

Assessment Schedule 2018**Mathematics and Statistics (Statistics): Evaluate statistically based reports (91584)****Evidence Statement**

Q	Expected Coverage	Achievement (c)	Merit (j)	Excellence (i)
ONE (a)	Control group: People who continue to use Facebook normally. Treatment group: People who stop using Facebook. Must be described with reference to Facebook use.	The control group AND the treatment group are described.		
(b)	People should be randomly allocated into the Facebook and non-use Facebook groups to attempt to create two unbiased (fair or balanced) groups for comparison. There will be other factors that affect happiness / life satisfaction that cannot be controlled by the researchers, and the use of random allocation allows the researchers to determine whether Facebook usage or not was the cause of any change in happiness / life satisfaction levels.	Comments on the use of random allocation as a good design feature of an experiment.	Explains how random allocation is used to reduce bias or create two fair or balanced groups. OR Clear discussion that not all factors that could affect life satisfaction can be controlled.	
(c)	The researchers could have found the difference between the mean of the treatment group (8.12) and the mean of the control group (7.75) and then carried out a randomisation test to assess the significance of the difference between the two group means (difference = 0.37). To claim “significantly higher”, the tail proportion from the randomisation test would need to be interpreted. OR The control group’s level of satisfaction increased by 0.08 points while the Treatment group increased by 0.56 over the week. This shows the Treatment group reported an increase in satisfaction of almost half a satisfaction point more than the control group. This is seven times larger, which seems big enough to ensure the difference is significant in comparison to chance.	Identifies need to test for difference of two means between control and treatment (Now). OR Identifies difference before and now for treatment group.	Makes statement comparing differences of means between treatment group and control group. OR Makes statement comparing the difference in Now means.	Full discussion of the use of randomisation test. OR Discussion of difference not just being due to chance. OR Justification of significant difference.

(d)	<p>Issue(s) could include:</p> <ul style="list-style-type: none"> • Issue with measuring response variables The claim states ‘positive effects’ – this could involve a variety of different aspects such as having more time, feeling happier, spending more time with family, less anxiety etc. The study has reported only on one aspect, life satisfaction, seeing a positive change and this would not necessarily be what others interpret as positive effects from the claim. The claim is very broad, therefore the limit in the question asked weakens the claim made. • Issue with isolating effect of treatment variable The assurance of each participant not using Facebook for a week in the treatment group. How confident could the study designers be that all the Treatment participants definitely did not use Facebook. The claim that “social media holiday” increases happiness could not be claimed in this case. • Short term effect The claim gives no time frame for the social media holiday. The design of the study only involved 1 week without Facebook, so therefore results may only apply in the short term and not be seen in a longer Facebook holiday. • Lack of blinding • Claim is social media, not just Facebook <p><i>NOTE: Discussion at all levels must be focused on claim made, not general discussion of study design features. Accept other valid design-related issues.</i></p>		<p>One relevant issue with study design described .</p>	<p>One relevant issue with study design fully described in terms of how it weakens the causal claim, with clear links to context, including quotes from the report or other relevant evidence.</p>
-----	---	--	---	--

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of c	2 of c OR 1 of j	3 of c OR 1 of c and 1 of j	2 of j	3 of j	1 of i	2 of i

Q	Expected Coverage	Achievement (c)	Merit (j)	Excellence (i)
TWO (a)(i)	<p>The margin of error is $\frac{1}{\sqrt{n}}$</p> <p>$= \frac{1}{\sqrt{2600}} = 0.0196$ or approximately 2%.</p>	<p>Calculation of reported MOE demonstrated.</p>		
(ii)	<p>For example –</p> <p>18% claiming not to do any exercise. This survey percentage is outside of the 30% to 70% range</p> <p>$\frac{1}{\sqrt{n}}$</p> <p>that the rule of thumb $\frac{1}{\sqrt{n}}$ can be applied to. The rule of thumb MOE will overestimate the size of the MOE.</p> <p>OR</p> <p>38% can't live without foundation. This will be a percentage for a subset of those surveyed (those who buy make-up) and so the sample size will be smaller, and the MOE will be bigger than the reported MOE.</p>	<p>Identifies any one of the reported survey percentages in context that are less than 30% or greater than 70%.</p> <p>AND</p> <p>refers to 30% to 70% range.</p> <p>OR</p> <p>Identifies any one of the reported survey percentages in context that are from a subset.</p> <p>AND</p> <p>Comments on the subgroup / sample size.</p>	<p>Identifies any one of the reported survey percentages that are less than 30% or greater than 70%.</p> <p>AND</p> <p>Explains that the rule of thumb MOE will overestimate the size of the MOE.</p> <p>OR</p> <p>Explains that the MOE will be bigger since the subgroup will be smaller than 2600.</p>	
(b)	<p>For a claim of “most New Zealand adults”:</p> <ul style="list-style-type: none"> • People have been surveyed via ‘Neighbourly’ – this will only include people using this social media platform. This is not a representative sample of all New Zealander adults. (NSE) <p>The social media group is likely to be made up of middle aged, house owners as they are concerned about their neighbourhood. This means the sample is likely to overrepresent wealthier, older people who will likely spend more to buy quality products. This means the percentage could be lower than the 56% quoted.</p> <ul style="list-style-type: none"> • People who are associated with Fresh Choice – likely this subgroup will not be representative of all New Zealanders. • It is probably safe to assume that not all the 2600 answered that particular question, so the margin of error would be bigger than 2%. If the margin of error was bigger than 6%, then the confidence interval would not sit entirely above 50%, so a “most” claim cannot be made (MOE). 		<p>Response acknowledges issue with either:</p> <p>WHO answered the question.</p> <p>AND</p> <p>Comment on representativeness.</p> <p>OR</p> <p>HOW many people answered the question.</p> <p>AND</p> <p>Subgroup involved (MOE).</p>	<p>Response acknowledges issues with either WHO OR HOW many people answered the question.</p> <p>AND</p> <p>Discusses this potential issue in reference to the “most New Zealand adults” claim.</p>

