived as more friendly than others by different participants</p> <p>used in the badge wearing / non-badge wearing conditions.</p> <p>Different interpretations of the 0-20 rating scale for friendliness by different participants.</p> <p>Different behaviour of the shop assistants (this would probably be standardised as it was a research set up) when interacting with customers (and how long spent in their company, etc).</p> <p>Social desirability bias could have affected participants differently, (make sure it is linked clearly to lack of consistency, not accuracy).</p> <p>Consistency could have been increased</p>	3 + 3	<p>Context = friendliness, friend, name, badge, shop assistants etc</p> <p>Annotation – up to 3 ticks per factor</p> <p>No credit for reference to population validity or generalisability.</p> <p>Credit responses that refer to replicability.</p> <p><u>Examiner’s Comments</u></p> <p>Most responses were able to identify at least one factor that could have affected the reliability of the data collected in this study. Common creditworthy responses focused on the rating scale being subjective (with different participants interpreting the scale differently), the shop assistants not acting in a consistent way, or following a standardised procedure, participant variables and the environment of the shop. However, the explanations were not always well developed with quite a bit of confusion between validity and reliability. Weaker responses were often brief with just a statement that the factor would affect either replicability or consistency of ratings without any explanations. Some responses would discuss issues that related to</p>

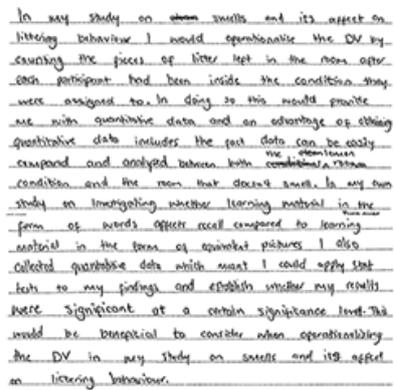
		<p>by:</p> <p>Standardised question increasing consistency as pre-set scale given.</p> <p>Shop assistants are likely to behave in a similar way with each customer so this could be considered standardised behaviour of shop assistants allowing replicability.</p> <p>Quantitative data is an objective measure so there is no inconsistency in the researcher interpreting the 0-20 ratings given by the participant.</p> <p>Up to 3 marks for each factor</p> <table border="1" style="width: 100%;"> <tr> <td colspan="2">Identification of a relevant factor with a clear and precise explanation in context - 3</td> </tr> <tr> <td>Identification of a relevant factor with a clear explanation but not in context - 2</td> <td>Identification of a relevant factor with an attempt at an explanation in context - 2</td> </tr> <tr> <td colspan="2">Identification of a relevant (contextual) factor but no/very weak explanation - 1</td> </tr> </table>	Identification of a relevant factor with a clear and precise explanation in context - 3		Identification of a relevant factor with a clear explanation but not in context - 2	Identification of a relevant factor with an attempt at an explanation in context - 2	Identification of a relevant (contextual) factor but no/very weak explanation - 1			<p>validity, and wouldn't explain why it related to consistency in sufficient detail (e.g. demand characteristics, socially desirable responses). There were a significant number of responses that explained issues which were not creditworthy e.g. generalisability, or confuse the stem and suggest that the shop assistants were rating how friendly the customers were. Many of the responses were an attempt in context and not properly developed to get full marks.</p> <p> Misconception</p> <p>Many responses for this question were confused between reliability and validity.</p>
Identification of a relevant factor with a clear and precise explanation in context - 3										
Identification of a relevant factor with a clear explanation but not in context - 2	Identification of a relevant factor with an attempt at an explanation in context - 2									
Identification of a relevant (contextual) factor but no/very weak explanation - 1										
		Total	6							
40		A		<p>a type 1 error</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates responded well to this question. Some candidates chose option B incorrectly.</p>						
		Total	1							
41		A		<p>dispersion around the mean</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates responded well to this question. Incorrect responses were varied.</p>						
		Total	1							
42		C		<p>35%</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates</p>						

					responded well to this question. Some candidates chose option D incorrectly.
			Total	1	
43			B		1/20 <u>Examiner's Comments</u> The majority of candidates responded well to this question. Some candidates chose option C incorrectly.
			Total	1	
44			B		3.5 <u>Examiner's Comments</u> The majority of candidates responded well to this question. Some candidates chose option C incorrectly.
			Total	1	
45			C		normal <u>Examiner's Comments</u> The majority of candidates responded well to this question. Incorrect responses were varied.
			Total	1	
46			B		Mann-Whitney U test <u>Examiner's Comments</u> The majority of candidates responded well to this question. Some candidates chose option C or D incorrectly.
			Total	1	
47			D		25.90 <u>Examiner's Comments</u> The majority of candidates responded well to this question. Some candidates chose option C incorrectly.

			Total		1									
48			A			concurrent Examiner's Comments The majority of candidates responded well to this question. Incorrect responses were varied.								
			Total		1									
49			C			mode Examiner's Comments The majority of candidates responded well to this question. Some candidates chose option B incorrectly.								
			Total		1									
50			<p>What you are being driven by is the left-hand column of the grid ('details of the required features (RFs)'. That is always your starting point and 'locator' for the appropriate mark band before considering the other two columns ('justification of decisions made' and 'reference to own practical work').</p> <table border="1"> <thead> <tr> <th>Level of response</th> <th>Details of required features (RFs) included</th> <th>Justification of decisions made</th> <th>Reference to own practical work</th> </tr> </thead> <tbody> <tr> <td>Good 12-15 marks</td> <td>All 4 required features (RFs) addressed in context Accurate and detailed knowledge and understanding of each feature in context Good evidence of</td> <td>Appropriate justification of all decisions and <i>some</i> is contextualised Well-developed line of reasoning that is clear and logically structured</td> <td>Explicit reference to own practical work and clear links between own work and the planned research for each required feature. e.g. specific mention of aim or procedural features For top band</td> </tr> </tbody> </table>	Level of response	Details of required features (RFs) included	Justification of decisions made	Reference to own practical work	Good 12-15 marks	All 4 required features (RFs) addressed in context Accurate and detailed knowledge and understanding of each feature in context Good evidence of	Appropriate justification of all decisions and <i>some</i> is contextualised Well-developed line of reasoning that is clear and logically structured	Explicit reference to own practical work and clear links between own work and the planned research for each required feature. e.g. specific mention of aim or procedural features For top band		Max = 15	<p>Context = aroma, smell, scent, fragrance, litter, any example of litter (e.g. wrappers, rubbish) cleanliness etc</p> <p>Annotations</p> <p>RF (in the left column AND see next page for descriptors of the levels for description of the RF) L4=Good; L3=Reasonable; L2=Limited; L1= Basic</p> <p>Annotate with CONT for context if RF in context. (under RF level annotation on left)</p> <p>Tick for justification within the response Do not annotate the level, note the level of justification to decide on the mark given within the band</p> <p>Examiner's Comments</p> <p>Responses varied a lot to this extended question, with the majority of candidates finding it difficult to gain the higher band marks. The best responses were characterised by taking each of the four required features in turn and</p>
Level of response	Details of required features (RFs) included	Justification of decisions made	Reference to own practical work											
Good 12-15 marks	All 4 required features (RFs) addressed in context Accurate and detailed knowledge and understanding of each feature in context Good evidence of	Appropriate justification of all decisions and <i>some</i> is contextualised Well-developed line of reasoning that is clear and logically structured	Explicit reference to own practical work and clear links between own work and the planned research for each required feature. e.g. specific mention of aim or procedural features For top band											

			<p>of required features</p> <p>OR more than one of the required features referred to but in a very brief and/or basic way</p>	<p>min 2 limited (L2) in context or 3-4 limited (L2) with no context</p> <p>L1 Basic – 1 basic (L1) (no context needed).</p> <p>THEN look at justifications Make judgement of which mark to give the response within the band based on the justifications</p> <p>L4 Good – at least 2 reasonable (L3) AND at least 2 of the justifications are in context (does not have to be the reasonable ones)</p> <p>L3 Reasonable – at least 2 limited (L2) AND at least 1 of the justifications is in context (8 marks if none contextualised OR meets the minimum justification requirement).</p> <p>L2 Limited – at least 1 limited (L1) (none have to be in context) OR If one required feature</p>	<p>some did not use acceptable strengths of the sampling technique. For example, stating that opportunity sampling method would give a representative sample. There was great variation in the response to RF2. Better responses gave a clear outline of how the dependent variable of 'amount of litter' would be operationalised. Such as 'counting the pieces of litter left in the room on a tally chart' or 'weighing the amount of litter left in the room in grams'. Sometimes the operationalisation of dependent variable lacked when and where the collection of data would occur or where the litter came from which sometimes limited the response to the 'reasonable' rather than 'good' band for this RF. Common justifications for this RF included making comparisons, use of bar charts and doing statistical analysis. Some responses gave weaknesses of their dependent variable which did not gain marks. Some justifications were contextualised but this was inconsistent in the responses. Candidates used a wide variety of ethical considerations in their response to RF3. Popular responses including debriefing which told the aim of the study and revealed any deceptions used as to the nature of the study, confidentiality and allowing the participants to withdraw their data. Giving informed consent was another successful response. Clear justifications contextualised the response and often referred to respect for participants or the necessity to follow the BPS guidelines. Weaker responses to this RF did not address the assessment request to address one ethical issue and often attempted to address many ethical issues. This led to less detail being given to each ethical issue raised in the response as well as difficulty in justifying any one ethical issue in</p>
--	--	--	---	---	---

			<p>addressed in detail (good) and justified in context and explicit links made to own practical work award 4 marks</p> <p>L1 Basic – no justification or basic justification</p>					
			<table border="1"> <thead> <tr> <th data-bbox="231 683 271 728">RF</th> <th data-bbox="271 683 774 728">Details of RF</th> </tr> </thead> <tbody> <tr> <td data-bbox="231 728 271 1960">1</td> <td data-bbox="271 728 774 1960"> <p>Sampling technique</p> <ul style="list-style-type: none"> • Good – Identified the sampling method and clearly explained where and how this has been carried out in their study. Details of how the sampling method is enacted/procedural details e.g. how the P is contacted, is approached or gets in touch with experimenter or becomes part of the sampling pool (e.g using everyone in that location). • Reasonable – Identified the sampling method, possibly defined AND reasonable attempt to explain how this has been carried out in their study. • Limited – Sampling method identified and defined OR unclear attempt to explain how this has been carried out in their study. • Basic – Just identifying the sampling technique </td> </tr> </tbody> </table>	RF	Details of RF	1	<p>Sampling technique</p> <ul style="list-style-type: none"> • Good – Identified the sampling method and clearly explained where and how this has been carried out in their study. Details of how the sampling method is enacted/procedural details e.g. how the P is contacted, is approached or gets in touch with experimenter or becomes part of the sampling pool (e.g using everyone in that location). • Reasonable – Identified the sampling method, possibly defined AND reasonable attempt to explain how this has been carried out in their study. • Limited – Sampling method identified and defined OR unclear attempt to explain how this has been carried out in their study. • Basic – Just identifying the sampling technique 	<p>detail. Some responses were in context, but this was less consistent compared to RF1 and RF2. The ethical consideration of confidentiality sometimes stated exactly how the researcher would keep the data from the participant confidential, such as using numbers in place of participants' names. However, many responses that had confidentiality as the ethical issue, often just stated they would keep it anonymous with no further explanation as to how they would do this. Some of the justifications were very brief, sometimes just stating it made the study 'more ethical'.</p> <p>There was a wide variety of suggestions given in responses to RF4. The best responses often detailed control of situational variables with a recognition that keeping the lemon smell/no smell in the rooms consistent for each participant was important. Popular ideas including opening windows and not allowing participants to wear any form of perfume during the study. Other ideas that often achieved well standardised the environment by having the layout of the rooms, location of the bins and time spent in rooms standardised among participants. There were also some good responses about participant variables with a recognition that anyone who had problems with smell or was allergic to lemons should be excluded from the study. Weaker responses often just identified the extraneous variable to be controlled or gave a very brief outline of how to control. Some responses were contextualised while others were not. Justifications were often linked to increased validity and/or reliability and some were able to give a clear explanation of why the control would lead to this improvement in the study. Better responses had specific evaluation points included. Weaker candidates</p>
RF	Details of RF							
1	<p>Sampling technique</p> <ul style="list-style-type: none"> • Good – Identified the sampling method and clearly explained where and how this has been carried out in their study. Details of how the sampling method is enacted/procedural details e.g. how the P is contacted, is approached or gets in touch with experimenter or becomes part of the sampling pool (e.g using everyone in that location). • Reasonable – Identified the sampling method, possibly defined AND reasonable attempt to explain how this has been carried out in their study. • Limited – Sampling method identified and defined OR unclear attempt to explain how this has been carried out in their study. • Basic – Just identifying the sampling technique 							

			or confuses sampling methods.			
		2	Operationalising DV	<ul style="list-style-type: none"> • Good – Clear details on how dependent variable will be operationalised. Outline how data is quantified and how/when the litter is counted (e.g. how weight is measured or unit of measure, comparison of amount before/after, number of pieces in the bin, tally of number of pieces of litter). • Reasonable – Reasonable details on how dependent variable will be operationalised that does lead to quantitative data e.g. count the number of pieces of litter/number of wrappers. May include a muddled/vague indication of where and how this has been carried out in their study. • Limited – Way DV is operationalised is quantitative and addressed in a limited/unclear way. E.g. does not indicate which litter is being counted, more than one measure indicated; Could indicate where the litter has come from and 'amount of litter left in the room'. • Basic – Vague indication of how DV would be measured (e.g. amount of litter left in the room). 	<p>justified by suggesting this would improve validity and/or reliability with no further elaboration. Links to own practical work were usually explicit with the responses outlining what they did in their study. Some candidates linked their own practical work to the choice of the RF while others just simply described what they did in their study. Many were able to further justify their choice of RF through their choices made in their own practical work.</p> <p>Exemplar 2</p>  <p>In my study on smell smells and its affect on littering behaviour I would operationalise the DV by counting the pieces of litter left in the room after each participant had been inside the condition they were assigned to. In doing so this would provide me with quantitative data and an advantage of obtaining quantitative data includes the fact data can be easily compared and analysed between the conditions the condition and the room that do not smell. In my own study on investigating whether learning material in the form of words affect recall compared to learning material in the form of equivalent pictures I also collected quantitative data which meant I could apply that test to my findings and establish whether my results were significant at a certain significance level. This would be beneficial to consider when operationalising the DV in my study on smell and its affect on littering behaviour.</p>	
		3	One ethical consideration addressed	Integrity (deception) Respect (privacy/confidentiality/consent/right to withdraw), Responsibility (debrief/no psychological or physical harm)		<p>Exemplar 2 is an example of a candidate who has addressed RF2 reasonably and in context and provided some reasonable justification but in context and makes explicit reference to their practical work. This candidate achieved a mark in the reasonable band overall as the RFs were one good in context, two reasonable in context and one basic in context. The justifications for their response were mixed with some good and some reasonable and mainly in context.</p>

			<p>unlikely to be creditworthy unless clear how their study could be psychologically or physically harmful),</p> <p>Competence (refers to the competence of the researcher e.g. get an expert in to check for sensory impairments)</p> <ul style="list-style-type: none"> • Good – Identifying the ethical consideration, explaining the ethical consideration and clarity on how it can be addressed. • Reasonable – Identifying the ethical consideration and reasonable explanation of how it can be addressed. Briefly addressed but lacks clarity. • Limited - Limited explanation with some understanding of the ethical consideration (e.g. outline of how to address ethical consideration possibly without identifying). • Basic – Just identifies the ethical consideration. <p>If candidate clearly does more than one consideration, credit the first one. Allow ethical considerations which are clearly connected to each other/influence each other.</p> <p>There may be a mislabelling of the ethical principle. Therefore, if the RF meets the requirements of the description, it can be put at this level.</p> <p>This RF needs to focus on the way the ethical consideration is addressed. Any information given on the reason is justification.</p>		
--	--	--	--	--	--

			<p>Control of one 4 Extraneous variable</p> <ul style="list-style-type: none"> • Good – Clear and somewhat detailed of how EV can be controlled. • Reasonable – Reasonable outline of how EV can be controlled. • Limited – Limited/brief outline of how EV can be controlled is unclear. • Basic – Identifies how EV can be controlled or is muddled. <p>If more than one extraneous variable, credit the first one.</p>		
Total				15	
51	a		<p>Ranking data refers to assigning numbers to denote position in an ordered sequence (lowest to highest or vice versa)</p> <p>Clear explanation - 2</p> <p>Attempted explanation/e.g. how to do ranking - 1</p> <p>The candidate has not provided any creditworthy information - 0</p>	Max 2	<p>For example - 1 mark – order the scores/numbers lowest to highest; highest to lowest; OR numerical order 2nd mark –explaining how the ranks are assigned e.g. 1 for the lowest, up to 20 for the highest OR explaining what ranking the data means</p> <p><u>Examiner’s Comments</u></p> <p>The majority of candidates showed knowledge of ranking that involves putting the data in numerical order from lowest to highest. Better responses showed knowledge of how to rank (i.e. lowest number is rank of 1 etc.)</p>
	b		<p>Because there are three friendliness ratings of 18, covering ranks 16, 17 and 18 collectively. Therefore, the ranks must be shared ($16 + 17 + 18 = 51/3 = 17$)</p> <p>Clear explanation - 2</p> <p>Attempted explanation - 1</p> <p>The candidate has not provided any creditworthy information - 0</p>	Max 2	<p>1 mark for identifying that identical numbers need to have the same rank OR identifying that there are three ratings of 18 2nd mark for how you calculate what that rank should be (mean or median or showing how this is done)</p> <p><u>Examiner’s Comments</u></p> <p>The majority of candidates identified that there were three</p>

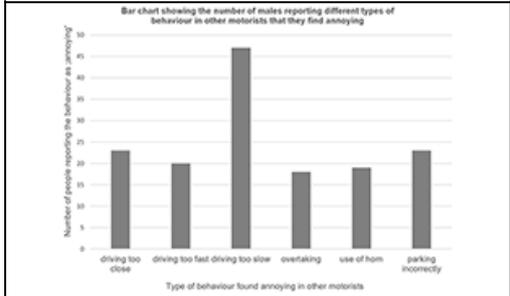
					ratings of 18 and therefore would need to have the same ranking. Many responses also explained that ranks 16, 17, 18 were assigned and the median/mean/middle value found. Some candidates did not gain the second mark due to not explaining how this was calculated (the mean/median of the rank positions).															
			Total	4																
52			<table border="1"> <tr> <td colspan="3">Wearing name badge condition = 14.8 Workings $18+14+10+17+16+18+5+11+20+19 = 148$ ($148/10 = 14.8$) No badge condition = 8.8 $13+2+6+4+8+18+7+15+12+3 = 88$ ($88/10 = 8.8$)</td> </tr> <tr> <td colspan="3">Correct calculation of mean in both conditions with workings for both - 3</td> </tr> <tr> <td>Correct calculation of mean in both conditions with workings for one - 2</td> <td>OR correct calculation of mean in both conditions but no workings - 2</td> <td>OR correct calculation of mean in one condition with workings - 2</td> </tr> <tr> <td>Correct calculation of mean for one condition with no workings - 1</td> <td colspan="2">Correct workings with incorrect mean - 1</td> </tr> <tr> <td colspan="3">The candidate has not provided any creditworthy information - 0</td> </tr> </table>	Wearing name badge condition = 14.8 Workings $18+14+10+17+16+18+5+11+20+19 = 148$ ($148/10 = 14.8$) No badge condition = 8.8 $13+2+6+4+8+18+7+15+12+3 = 88$ ($88/10 = 8.8$)			Correct calculation of mean in both conditions with workings for both - 3			Correct calculation of mean in both conditions with workings for one - 2	OR correct calculation of mean in both conditions but no workings - 2	OR correct calculation of mean in one condition with workings - 2	Correct calculation of mean for one condition with no workings - 1	Correct workings with incorrect mean - 1		The candidate has not provided any creditworthy information - 0			Max 3	<p>Context not required Credit workings for full marks as (both included)- $148/10 = 14.8$ $88/10 = 8.8$ Credit all results for one condition on one line divided by 10 on the next e.g. $18+14+10+17+16+18+5+11+20+19 / 10 = 14.8$ Credit rounding up 15 and 9</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates were correct and most achieved full marks by including the correct workings and correct mean scores for both conditions. Some responses went beyond what was needed and also calculated the mean of both conditions together which was not a requirement of the question. A minority of responses incorrectly worked out the mean of the ranks rather than the ratings of friendliness which did not gain marks. There were a very small number of responses that made an error in their calculation and presented the incorrect mean score for one of the conditions.</p>
Wearing name badge condition = 14.8 Workings $18+14+10+17+16+18+5+11+20+19 = 148$ ($148/10 = 14.8$) No badge condition = 8.8 $13+2+6+4+8+18+7+15+12+3 = 88$ ($88/10 = 8.8$)																				
Correct calculation of mean in both conditions with workings for both - 3																				
Correct calculation of mean in both conditions with workings for one - 2	OR correct calculation of mean in both conditions but no workings - 2	OR correct calculation of mean in one condition with workings - 2																		
Correct calculation of mean for one condition with no workings - 1	Correct workings with incorrect mean - 1																			
The candidate has not provided any creditworthy information - 0																				
			Total	3																
53			B		<p>competence</p> <p><u>Examiner's Comments</u></p> <p>The majority of candidates</p>															

					responded well to this question. Incorrect responses were varied.
			Total	1	
54			A		8:3 <u>Examiner's Comments</u> The majority of candidates responded well to this question. Incorrect responses were varied.
			Total	1	
55			D		percentage <u>Examiner's Comments</u> The majority of candidates responded well to this question. Incorrect responses were varied.
			Total	1	
56			A		Chi-square <u>Examiner's Comments</u> The majority of candidates responded well to this question. Incorrect responses were varied.
			Total	1	
57			D		Wilcoxon <u>Examiner's Comments</u> The majority of candidates responded well to this question. Some candidates chose option B incorrectly.
			Total	1	
58			B		data that is obtained directly from the sample by the researcher(s) <u>Examiner's Comments</u> The majority of candidates responded well to this question. Incorrect responses were varied.
			Total	1	