(c)	<p>Comparison within one group: $2 \times \text{MoE} = 2 \times 2\% = 4\%$ Difference in poll percentage = $55\% - 32\% = 23\%$ Confidence interval generated as a result [19%, 27%] <u>I'm pretty sure that the percentage of New Zealand adults who get between seven to eight hours of sleep each night is somewhere between 19% to 27% higher than those who get five to six hours of sleep. (Should use percentage points.)</u> Both limits of the confidence interval are positive, so yes, it can be claimed that a higher percentage of NZ adults sleep 7 to 8 hours than 5 to 6 hours.</p>	<p>ONE confidence interval for the difference of two proportions correctly calculated OR TWO confidence intervals correctly calculated and used in context to discuss claim about higher percentage.</p>	<p>ONE confidence interval for the difference of two proportions correctly calculated. AND ATTEMPTED to either: Interpret in context. OR Used to discuss claim in context.</p>	<p>ONE confidence interval for the difference of two proportions correctly calculated. AND Interpreted in context. AND Used and Justified to discuss claim in context.</p>
-----	---	--	---	--

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of c	2 of c OR 1 of j	3 of c OR 1 of c and 1 of j	2 of j	3 of j	1 of i	2 of i

Q	Expected Coverage	Achievement (c)	Merit (j)	Excellence (i)
THREE (a)	<p>Explanatory variable: Number of hours of screen time per day (accept TV time). Response variable: Number of GCSE points. Must be clear both variables are numeric.</p>	<p>The explanatory variable. AND Response variable stated.</p>		
(b)	<p>The variables <i>Number of hours of screen time per day</i> and <i>Number of GCSE points</i> could have been used to construct a scatterplot, with screen time as the explanatory variable, and GCSE points as the response variable. A linear model would be fitted to the data, and the gradient of this model would have been used to quantify the relationship between screen time and GCSE points. To evaluate this claim, you would need to know the strength of the linear relationship between these two variables.</p>	<p>Use of scatter plot to display data.</p>	<p>The gradient of the linear model is linked to the claim / inference made. OR The need to assess the strength of evidence is discussed, in terms of being uncertain about the amount of variation.</p>	
(c)	<p>A potential confounding variable could be parental influence. Students might have strict parents that both limit their screen time and also encourage them to study for their examinations, which would lead to better examination results. Another potential confounding variable could be health. Students with poor health might have more screen time as this is their entertainment (not fit / healthy so can't participate in other activities) and poor health also affects grades. Time spent studying or doing homework <i>Accept other relevant potential confounding variables.</i></p>		<p>A potential confounding variable is identified. AND Must discuss how it impacts / interacts with one of the variables AND Must comment on all three variables.</p>	<p>A relevant potential confounding variable is identified, with a clear reason given as to why this variable may be confounding. AND How it would interact with the two variables.</p>
(d)	<p>Report 3A headline is <i>Every hour you spend in front of a screen is linked to poorer examination results</i> and Report 3B headline is <i>Screen time affects grades</i>. As the study is an observational study, not an experiment, causal claims should not be made. Report 3B reads like a causal relationship has been identified as it uses the word "affects". Therefore, Report 3A has the more appropriate headline as it uses the word "linked", which indicates a correlation between the screen time and grades.</p>	<p>Identifies Report 3A has more appropriate headline, with supporting comments, e.g. Negative Trend. OR Casual / correlation. OR Observational study OR Affects / links.</p>	<p>Explains why Report 3A has more appropriate headline, with clear justification based on two of the three: Casual / Correlation AND / OR Observational study. AND / OR Affects / links.</p>	<p>Explains why Report 3A has more appropriate headline with clear discussion of insight into observational study design . AND Discussion of causal / correlation in terms of the language used in headlines.</p>

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of c	2 of c OR 1 of j	3 of c OR 1 of c and 1 of j	2 of j	3 of j	1 of i	2 of i

Cut Scores

Not Achieved	Achievement	Achievement with Merit	Achievement with Excellence
0 – 7	8 – 12	13 – 18	19 – 24