59			<p>The probability of the null hypothesis being true (or results being due to chance) is less than 5% (so in this study, there is a significant difference in what male and female motorists find annoying in other motorists)</p> <p>Correct answer clearly explained - 2</p> <p>Basic response such as 'There is a significant difference' - 1</p> <p>The candidate has not provided any creditworthy information - 0</p>	Max 2	<p>Context is not required for full marks.</p> <p>For 1 mark accept 'The null hypothesis will be rejected' or 'The Alternate hypothesis will be accepted'.</p> <p><u>Examiner's Comments</u></p> <p>Many candidates didn't understand this question or didn't attempt it and gained zero marks. Some candidates gained 1 mark and recognised that data would be significant or that the null hypothesis would be rejected. The minority of candidates accessed full marks on this question.</p>
			Total	2	
60		D		1	<p><u>Examiner's Comments</u></p> <p>The majority of candidates responded correctly however a minority of students chose option C.</p>
			Total	1	
61		C		1	<p><u>Examiner's Comments</u></p> <p>The majority of candidates responded correctly with the minority of candidates incorrectly choosing option A.</p>
			Total	1	
62		C		1	<p><u>Examiner's Comments</u></p> <p>Some candidates were unable to give a correct response to this question. Probably because more focus was given to identifying ordinal and nominal data.</p>
			Total	1	
63		B		1	<p><u>Examiner's Comments</u></p> <p>The majority of candidates responded correctly.</p>
			Total	1	

64		A	1	<u>Examiner's Comments</u> The majority of candidates responded correctly.								
		Total	1									
65		B	1	<u>Examiner's Comments</u> The majority of candidates responded correctly.								
		Total	1									
66		<p>Weaknesses could include: problems understanding rating scale (if used) by participants; problems responding to forced choice, lack of insight and meaning relating to why participants behaved / responded as they did etc Up to 3 marks for each weakness</p> <table border="1"> <tr> <td colspan="2">Clear outline of weakness in context - 3</td> </tr> <tr> <td>Clear outline of weakness but not in context - 2</td> <td>OR attempted outline of weakness in context - 2</td> </tr> <tr> <td colspan="2">Brief and/or weak attempt to outline weakness (whether in context or not) - 1</td> </tr> <tr> <td colspan="2">The candidate has not provided any creditworthy information - 0</td> </tr> </table>	Clear outline of weakness in context - 3		Clear outline of weakness but not in context - 2	OR attempted outline of weakness in context - 2	Brief and/or weak attempt to outline weakness (whether in context or not) - 1		The candidate has not provided any creditworthy information - 0		Max 3+3	<p>-Context = music, prosocial, love, kindness, helpful etc</p> <p><u>Examiner's Comments</u></p> <p>Most candidates were able to outline one weakness of quantitative data but were unable to outline a second weakness without becoming repetitive in their response.</p> <p>Exemplar 2</p> <p><i>1. One weakness is having quantitative data is that it lacks rich detail about how prosocial lyrics influence how people are. This is as pros. lyrics about to give information on how prosocial lyrics make people feel or how they may feel more helpful about.</i></p> <p><i>2. Another weakness is that results may not be valid as people may feel rejected in the questions and options being provided and so may not give accurate, valid or reliable results on if prosocial lyrics make them feel more helpful.</i></p> <p>This exemplar demonstrates how candidates can address the question without repetition of their first weakness. The responses identify separate weaknesses, explain why they are a weakness and are in context of the stem given.</p>
Clear outline of weakness in context - 3												
Clear outline of weakness but not in context - 2	OR attempted outline of weakness in context - 2											
Brief and/or weak attempt to outline weakness (whether in context or not) - 1												
The candidate has not provided any creditworthy information - 0												
		Total	6									
67		<table border="1"> <tr> <td>Closed question. Reason = the data is just frequency counts of the number of responses in each category of</td> </tr> </table>	Closed question. Reason = the data is just frequency counts of the number of responses in each category of	Max 2	<p>Context is not required for full marks</p> <p><u>Examiner's Comments</u></p> <p>Most candidates got this answer</p>							
Closed question. Reason = the data is just frequency counts of the number of responses in each category of												

		<p>things that motorists may find annoying. The data is participants responses to fixed choice response options with no opportunity to respond in any different way.</p> <p>Correct question type identified and clearly justified - 2</p> <table border="1"> <tr> <td>Correct question type identified - 1</td> <td>OR correct question type identified and unclear attempt to justify why - 1</td> </tr> </table> <p>The candidate has not provided any creditworthy information - 0</p>	Correct question type identified - 1	OR correct question type identified and unclear attempt to justify why - 1		<p>correct and gained 2 marks. For those who only gained 1 mark, they were unable to provide an appropriate reason for it being a closed question.</p>
Correct question type identified - 1	OR correct question type identified and unclear attempt to justify why - 1					
		Total	2			
68		<p>For example: Male and female motorists find different things annoying (males 'people driving too slow' was the most frequently reported category, and for females 'people driving too close'), which suggests the way men and women approach driving and their experiences of it are very different; Many more females (51) than males (20) reported being annoyed by other motorists driving too fast, which suggests females are more safety conscience than males and think more about the consequences of driving too fast; Etc Up to 2 marks for each conclusion ...</p> <p>Clear outline of conclusion in context - 2</p> <table border="1"> <tr> <td>Clear outline of conclusion but not in context - 1</td> <td>OR attempted outline of conclusion in context - 1</td> </tr> </table> <p>The candidate has not provided any creditworthy information - 0</p>	Clear outline of conclusion but not in context - 1	OR attempted outline of conclusion in context - 1	<p>Max 2+2</p>	<p>-Context = driving, car(s), anger, frustration, annoying etc</p> <p>-Must be conclusions (interpretation of data), and not simply stating findings (if so, cap at 1 mark max, whether one or two findings stated)</p> <p><u>Examiner's Comments</u></p> <p>This question gained a variety of responses with few candidates accessing the full range of marks. Most candidates were restricted in their marks due to only identifying findings rather than a conclusion.</p> <p> Assessment for learning</p> <p>Candidates would benefit from being given opportunities to examine raw data tables and practice extrapolating conclusions that could be drawn from them by looking at the nature of the data and any patterns that can be inferred. These skills can also be illustrated through data collected through practical investigations conducted in preparation for this exam.</p> <p>The difference in findings and conclusions from the core studies component could be highlighted to assist in demonstrating to candidates how to differentiate between findings and conclusions.</p>
Clear outline of conclusion but not in context - 1	OR attempted outline of conclusion in context - 1					

		Total	4						
69		<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;">  <p style="text-align: center; font-size: small;">Bar chart showing the number of males reporting different types of behaviour in other motorists that they find annoying.</p> </div> <p>1 mark is awarded for correctly presenting by value each bar representing all 6 categories of motorists' behaviour found annoying for males</p> <p>1 mark is awarded for named categories on the x axis</p> <p>1 mark is awarded for clear labelling of the y axis (must be contextualised to 'annoying behaviour')</p> <p>1 mark is awarded for a clearly contextualised title</p> <table border="1" style="width: 100%; text-align: center; font-size: small;"> <tr><td>All features included - 4</td></tr> <tr><td>3 features included - 3</td></tr> <tr><td>2 features included - 2</td></tr> <tr><td>1 feature included - 1</td></tr> <tr><td>The candidate has not provided any creditworthy information - 0</td></tr> </table>	All features included - 4	3 features included - 3	2 features included - 2	1 feature included - 1	The candidate has not provided any creditworthy information - 0	Max 4	<p>-Context = driving, car(s), anger, frustration, annoying etc Full Contextualisation would include driving/motorists and annoying</p> <p>If presented as a histogram cap at 2m</p> <p>If both males and females are included on the graph cap at 2m</p> <p>If a clearly contextualised title has been included, this could make up for any shortfall of labelling on an x and y axis</p> <p><u>Examiner's Comments</u></p> <p>Many candidates could draw a bar chart but few accessed full marks due to not providing a title, labelled axes or incorrectly presenting data as a histogram rather than a bar chart.</p>
All features included - 4									
3 features included - 3									
2 features included - 2									
1 feature included - 1									
The candidate has not provided any creditworthy information - 0									
		Total	4						
70		<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"> <p style="text-align: center;">14%</p> <p style="text-align: center;">Workings ...</p> <p style="text-align: center;">18 (males) + 24 (females) = 42 in total</p> <p style="text-align: center;">So ... $42/300 \times 100 = 14$</p> </div> <table border="1" style="width: 100%; text-align: center; font-size: small;"> <tr><td>Correct answer with full workings shown - 3</td></tr> <tr><td>Correct answer with some workings shown - 2</td></tr> <tr><td>Correct answer shown with no workings OR some correct workings shown but incorrect final answer - 1</td></tr> <tr><td>The candidate has not provided any creditworthy information - 0</td></tr> </table>	Correct answer with full workings shown - 3	Correct answer with some workings shown - 2	Correct answer shown with no workings OR some correct workings shown but incorrect final answer - 1	The candidate has not provided any creditworthy information - 0	Max 3	<p>18 + 24 is not a necessary working to show for full marks.</p> <p><u>Examiner's Comments</u></p> <p>Most candidates achieved full marks on this question. A small number of candidates lost a mark for not showing their full workings.</p>	
Correct answer with full workings shown - 3									
Correct answer with some workings shown - 2									
Correct answer shown with no workings OR some correct workings shown but incorrect final answer - 1									
The candidate has not provided any creditworthy information - 0									
		Total	3						
71	a	<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"> <p style="text-align: center;">Reasons:</p> <p style="text-align: center;">-test of difference (the study investigated the difference</p> </div>	Max 3	<p>-Context = driving, car(s), anger, annoying, frustration, males and females etc</p> <p><u>Examiner's Comments</u></p>					

			<p>between what males and females found annoying in other motorists' behaviour)</p> <p>-nominal data collected (frequency count of different categories of annoying behaviour in motorists)</p> <p>-independent measures design (males v females)</p> <p>Two or three appropriate reasons provided, at least one in context - 3</p> <p>Two appropriate reasons provided without context OR one appropriate reason in context - 2</p> <p>One appropriate reason provided without context - 1</p> <p>The candidate has not provided any creditworthy information - 0</p>		<p>Most students managed to identify two criteria for use of the Chi square test. Those who gained full marks were able to contextualise at least one of these criteria explicitly to the scenario.</p>
	b		<p>df = 5</p> <p>Workings ...</p> <p>$(R-1) \times (C-1)$</p> <p>$(2-1) \times (6-1) = 5$</p> <p>Correct answer (no need for workings) - 1</p> <p>The candidate has not provided any creditworthy information - 0</p>	Max 1	<p><u>Examiner's Comments</u></p> <p>Many candidates got this question incorrect or didn't write a response.</p>
	c		<p>A clear explanation</p> <p>The 'degrees of freedom' is mapped against levels of significance on a table of critical values to establish the critical value necessary to judge the significance of the study.</p> <p>Basic explanation: e.g. It is used to find the critical value. - 1</p> <p>The candidate has not provided any creditworthy information - 0</p>	Max 2	<p>For two marks candidates must refer to level of significance or probability AND finding the critical value.</p> <p>Reference to number of participants is not creditworthy.</p> <p><u>Examiner's Comments</u></p> <p>Many candidates got this question incorrect or didn't write a response. Some candidates gained 1 mark for reference to finding a critical value. A minority of candidates gained full marks on this question.</p>
			Total	6	
72		C		1	<p><u>Examiner's Comments</u></p> <p>The majority of candidates responded correctly.</p>
			Total	1	
73		C		1	<p><u>Examiner's Comments</u></p> <p>The majority of candidates responded correctly.</p>

			Total	1	
74			B	1	<p><u>Examiner's Comments</u></p> <p>The minority of candidates found it difficult to identify this graph as a histogram.</p> <p> Misconception</p> <p>There was a common misconception here that this is a bar chart rather than a histogram.</p>
			Total	1	
75			<p>Accept -0.639 and -0.64</p> <p>Workings ...</p> $r_s = 1 - \frac{6(\sum d^2)}{n(n^2-1)}$ <p>$6 \times 270.5 = 1623$ $1623 / 990 = 1.64$ $1 - 1.64 = -0.64$</p> <p>3 mark answer: Correct answer with all workings</p> <p>2 mark answer: Correct answer with some workings</p> <p>1 mark answer: Correct calculation with no workings OR Some correct workings, but incorrect answer</p> <p>0 marks: No credit worthy information</p>	3	<p><u>Examiner's Comments</u></p> <p>There was a spread of marks for this question. Candidates generally were able to calculate the value of r_s however were sometimes penalised for not including their workings.</p>
			Total	3	
76			<p>The probability of there being no relationship between self-ratings of extroversion and ratings by a friend is greater than 5%, meaning that this is not a statistically significant result. There is no relationship between self-ratings of extroversion and ratings made by a friend.</p> <p>3 mark answer:</p>	3	<p>Context = personality, extroversion etc</p> <p><u>Examiner's Comments</u></p> <p>There were quite a few non-responders to this question, in addition to many stating that this would mean there was a significant relationship. Some candidates</p>

			<p>Clear explanation in context</p> <p>2 mark answer: Clear explanation but not in context</p> <p>1 mark answer: Brief and/or weak explanation, whether in context or not</p> <p>0 marks: No credit worthy information</p>		<p>stated that there would not be a significant relationship but did not go beyond this. Very few included context and a minority of candidates were able to give a clear explanation.</p>
			Total	3	
77			C	1	<p><u>Examiner's Comments</u></p> <p>Most candidates answered this correctly.</p>
			Total	1	
78			D	1	<p><u>Examiner's Comments</u></p> <p>Many candidates answered this correctly with some giving an inaccurate response.</p>
			Total	1	
79			A	1	<p><u>Examiner's Comments</u></p> <p>Many candidates answered this correctly with some inaccurately answering C as their response.</p>
			Total	1	
80			A	1	<p><u>Examiner's Comments</u></p> <p>Some inaccurate responses were given to this question.</p> <p> Assessment for learning</p> <p>Candidates should undertake calculations of inferential statistical tests to understand the outcomes.</p>
			Total	1	
81			For example: inclusion of distractor questions; ensuring anonymity of respondents; disguising aim of research	6	Context = mental illness, psychiatric, attitudes etc

		<p>etc</p> <p>3 mark answer: Clear outline in context</p> <p>2 mark answer: Clear outline but not in context OR Attempted outline in context</p> <p>1 mark answer: Brief and/or weak attempt to outline (whether in context or not)</p> <p>0 marks: No credit worthy information</p>		<p>3 marks for each way outlined to reduce demand characteristics.</p> <p><u>Examiner's Comments</u></p> <p>There was a mixed response to this question. Some candidates gained 3 marks for one suggestion but struggled to write an equally successful response for another suggestion. Some candidates were penalised due to the lack of detail or context in their response. Some candidates were unable to give a response or gave a response that wasn't relevant to the self-report study such as using independent measures design.</p>
		Total	6	
82		<p>Conclusions could include: there was a variety in how consistent some people rated their own personality compared to the rating from their best friend, suggesting that some people think about their personality in a similar way to how their friend perceives them, whereas other people think very differently. The differences may be due to how well or how long the person has known their best friend. Judging personality is a difficult thing to do, and the way we perceive our own personality may be different because it is based on different things to what anyone else used to judge personality, such as things that are private and only known to the individual etc</p> <p>3 mark answer: Clear detailed response, in context</p> <p>2 mark answer: Clear detailed response, but not in context OR Attempted response, in context</p> <p>1 mark answer: Attempted response, not in context OR simply stating a finding</p>	6	<ul style="list-style-type: none"> Context = personality, extroversion etc Clear (explicit) interpretation of findings (not simply stating a finding) is required for top band <p>Accept any other appropriate conclusions here. 3 marks for each conclusion</p> <p><u>Examiner's Comments</u></p> <p>This question gained a variety of responses with not many candidates accessing the full range of marks. Candidates were required to interpret a table of data and infer a finding. Candidates needed to go beyond merely restating an individual finding which many did not do. Credit was given for a finding but this restricted candidates to low band marks.</p> <p> Assessment for learning</p> <p>Candidates would benefit from</p>

		<p>0 marks: No credit worthy information</p>		<p>being given opportunities to examine raw data tables and practice extrapolating conclusions that could be drawn from them by looking at the nature of the data and any patterns that can be discerned. These skills can also be illustrated through practical work data gathered in preparation for this exam.</p> <p>The difference in findings and conclusions from the core studies component could be highlighted to assist in demonstrating to candidates how to differentiate between findings and conclusions.</p>
		Total	6	
83		<p>Strengths include easy to analyse, record and present findings on ratings of extroversion; easy to plot visually (scatter diagram) etc</p> <p>Weaknesses include lack of insight into why ratings on extroversion mean, or why given; possible differences in use / interpretation of extroversion scale between participants etc</p> <p>3 mark answer: Clear outline of strength/weakness, in context</p> <p>2 mark answer: Attempted outline of strength/weakness, in context OR Clear outline of strength/weakness, not in context</p> <p>1 mark answer: Brief and/or weak outline of strength/weakness, whether in context or not</p> <p>0 marks: No credit worthy information</p>	6	<p>Context - personality, extroversion etc</p> <p>3 marks for strength 3 marks for weakness</p> <p><u>Examiner's Comments</u></p> <p>Most candidates were able to identify a strength and a weakness of quantitative data. For those not achieving full marks, often candidates were penalised for not contextualising their response or for not elaborating on the strength/weakness. For example, candidates need to go beyond 'it lacks detail' for a weakness.</p> <p> Assessment for learning</p> <p>Centres should encourage candidates to make sure they always contextualise their responses when a question includes the phrase "in this study".</p> <p>Candidates could be encouraged to self-reflect on their own work and highlight their use of context. This will assist candidates in identifying where they are losing marks if they</p>

					have not included context in their response.
			Total	6	
84			<p>Reasons:</p> <ul style="list-style-type: none"> test of correlation/relationship (and the study investigated the relationship between self-rating and friends rating of extroversion) ordinal data collected (ratings of extroversion on 0 to 100 scale) <p>3 mark answer: Two appropriate reasons provided, both in context</p> <p>2 mark answer: Two appropriate reasons provided, but only one, or neither in context OR One appropriate reason provided in context</p> <p>1 mark answer: One appropriate reason provided but not in context</p> <p>0 marks: No credit worthy information</p>	3	<p>Context - personality, extroversion etc</p> <p><u>Examiner's Comments</u></p> <p>Candidates answered this question well in terms of knowledge. Very few candidates contextualised their response so were often penalised here.</p>
			Total	3	
85			C	1	<p><u>Examiner's Comments</u></p> <p>This question was answered correctly by most candidates.</p>
			Total	1	
86			D test-retest		<p><u>Examiner's Comments</u></p> <p>The majority of candidates answered correctly. Some candidates chose option A incorrectly.</p>
			Total	1	

87			B nominal		<u>Examiner's Comments</u> This was answered correctly by the majority of candidates. Incorrect choices by candidates were varied.
			Total	1	
88			C 3:2	Enter text here.	<u>Examiner's Comments</u> Most candidates were successful in this question. A small number of candidates chose option D incorrectly.
			Total	1	
89			B 0.04		<u>Examiner's Comments</u> Answered correctly by most candidates. Some candidates chose option A incorrectly.
			Total	1	
90			B Binomial Sign		<u>Examiner's Comments</u> Answered correctly by less than half of all candidates. Many candidates chose option D incorrectly.
			Total	1	
91			C sample size		<u>Examiner's Comments</u> Answered correctly by over half of the candidates. Some candidates chose option A incorrectly.
			Total	1	
92			D positively skewed		<u>Examiner's Comments</u> Answered correctly by less than half of the candidates. Some candidates chose option B or C incorrectly.
			Total	1	
93			Possible examples Disguising aim of study Asking additional, unrelated questions etc. <div style="border: 1px solid black; padding: 2px; width: fit-content;">One way to reduce demand characteristics clearly presented in context</div>	Max 2 2 1	Context = friendliness, hand-shaking etc. Experimenter isn't in room when study is taking place Independent measures (attempt unless explained) Experimenter keeps a neutral

		<table border="1"> <tr> <td>One way to reduce demand characteristics clearly presented but not in context</td> <td>OR attempted way to reduce demand characteristics (whether in context or not)</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	One way to reduce demand characteristics clearly presented but not in context	OR attempted way to reduce demand characteristics (whether in context or not)	0	<p>expression (attempt unless explained) Ensuring anonymity of responses.</p> <p><u>Examiner's Comments</u></p> <p>There were many good responses which used deception or distractor questions to disguise the aim or use independent measures design. Better responses used context examples as part of their explanation. Lower scoring responses gave a basic outline of how to reduce demand characteristics such as disguising the aim without outlining either how this could be done (e.g. being told a different aim) or why this would lead to a reduction in demand characteristics. A significant number of responses were confused between demand characteristics and social desirability or suggested deceiving the participants in the debrief as the candidate was clearly confused about when a debrief takes place.</p>						
One way to reduce demand characteristics clearly presented but not in context	OR attempted way to reduce demand characteristics (whether in context or not)											
		Total	2									
94		<p>Conclusions could include: males gesture more, perhaps because they are less self-conscious than females, or that they are more visual in how they think and this is reflected in conversations they have whilst on the phone with other people making them more animated etc. Accept any other appropriate conclusions here.</p> <table border="1"> <tr> <td colspan="2">3 marks for each conclusion</td> </tr> <tr> <td colspan="2">Clear, detailed conclusion in context</td> </tr> <tr> <td>Clear, detailed conclusion but not in context</td> <td>OR attempt in context</td> </tr> <tr> <td>Brief and/or weak attempt (whether in context or not)</td> <td>OR simply stating a finding</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	3 marks for each conclusion		Clear, detailed conclusion in context		Clear, detailed conclusion but not in context	OR attempt in context	Brief and/or weak attempt (whether in context or not)	OR simply stating a finding	<p>3</p> <p>3</p> <p>2</p> <p>1</p> <p>0</p>	<p>Context = mobile, phone, gestures etc.</p> <p>Must be an interpretation of findings (not simply presentation of finding).</p> <p>1 mark for finding, 2 marks for explanation. Example of a finding - 'Males gesture more than females' (1)</p> <p>Zero marks if just data is given.</p> <p><u>Examiner's Comments</u></p> <p>Most responses were able to gain 1 mark for this question by concluding that men use more gestures than women while making or receiving a mobile phone call. However, many candidates did not understand the term 'conclusion' and often repeated the data from</p>
3 marks for each conclusion												
Clear, detailed conclusion in context												
Clear, detailed conclusion but not in context	OR attempt in context											
Brief and/or weak attempt (whether in context or not)	OR simply stating a finding											

					the table. Better responses gave an explanation as to why men might use more gestures such as men being more passionate or enthusiastic; females being more verbally articulate so did not need to gesture to emphasise their point.								
			Total	3									
95			<p>A pie chart showing the proportion of males compared to females who made gestures whilst on their mobile phone</p> <table border="1"> <tr> <td>Each feature can be awarded one mark as follows ...</td> </tr> <tr> <td>1 mark is awarded for correctly calculating what % of the circle should represent the males and females who made gestures.</td> </tr> <tr> <td>1 mark is awarded for drawing the sectors in correct proportion to data</td> </tr> <tr> <td>1 mark is awarded for clear labelling of each sector of the pie chart</td> </tr> <tr> <td>1 mark is awarded for a clear and appropriate title</td> </tr> <tr> <td>4 features included</td> </tr> <tr> <td>3 features included</td> </tr> <tr> <td>2 features included</td> </tr> <tr> <td>1 feature included</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	Each feature can be awarded one mark as follows ...	1 mark is awarded for correctly calculating what % of the circle should represent the males and females who made gestures.	1 mark is awarded for drawing the sectors in correct proportion to data	1 mark is awarded for clear labelling of each sector of the pie chart	1 mark is awarded for a clear and appropriate title	4 features included	3 features included	2 features included	1 feature included	<p>Max 4</p> <p>1 mark for % 1 mark for proportions 1 mark – male/female labelled (or key) 1 mark title</p> <p>Males $11/16 \times 100 = 68.75\%$ accept 69 Females $5 / 16 \times 100 = 31.25\%$ accept 31</p> <p>Title should refer to males and females and gestures. If no reference is made to percentage/proportion in title there must be % within the pie chart to be creditworthy. Reference to number of males/females is not creditworthy.</p> <p>If all data included – can be awarded full marks (see pie chart below) the candidate may alter the percentage slightly for one of the categories to show the pie chart adds up to 100% (43% below illustrates this) which is still creditworthy. Credit the whole numbers (15%, 42/43%, 9% and 33%) regardless of what is written after the decimal point.</p> <p>4 3 Yes Female – 15% (15 2) 2 No Female – 42% (42.4)/43% 1 No Male – 9% (9.1) Yes Male – 33% (33.3) 0 OR drawing 2 pie charts – one for males (21.4%, 78.6% or 21%, 79%) and one for females (26.3%, 73.7%, 26%, 74%)also can be awarded 4 marks</p> <p>Drawing a pie chart where the total for males (79%) and total for</p>
Each feature can be awarded one mark as follows ...													
1 mark is awarded for correctly calculating what % of the circle should represent the males and females who made gestures.													
1 mark is awarded for drawing the sectors in correct proportion to data													
1 mark is awarded for clear labelling of each sector of the pie chart													
1 mark is awarded for a clear and appropriate title													
4 features included													
3 features included													
2 features included													
1 feature included													

				<p>females (26%) is not creditworthy.</p> <p>Be generous with drawing of the circle and indication of the proportions e.g., for the 31% need to be over a quarter of the circle.</p> <p><u>Examiner's Comments</u></p> <p>Many candidates achieved 3 or 4 marks for this question by working out how many males and females gestured overall and then the percentage of males and females within this group. There were many responses that gave a clear title and labelled all sections of their pie chart as well as putting in the percentages for each section. Some responses provided a key rather than a label within the pie chart which was creditworthy. Less successful responses often had no title or an inappropriate title such as 'number of males and females..' where the pie chart showed the percentage. Those responses that provided a pie chart for all the data from the table were also able to achieve full credit for this question.</p>										
		Total	4											
96		<p>A pie chart of percentages of males and females who made gestures or not whilst on their mobile phone</p> <table border="1"> <caption>Data from Pie Chart</caption> <thead> <tr> <th>Category</th> <th>Percentage</th> </tr> </thead> <tbody> <tr> <td>Females NO</td> <td>42.42%</td> </tr> <tr> <td>Males YES</td> <td>31.33%</td> </tr> <tr> <td>Females YES</td> <td>15.15%</td> </tr> <tr> <td>Males NO</td> <td>9.09%</td> </tr> </tbody> </table> <p>Any one reason in context from: nominal/categorical data obtained; looking for a difference; independent/unrelated groups One appropriate reason in appropriate context</p> <p>One appropriate reason but not in context</p> <p>The candidate has not provided any creditworthy information</p>	Category	Percentage	Females NO	42.42%	Males YES	31.33%	Females YES	15.15%	Males NO	9.09%	<p>Max 2</p> <p>2</p> <p>1</p> <p>0</p>	<p>Nominal data/categories – whether the participant makes a gesture or not OR yes/no</p> <p>Independent measures – male/female</p> <p>Looking for a difference – between males and females OR difference between the number of gestures made between males and females.</p> <p>Reason and context must be linked.</p> <p><u>Examiner's Comments</u></p> <p>Many candidates achieved full marks for this question by either identifying that the data was nominal with the example of the categories from the study (use of</p>
Category	Percentage													
Females NO	42.42%													
Males YES	31.33%													
Females YES	15.15%													
Males NO	9.09%													

					gestures or not) or identified that the study had an independent measures design with the two conditions of male and female. A small number of responses identified that the study was investigating a difference but often this lacked context. Less successful responses often gave two reasons such as nominal and independent measures without contextualising either of these.						
			Total	2							
97	a		<p>9.79</p> <p>Row total (17) × column total (19) / overall total (33) = 9.78787878</p> <table border="1"> <tr> <td colspan="2">9.79 with workings shown</td> </tr> <tr> <td>9.78787878 with workings (not to 2 dps)</td> <td>OR 9.79 without workings</td> </tr> <tr> <td colspan="2">9.78787878 without workings (not to 2 dps) OR correct workings only presented</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	9.79 with workings shown		9.78787878 with workings (not to 2 dps)	OR 9.79 without workings	9.78787878 without workings (not to 2 dps) OR correct workings only presented		<p>Max 3</p> <p>3</p> <p>2</p> <p>1</p> <p>0</p>	<p>$\frac{17 \times 19}{33} = 9.79$ Award 3 marks</p> <p>14-4.21 = 9.79 Award 3 marks</p> <p>33-23.21 = 9.79 Award 3 marks</p> <p>Many candidates are putting workings out for 30(b) in the answer to this question. If possible, ignore these workings out for this question.</p> <p>9.79 in correct column in table (regardless of what else is written) = 1</p> <p><u>Examiner's Comments</u></p> <p>Candidates were well prepared for this question. A common good response was achieved via $17 \times 19 / 33$ method. Common error was not rounding up to 2 decimal places or just knocking off the remaining numbers and using 9.78 rather than rounding up to 9.79.</p>
9.79 with workings shown											
9.78787878 with workings (not to 2 dps)	OR 9.79 without workings										
9.78787878 without workings (not to 2 dps) OR correct workings only presented											
	b		<p>$\chi^2 = 8.8$</p> <p>Workings ... Cell (d) $O - E = 14 - 9.79 = 4.21$ $(O - E)^2 = 4.21 \times 4.21 = 17.72$ $(O - E)^2 / E = 17.72 / 9.79 = 1.81$</p> <p>So, the sum of $(O - E)^2 / E = 2.61 + 1.92 + 2.46 + 1.81 = 8.8$</p> <table border="1"> <tr> <td>Correct calculation of overall chi square value with workings shown</td> </tr> </table>	Correct calculation of overall chi square value with workings shown	<p>Max 3</p> <p>3</p>	<p>Can credit cell calculations in the table</p> <p>$17.72 / 9.79 = 1.81$ (1) – on its own</p> <p>$2.61 + 1.92 + 2.46 + 1.81 = 8.8$ award 3 marks</p> <p>If incorrect frequency worked out in 30(a) will mean that the calculation will be incorrect = 0</p> <p><u>Examiner's Comments</u></p>					
Correct calculation of overall chi square value with workings shown											

		<table border="1"> <tr> <td colspan="3">Correct calculation with some workings shown</td> </tr> <tr> <td>Correct calculation.</td> <td>All workings shown correctly but no calculation given (or incorrect calculation)</td> <td>Cell D calculated correctly.</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	Correct calculation with some workings shown			Correct calculation.	All workings shown correctly but no calculation given (or incorrect calculation)	Cell D calculated correctly.	<p>2</p> <p>1</p> <p>0</p>	<p>Candidates were well prepared for this question. The most common good response: $2.61 + 1.92 + 2.46 + 1.81 = 8.8$. A common error was to work out 1.81 to 1.82 instead.</p>
Correct calculation with some workings shown										
Correct calculation.	All workings shown correctly but no calculation given (or incorrect calculation)	Cell D calculated correctly.								
	c	<p>df = 1 Workings ... $(2-1) \times (2-1) = 1$</p> <table border="1"> <tr> <td colspan="2">Correct answer with workings</td> </tr> <tr> <td>Correct answer but not workings (or workings incomplete or unclear)</td> <td>Workings but no answer</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	Correct answer with workings		Correct answer but not workings (or workings incomplete or unclear)	Workings but no answer	<p>Max 2</p> <p>2</p> <p>1</p> <p>0</p>	<p>No marks for just formula. $(R-1) \times (C-1)$</p> <p><u>Examiner's Comments</u></p> <p>Well answered, most candidates understood how to calculate this and got the correct answer.</p> <p>Some counted the headings of the rows and columns in the table OR counted the rows and columns in the table used to work out cell d, making the response incorrect.</p>		
Correct answer with workings										
Correct answer but not workings (or workings incomplete or unclear)	Workings but no answer									
	d	<p>6.635 Correct answer provided The candidate has not provided any creditworthy information</p>	<p>Max 1</p> <p>1</p> <p>0</p>	<p>Award 1 mark if candidate has circled (or otherwise correctly indicated) correct answer in table.</p> <p><u>Examiner's Comments</u></p> <p>Candidates achieved well on this question with many identifying the correct critical value of 6.635. Small number of candidates had the common error of using the 0.05 critical value so clearly misinterpreting its meaning. In addition, some chose the incorrect critical value due to having calculated the incorrect degree of freedom in Question 30 (d).</p>						
	e	<p>Any three from the following statements.</p> <ul style="list-style-type: none"> - There is a statistically significant difference between the number of males and females who made gestures whilst on the phone - The probability of there being no difference between the number of 	<p>Max 3</p>	<p>Context = mobile, phone, gestures etc</p> <p><u>Examiner's Comments</u></p> <p>There were many good responses to this question where candidates correctly interpreted that the</p>						

		<p>males and females who made gestures whilst on the phone is less than 1%</p> <ul style="list-style-type: none"> - Males are more likely to make gestures whilst on the phone compared to females - Because calculated is higher than the critical (a significant difference has been found). - Null hypothesis can be rejected/alternative hypothesis can be accepted. - any probability level other than 0.001 (e.g., $p \leq 0.01$) <p>3 correct statement – with at least 1 in context 2 correct statements 1 correct statements The candidate has not provided any creditworthy information</p>	<p>3 2 1 0</p>	<p>calculated value was higher than the critical value and so the result of the test was significant. Better responses contextualised this in light of the study's overall conclusion. Many then correctly identified that the null hypothesis was rejected. Less successful responses often lacked context, accepting the null hypothesis, and not knowing how to refer to the meaning of the significance level accurately i.e. probability of confidence in the IV impacting the DV/ probability results were a fluke/chance.</p>						
		Total	12							
98	a	<p>Answers could include: easier data analysis and comparison across groups (males and females); easier to record data; easier to interpret, no researcher bias to interpret, etc.</p> <table border="1" style="width: 100%;"> <tr> <td colspan="2">Clear outline of strength in context</td> </tr> <tr> <td>Clear outline of strength but not in context</td> <td>Attempted outline of strength in context</td> </tr> <tr> <td colspan="2">Brief and/or weak attempt at an outline of strength (whether in context or not)</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	Clear outline of strength in context		Clear outline of strength but not in context	Attempted outline of strength in context	Brief and/or weak attempt at an outline of strength (whether in context or not)		<p>Max 3</p> <p>3 2 1 0</p>	<p>Context = mobile, phone, gestures, male/female etc.</p> <p>No credit for reference to being able to calculate the mean as this is not relevant for this data.</p> <p><u>Examiner's Comments</u></p> <p>There were many accurate strengths identified by candidates in this question such as easy to analyse or to compare. Some responses were clearly outlined as to why this was a strength in the context of the study. Lower achieving responses were where candidates did not fully elaborate on a point and just mentioned that this made the study more reliable or more objective, for example, without an explanation as to why. In addition, some candidates did not recognise that the data to which they were referring was nominal and mentioned factors such as calculating mean.</p>
Clear outline of strength in context										
Clear outline of strength but not in context	Attempted outline of strength in context									
Brief and/or weak attempt at an outline of strength (whether in context or not)										
	b	<p>Answers could include: doesn't inform us about reasons why there is a difference in amount of gesturing made whilst on phone between males and</p>	<p>Max 3</p>	<p>Context = mobile, phone, gestures, male/female etc.</p> <p><u>Examiner's Comments</u></p>						

		<p>females, mean cannot be calculated, etc.</p> <table border="1"> <tr> <td colspan="2">Clear outline of weakness in context</td> </tr> <tr> <td>Clear outline of weakness but not in context</td> <td>Attempted outline of weakness in context</td> </tr> <tr> <td colspan="2">Brief and/or weak attempt at an outline of weakness (whether in context or not)</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	Clear outline of weakness in context		Clear outline of weakness but not in context	Attempted outline of weakness in context	Brief and/or weak attempt at an outline of weakness (whether in context or not)		<p>3</p> <p>2</p> <p>1</p> <p>0</p>	<p>There were many good responses to this question which often referred to lack of insight or detail. The best responses included some relevant detail such as quantitative data being limited because we are unaware of why gestures were being made such as how annoyed the participant on the phone was or the content of their call. Responses that did not score as well did not outline why a lack of detail is a problem or did not include context.</p>		
Clear outline of weakness in context												
Clear outline of weakness but not in context	Attempted outline of weakness in context											
Brief and/or weak attempt at an outline of weakness (whether in context or not)												
		Total	6									
99		<p>Evaluation points could include: higher validity as participants unlikely to know they are being observed; higher ecological validity as in natural setting of high street; low validity due to possibly missing/misinterpreting some use of gesturing if high street busy, lower validity due to more extraneous variables in the natural environment, etc.</p> <p>Accept any other appropriate conclusions here.</p> <table border="1"> <tr> <td colspan="2">3 marks for each evaluation point</td> </tr> <tr> <td colspan="2">Clear, detailed evaluation in context</td> </tr> <tr> <td>Clear, detailed evaluation but not in context</td> <td>OR attempted evaluation in context</td> </tr> <tr> <td colspan="2">Brief and/or weak attempt at evaluation (whether in context or not)</td> </tr> </table> <p>The candidate has not provided any creditworthy information</p>	3 marks for each evaluation point		Clear, detailed evaluation in context		Clear, detailed evaluation but not in context	OR attempted evaluation in context	Brief and/or weak attempt at evaluation (whether in context or not)		<p>6</p> <p>3</p> <p>2</p> <p>1</p> <p>0</p>	<p>Context = mobile, phone, gestures, male/female, high street, bench etc.</p> <p>Full credit can be given for two positive effects on validity, two negative effects or one of each.</p> <p>Lacks external validity - from a busy high street to different locations is creditworthy</p> <p>Do not credit population validity.</p> <p>Do not credit references to reliability.</p> <p><u>Examiner's Comments</u></p> <p>The marks achieved by candidates for this question covered the full range of the mark scheme. Common points included ecological validity, lack of demand characteristics which improved validity and lack of control of extraneous variables which reduced validity. Examples of good responses fully explained how the natural environment enabled the observers to record true to life gesture behaviour or how behaviour was unaffected by the demand characteristics that occur within the laboratory. Candidates provided good examples of possible extraneous variables that occur within a busy street that could affect the gestures that a person</p>
3 marks for each evaluation point												
Clear, detailed evaluation in context												
Clear, detailed evaluation but not in context	OR attempted evaluation in context											
Brief and/or weak attempt at evaluation (whether in context or not)												

					<p>makes while on the phone (e.g bumping into people, little space, used phone call content well). Less successful responses often identified the issue but did not explain how validity would actually be affected, lacked detail or did not include context; all of which restricted the marks available to them. A significant number of responses identified population validity as an issue in this study which is not specific to naturalistic observations and therefore was not creditworthy.</p> <p>Exemplar 3</p> <p>1. As it is a nonnaturalistic observation it takes place in the participants' natural environment so there is low control over extraneous variables (such as who the person is talking on the phone to) which could have an effect on the dependent variable. Therefore the internal validity will be lower as you can't be certain that the independent variable is truly affecting the dependent variable.</p> <p>2. As it is a naturalistic observation and someone's phone use in their environment is not manipulated and occurs in a place where the behaviour (talking on the phone) is normal the ecological validity will be higher therefore the data findings can be generalised to other places and environments where someone might be on the phone.</p> <p>Exemplar 3 is a full mark response where the two points raised are clear, detailed and in context.</p>																				
			Total	6																					
100			A Mann-Whitney U		<p>Examiner's Comments</p> <p>Most candidates responded correctly to this question. Some candidates chose option C incorrectly.</p>																				
			Total	1																					
101			B 6		<p>Examiner's Comments</p> <p>Most candidates answered this question correctly. Incorrect choices by candidates were varied.</p>																				
			Total	1																					
102	a		<p>Answer T = 5 Workings</p> <table border="1"> <thead> <tr> <th colspan="6">Number of times snacks taken from fridge during one week</th> </tr> <tr> <th>p</th> <th>(a) Without eyes on</th> <th>(b) With eyes on</th> <th>Diff (a - b)</th> <th>Ranks of diffs</th> <th>Sum of ranks of neg' diffs</th> <th>Sum of ranks of pos' diffs</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>	Number of times snacks taken from fridge during one week						p	(a) Without eyes on	(b) With eyes on	Diff (a - b)	Ranks of diffs	Sum of ranks of neg' diffs	Sum of ranks of pos' diffs								5	Allow Diff to be b-a
Number of times snacks taken from fridge during one week																									
p	(a) Without eyes on	(b) With eyes on	Diff (a - b)	Ranks of diffs	Sum of ranks of neg' diffs	Sum of ranks of pos' diffs																			

	fridge door	fridge door				
a	14	10	4	7		7
b	21	18	3	5		5
c	32	24	8	10		10
d	28	24	4	7		7
e	20	16	4	7		7
f	8	10	-2	3.5	3.5	3.5
g	26	24	2	3.5		3.5
h	12	12	0	-		
i	22	23	-1	1.5	1.5	1.5
j	18	12	6	9		9
k	20	19	1	1.5		1.5
l	21	21	0	-		
					$\Sigma r =$ 5	$\Sigma r =$ 50

1 mark for finding difference a-b
1 mark for ranking differences
1 mark for sum of negative ranks
1 mark for sum of positive ranks
1 mark for correct identification of T

b	<p>T = 5 , n = 10, p<0.05</p> <p>OR In words ... e.g. the null hypothesis (i.e. that there is no difference in amount of food consumed with eyes on fridge compared to without) is rejected as the calculated value is less than the critical value. There is a significant finding / the hypothesis is accepted (eyes on fridge <i>does</i> make a difference to amount of food consumed).</p> <p>2 mark answer: Correctly written significance statement (calculated value, sample size and probability level) OR written in words rather than a formal statement</p> <p>1 mark answer: just stating p<0.05</p> <p>OR weak and / or brief written response</p> <p>0 mark answer: No creditworthy response</p>	Max 2	<p>Context = food, snacks, eat(ing), fridge etc OR the correct numerical data (e.g. T=5).</p> <p>1 mark for correctly saying null rejected, hypothesis accepted.</p>
---	--	-------	--

			Total	7	
103	a		<p>Table critical value = 8</p> <p>2 mark answer: Critical value correctly stated</p> <p>1 mark answer: Correct sample size used (10) but wrong critical value quoted</p> <p>0 mark answer: No creditworthy response</p>	2	
	b		<p>Answers could include:</p> <ul style="list-style-type: none"> • Population / scores normally distributed (68% of our data should be less than ± 1 standard deviation around the mean) • At least interval / ratio (continuous) data • No extreme scores <p>1 mark per criteria: Appropriate criteria identified</p> <p>0 marks: No creditworthy response</p>	2	
			Total	4	
104	a		<p>Answer could include: ability to analyse data about snacking more easily; can compare across conditions and individual participants easily etc</p> <p>3 mark answer: Clear outline of advantage in context</p> <p>2 mark answer: Clear outline of advantage, but not in context OR Attempted outline of advantage in context</p> <p>1 mark answer: Brief and / or weak to outline advantage (whether in context or not)</p> <p>0 mark answer: No credit worthy response</p>	3	Context = food, snacks, eat(ing), fridge etc

	b	<p>Answer could include: no reasons for why snacks consumed obtained</p> <p>3 mark answer: Clear outline of disadvantage in context</p> <p>2 mark answer: Clear outline of disadvantage, but not in context OR Attempted outline of disadvantage in context</p> <p>1 mark answer: Brief and / or weak to outline disadvantage (whether in context or not)</p> <p>0 mark answer: No credit worthy response</p>	3	Context = food, snacks, eat(ing), fridge etc
		Total	6	
105	a	<p>Total 14,21,32,28,20,8,26,12,22,18,20,21 =242 /12 Mean = 20.16666 So to 3 SFs = 20.2</p> <p>1 mark for correct working 1 mark for correct mean 1 mark for correctly stating to 3 significant figures</p> <p>0 marks: No creditworthy response</p>	3	
	b	<p>Order: 10, 10, 12, 12, 16, <u>18, 19</u>, 21, 23, 24, 24, 24 Median = 18.5 Range = 24-10 = 14 OR 24-10 = 14+1 = 15</p> <p>1 mark for ordering data correctly 1 mark for identifying median correctly 1 mark for calculating range correctly</p> <p>0 marks: No creditworthy response</p>	3	
		Total	6	
106		Answer = 25% less food consumed when eyes on fridge compared to when no eyes on fridge	3	

		<p>Workings 8 items less when eyes on fridge (32-24) 8 as a percentage of 32 = $8/32 \times 100 = 25\%$</p> <p>3 mark answer: Percentage correctly stated with full workings shown</p> <p>2 mark answer: Percentage correctly stated with some workings shown OR Percentage only shown (no workings)</p> <p>1 mark answer: Partial workings shown e.g. 8/32 or 32-24</p> <p>0 mark answer: No credit worthy response</p>		
		Total	3	
107	a	<p>There was more variation in the consumption of snacks when eyes were not on the fridge compared to when they were. This informs us that having eyes on the fridge has a fairly consistent effect on the amount of food consumed across the different participants. Without eyes on the fridge there is more variation, which informs us that some people may be better than others at restricting taking food from the fridge.</p> <p>3 mark answer: Clear explanation in context</p> <p>2 mark answer: Clear explanation, but not in context OR Attempted explanation in context</p> <p>1 mark answer: Brief and / or weak explanation (whether in context or not)</p> <p>0 mark answer: No credit worthy response</p>	3	Context = food, snacks, eat(ing), fridge etc
	b	<p>The standard deviation is the square root of the variance. The advantage of this is that the answer is a figure that is much more typical of the actual</p>	3	Context = food, snacks, eat(ing), fridge etc Responses that refer to the

		<p>difference (actual data collected) in the amount of snacking taking place with or without eyes on the fridge. The variance is a much bigger (untypical) figure as it is the sum of the differences in each individual's snacking compared to the mean which is then squared (squaring makes the answer a much bigger figure).</p> <p>3 mark answer: Clear outline in context</p> <p>2 mark answer: Clear outline, but not in context OR Attempted outline in context</p> <p>1 mark answer: Brief and / or weak outline (whether in context or not)</p> <p>0 mark answer: No credit worthy response</p>		standard deviation as more valid / accurate than the variance are not creditworthy.
		Total	6	
108		B	1	
		Total	1	
109		C	1	
		Total	1	
110		A	1	
		Total	1	
111		B	1	
		Total	1	
112		A	1	
		Total	1	
113		B	1	
		Total	1	
114		D	1	
		Total	1	
115	a	A	1	
	b	B	1	

		Total	2	
116		C	1	
		Total	1	
117		<p>Outline one strength of the type of data collected in this study.</p> <p>Possible strengths:</p> <ul style="list-style-type: none"> • Easy to make comparisons / identify patterns • Objective / less open to interpretation • Easy to replicate <p><u>Example of 1 mark answer</u> Easy to compare data from each condition.</p> <p><u>Example of 2 mark answer</u> It was easy to make comparisons (1) to see if one group's hippocampi were structured differently from the other's (1).</p>	2	<p>2 marks for a response which identifies a relevant strength which is then developed in the context of the study.</p> <p>1 mark for identifying a relevant strength either explicitly, or implicitly through reference to the study.</p> <p>0 marks – no creditworthy response.</p>
		Total	2	
118		<p>Briefly discuss the extent to which Kohlberg's (1968) study into moral development may be considered ethnocentric.</p> <p>Possible content:</p> <ul style="list-style-type: none"> • More focus on American – through sampling and longitudinal research. • However, five other cultures studied to support universality of theory. • Still Western bias in overall sample. • Stories may be more relevant to some cultures than others. • Story was changed for Taiwanese participants showing some awareness of cultural differences. • Implication that some cultures develop slower than others, implying a deficiency in cultural experience or possibly intellect. 	4	<p>3-4 marks for a clear, detailed and well informed discussion of the extent of ethnocentrism in the context of the study.</p> <p>1-2 marks for a brief or vague discussion of the extent of ethnocentrism in the study. There may be some muddling of ideas or applications as part of the discussion.</p> <p>0 marks – no creditworthy response.</p> <p>NB The question requires a <i>brief</i> discussion and marks should be awarded according to this.</p> <p>A full mark response does not necessarily have to consider both sides of the argument – a good answer could argue that the study is or is not ethnocentric alone.</p> <p>The discussion can focus on one</p>

					aspect of the study (e.g. sample) and still earn full marks.
			Total	4	
119	i	<p>Explain one way in which the procedure of Grant <i>et al</i>'s (1998) study into memory increased the reliability of the research.</p> <p>Possible answers:</p> <ul style="list-style-type: none"> • same background noise within condition • headphones worn for both conditions • all participants read the same article • tests for recall were the same • order of testing was the same • break between study and test phase was about 2 minutes each time <p><u>Example of 1 mark answer</u> The break between study and testing was about the same time each trial.</p> <p><u>Example of 2 mark answer</u> All participants read with headphones on (1) to ensure consistency in experience (1).</p> <p><u>Examples of 3 mark answer</u> Standardisation was used (1) where experimenters used the same article on psychoimmunology (1) to ensure that what participants had to learn was not an extraneous variable affecting recall (1).</p> <p>The experimenters used an article on psychoimmunology (1) which was published and accessible to anyone (1) so that the study could be easily replicated to establish reliability (1).</p>	3	<p>3 marks for a clear response which outlines what is meant by reliability (e.g. consistency or replicability), identifies how this is achieved in study, and then applies specifically to this study through example.</p> <p>2 marks for a vague response with all three of the above features or for a clear response with two of the features.</p> <p>1 mark for identifying a relevant way the procedure addressed the issue of reliability.</p>	
	ii	<p>Outline one conclusion that was drawn from this study.</p> <p>Possible answers:</p> <ul style="list-style-type: none"> • There are context-dependency effects for newly learned meaningful material regardless 	2	<p>2 marks for a clearly identified and relevant conclusion arising from the study.</p> <p>1 mark for a vague, brief or muddled conclusion arising from the study.</p>	

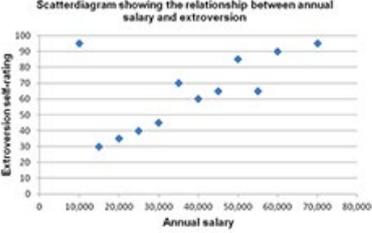
			<p>of whether a short-answer test or a multiple-choice test is used to assess learning.</p> <ul style="list-style-type: none"> Studying and testing in the same environment leads to enhanced performance in recall through access to relevant cues. Students are likely to perform better in exams if they study for them with a minimum of background noise because, although there was no overall effect of noise on performance, the fact that there was evidence for context-dependency suggests they are better off studying without background noise as it will not be present during actual testing. 		0 marks – no creditworthy response.
			Total	5	
120	i	<p>Explain one weakness of the type of data collected in Loftus & Palmer's (1974) study into eye witness testimony.</p> <p>Possible answer: Because the data was quantitative (1) there is a risk that it suggests similarity in response when there may not be (1) so for example a mean estimate of speed does not tell us whether certain individuals' estimates were similar across conditions (1). NB Other appropriate responses should be credited.</p>	3	<p>3 marks for a clear answer which;</p> <ul style="list-style-type: none"> identifies the data is quantitative or similar identifies a weakness of quantitative data identifies this weakness in the context of the study <p>2 marks for an answer which addresses at least two of the above points.</p> <p>1 mark for a partial or vague answer which addresses at least one of the above points.</p> <p>0 marks – no creditworthy response.</p> <p>No credit for use of qualitative data.</p>	
	ii	<p>Outline one conclusion that was drawn from this study.</p> <p>Possible answers:</p> <ul style="list-style-type: none"> The verb used in a question influences a participant's response i.e. the way a question is phrased influences the answer given. 	2	<p>2 marks for a clearly identified and relevant conclusion arising from the study.</p> <p>1 mark for a vague, brief or muddled conclusion arising from the study.</p> <p>0 marks – no creditworthy response.</p>	

		<ul style="list-style-type: none"> • People are not very good at judging vehicular speed regardless of the situation they find themselves in. • Misleading post event information can have a distorting effect on an individual's memory. <p>Information gleaned during the perception of the original event and the post-event information integrate over time to affect memory.</p> <p><u>Example of a 1 mark answer:</u> 'leading questions affected memory', 'eye witness testimonies are unreliable'.</p>		No credit for just findings.
		Total	5	
121	a	$r_s = 1 - \frac{6(\sum d^2)}{n(n^2-1)}$ <p>$R_s = +0.517$</p> <p>1 mark for calculation of differences between ranks in each condition 1 mark for calculation of sum of differences squared ($\sum d^2$) = 138 1 mark for correct application / use of formula provided/substituting at least one number correctly 1 mark for all workings at each stage shown 1 mark for correct final overall answer ($r_s = +0.517$)</p>	5	Won't necessarily show all these steps
	b	<p>Table critical value = 0.587</p> <p>2 mark answer: Critical value correctly stated</p> <p>1 mark answer: Correct sample size (12) used or 0.05 column, but wrong critical value quoted</p> <p>0 marks: No credit worthy information</p>	2	
	c	<p>As the R_s value (+0.517) is smaller than the critical value (0.587) with 12 participants and $p < 0.05$, the null hypothesis is supported that there is no relationship between extroversion self-rating and salary earned).</p>	2	<p>Allow this written in numerical form</p> <p>As $0.517 < 0.587$ for $n=12$, $p > 0.05$</p>

		<p>2 mark answer: Correctly written significance statement with reference to calculated value, critical value, sample size and probability level</p> <p>1 mark answer: Weak and/or brief written response e.g. the results are not significant or there is no relationship between extroversion and salary</p> <p>0 marks: No credit worthy information</p>		
		Total	9	
122		<p>Likely answers: don't know reasons why personality may / may not be related to salary earned; doesn't differentiate between different types of extroversion;</p> <p>3 mark answer: Clear outline of weakness in context</p> <p>2 mark answer: Clear outline of weakness but not in context OR attempted outline in context</p> <p>1 mark answer: Brief and/or weak attempt (whether in context or not)</p> <p>0 marks: No credit worthy information</p>	3	<p>Context = personality, introvert, extrovert, money, salary etc</p> <p>Do not credit weaknesses of correlations</p>
		Total	3	
123		<p>Likely answers: demand characteristics affecting accuracy of extroversion self-rating; honesty regarding salary details; problems interpreting extroversion rating scale etc</p> <p>3 marks for each thing</p> <p>3 mark answer: Clear outline in context</p> <p>2 mark answer: Clear outline but not in context OR attempted outline in context</p>	6	<p>Context = personality, introvert, extrovert, money, salary etc</p> <p>Can be strengths or weakness.</p> <p>Could outline two things that have a positive effect on the validity. Two things that have a negative affect on validity. One thing that is positive and one that is negative</p> <p>Could also outline a thing that might have affected the validity and the response explains that it is unclear</p>

		<p>1 mark answer: Brief and/or weak attempt (whether in context or not)</p> <p>0 marks: No credit worthy information</p>		<p>whether this affect would have been positive or negative.</p> <p>All types of validity creditworthy including, for example, population validity and temporal validity</p>
		Total	6	
124	a	<p>Example response: Why did you eat the food during the study? Describe to me why you think the colour of the plate might have influenced how much food you ate.</p> <p>2 mark answer: Clear outline in context</p> <p>1 mark answer: Clear outline but not in context OR Attempt in context</p> <p>0 marks: No credit worthy information</p>	2	<p>-Context = food, meal(s), eating, plate(s), crockery etc</p> <p>No marks awarded to closed questions.</p> <p>Can be awarded mark for both a question or a statement</p> <p>Do not necessarily have to ask a specific question for full marks</p>
	b	<p>Likely answers: provides rich and detailed responses, allows participants to explain their behaviour in the experiment, doesn't allows differences between participants to be ranked, is subjective, does not enables use of descriptive statistics (e.g. mean), some motivated participants may write more detailed responses compared to other participants, less reliable etc</p> <p>3 mark answer: Clear evaluation in context</p> <p>2 mark answer: Clear evaluation but not in context OR Attempt in context</p> <p>1 mark answer: Brief and/or weak attempt (whether in context or not)</p> <p>0 marks: No credit worthy information</p>	3	<p>-Context = food, meal(s), eating, plate(s), crockery etc</p> <p>Accept positive and/or negative evaluation points as creditworthy</p> <p>Can get full credit for one point although this will need to be a detailed point in context.</p>
		Total	5	

125		<p>External validity can refer to aspects of ecological validity and or population validity.</p> <p>5-6 mark answer: Clear evaluation with two or more points with some context (for 6 marks two of the points must be in context)</p> <p>3-4 mark answer: Clear evaluation with two or more points but not in context OR One clear evaluation point in context OR Two points, with one in context</p> <p>1-2 mark answer: Attempted evaluation (whether in context or not)</p>	6	<p>Context = food, meal(s), eating, plate(s), crockery etc</p> <p>Accept positive and/or negative evaluation points as creditworthy</p> <p>Do not accept evaluation points related to the methodology used in general</p> <p>Ignore reference to usefulness</p>
		Total	6	
126	a	<p>Ranking data refers to assigning numbers to denote position in an ordered sequence (lowest to highest or vice versa)</p> <p>2 mark answer: An explanation in context</p> <p>1 mark answer: An attempted explanation whether in context or not</p> <p>0 marks: No credit worthy information</p>	2	Context = personality, introvert, extrovert, money, salary etc
	b	<p>They both have the same extroversion rating (65), which in an ordered sequence of all the extroversion ratings occupy rank positions 6 and 7, so ranks 6 and 7 have been shared (added together and divided by two), resulting in a rank of 6.5 for both.</p> <p>2 mark answer: Clear explanation</p> <p>1 mark answer: Attempted explanation</p> <p>0 marks: No credit worthy information</p>	2	Could refer to having a tied rank.

			Total	4	
127	a	 <p>Scatter diagram showing the relationship between annual salary and extroversion</p> <p>1 mark for accurate plotting of data 1 mark for title with reference to both co-variables and relationship/correlation 1 mark for clear labelling of both axes 1 mark for scales on both axes</p>	4	- If points joined up no mark for 'accurate plotting of data'	
	b	<p>Any 2 of the following:</p> <ul style="list-style-type: none"> - There is a positive correlation between annual salary and extroversion/extroversion self-rating. (2 marks) <p>There is a correlation between annual salary and extroversion. (1 mark)</p> <ul style="list-style-type: none"> - Extroversion does not necessarily cause a high salary as correlations do not show cause and effect (2 marks) - There is an anomaly with one participant has a high extroversion score but a low annual salary.(2 marks) <p>2 mark answer: Clear conclusion in context</p> <p>1 mark answer: Clear conclusion but not in context OR Attempted conclusion in context</p> <p>0 marks: No credit worthy information</p>	4	-Context = personality, introvert, extrovert, money, salary etc	
			Total	8	
128		<p>$775/12=$ Mean 64.583333 Two SFs = 65</p> <p>3 mark answer: Mean correctly stated to two significant figures with all workings shown</p>	3		

		<p>2 mark answer: Mean correctly stated to two significant figures but with no workings shown OR Workings shown but mean not written to two significant figures</p> <p>1 mark answer: Mean only stated but not to two significant figures</p> <p>0 marks: No credit worthy information</p>		
		Total	3	
129		<p>Mean is a more sensitive measure of central tendency to use as there are no outliers in the results for extroversion self-ratings.</p> <p>The mean takes into account all of the data and is more mathematically accurate than the median which does not take into account all of the scores.</p> <p>2 mark answer: Clear explanation in context</p> <p>1 mark answer: Attempted explanation whether in context or not</p> <p>0 marks: No credit worthy information</p>	2	Context = personality, introvert, extrovert, money, salary etc
		Total	2	
130		A dispersion around the mean	1	
		Total	1	
131		C nominal	1	
		Total	1	
132		D 0.05	1	
		Total	1	
133		B 4	1	

			Total	1	
134			D positively skewed	1	
			Total	1	
135			C data is normally distributed in the population	1	
			Total	1	
136			C 0.0061	1	
			Total	1	
137			D raw	1	
			Total	1	
138			B median	1	
			Total	1	
139			B median	1	
			Total	1	
140			A Chi-square	1	
			Total	1	
141			C Mann-Whitney U Test	1	
			Total	1	
142			C $p < 0.05$	1	
			Total	1	
143			B level of data	1	
			Total	1	
144			Explain one way in which Freud's study of Little Hans could be criticised for its lack of reliability. <u>Possible content:</u>	3	3 marks for a clearly described issue associated with reliability which is well developed in the context of the study.

		<ul style="list-style-type: none"> • use of case study method • use of secondary data • many ideas open to interpretation • unrepresentative sample <p><u>Example of a 1-mark answer:</u></p> <p>The use of the case study method produces unreliable findings as it cannot be replicated (1). (No context)</p> <p><u>Example of a 2-mark answer:</u></p> <p>Freud only studied Little Hans so his sample was unrepresentative (1) making the study unreliable (1).</p> <p><u>Example of a 3-mark answer:</u></p> <p>The case study method is unreliable as it not replicable (1) as we only have Freud's account and interpretation of Little Hans' experiences at the time (1) and it would be difficult for someone else to revisit this after the event to assess for reliability (1).</p> <p>Freud investigated Little Hans' fears and fantasies using the case study method which lacks reliability as it does not have a standardised procedure (1). The way in which Hans was observed and questioned by his father was not standardised as the situation was unique to Hans (1). The procedure cannot therefore be replicated to assess whether other boys experience the Oedipus complex/go through the five psychosexual stages of development (1).</p> <p>Other appropriate answer.</p>		<p>2 marks for a reasonably described and relevant issue associated with reliability loosely or vaguely linked to Freud's study</p> <p>1 mark for the mere identification of a relevant issue which is not in the context of the study.</p> <p>0 marks – no creditworthy response.</p>
		Total	3	