Assessment Schedule 2020
Mathematics and Statistics (Statistics): Evaluate statistically based reports (91584)
Evidence Statement

Q	Expected Coverage	Achievement (c)	Merit (j)	Excellence (i)
ONE (a)(i)	Using rule of thumb: $$\frac{1}{\sqrt{876}} = 0.0338$$ So margin of error = 3.4%	Correct margin of error.		
(ii)	The margin of error is needed to take into account the variation we would expect to see from survey to survey just through the process of sampling.	Provides a generic template answer about the variation expected due to sampling.		
(iii)	If the sample size was increased, then the margin of error would become smaller. We want $\frac{1}{\sqrt{n}} \leq 0.02$, that is, the margin of error = at most 2% The CTU would need to increase their sample size to 2500 in order to reduce the margin of error to at most 2%.	States that the sample size would need to be increased	Calculates the sample size correctly. Working not required.	

| (b) | According to the report, 56% of the survey respondents were unable to afford necessities like rent, power and petrol.

Using margin of error of 3.4%:

Confidence interval produced = 56% ± 3.4% [52.6% , 59.4%]

We can be fairly sure that the true percentage of workers who claim that they cannot afford basic necessities is somewhere between 52.6% and 59.4%.

There is evidence to support the CTU's claim over half of the workers cannot afford basic necessities as these percentages are both above 50%. (or CI > 50%.)

Accept other expressions of some uncertainty with the confidence interval provided such as "It's a fairly safe bet…" or "With 95% confidence…"

95% certain → illustrates uncertainty.

Context → must mention workers and basic necessities. | Calculates correct confidence interval. | Confidence interval **correctly** calculated.

AND

Interpretation in **context** given.

OR

Used to make a majority claim in **context**. | Confidence interval **correctly** calculated.

AND

Interpretation in **context** given.

AND

Used to make a majority claim in **context**.

Note: Both the claim and the judgement must be separate statements, and both must be in context. |
| (c) | Margin of error for a confidence interval comparing two independent groups

2019, $n = 1176$, so MoE $= \dfrac{1}{\sqrt{1176}} = 0.0292$, so 2.9%

2020, $n = 876$, so MoE $= \dfrac{1}{\sqrt{876}} = 0.0338$, so 3.4%

Average MoE = 3.15%

Average MoE × 1.5 = 4.7%

Difference in survey percentages from 2019 to 2020 = 11%

Average MoE × 1.5 applied to difference in poll percentage to give 95% confidence interval = [6.3% , 15.7%]

We can be fairly sure that the true proportion of workers in 2020 who think their weekly hours of work has got worse is between 6.3% and 15.7% lower than in 2019.

As the confidence interval is entirely positive, this means the proportion of workers who think their hours of work has got worse has decreased from 2019 to 2020. Claim is supported.

Accept other expressions of some uncertainty with the confidence interval provided such as "It's a fairly safe bet…" or "With 95% confidence…"

95% certain → illustrates uncertainty.

Context → must mention workers in 2019 / 2020 and weekly hours of work. | Average MOE × 1.5 AND difference between survey percentages (2019 and 2020) and confidence interval correctly calculated.

OR

Or student has correctly calculated an incorrect CI and both interpretation and claim is in context. | Confidence interval correct

AND EITHER

Interpret in **context**.

OR

A response to the claim is stated using the **context** of hours of work has got worse between 2019 and 2020. | Confidence interval correct

AND

Interpret in **context**.

AND

A response to the claim is stated using the **context** of hours of work has got worse between 2019 and 2020.

Note: Both the claim and the judgement must be separate statements, and both must be in context |

| (d) | The sampling frame is those workers who are members of unions connected to the CTU.

Because some workers may not have internet access as they can't afford this necessity, this means that the survey may not be representative of the general workforce of NZ

Selection bias is an issue for this study because the sampling frame does not match the target population. The survey is only available to workers who are members of unions affiliated to the CTU. This is not representative of the general workforce in New Zealand.

The survey was undertaken between 3-5 January. This is a time when most people are on holiday, and that maybe at this time, those that are working may be more likely to be CTU members or affiliated and that this is not representative of the general NZ working population | Correctly identifies sampling frame.

OR

Discussion regarding those workers that don't have internet because they can't afford necessities so not representative.

OR

Discussion regarding the fact that the survey was taken 3–5 January. Must give reason why this is an issue with respect to workers at that time not being representative to the general NZ working population. | States selection bias is present due to the sampling frame not matching the target population. (CTU members only). | |

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 j and 1 c	2 of j Or 1 j and 2 c	3 of j	1 of i	2 of i

Q	Expected Coverage	Achievement (c)	Merit (j)	Excellence (i)
TWO (a)	It is an experiment because the scientists deliberately manipulated the amount of sleep the subjects had during one of the phases of the experiment (normal sleep vs reduced sleep).	Correctly gives a description as to why this is an experiment. Must be in **context** of this experiment. Generic statements about experiments not sufficient.		
(b)	The method of showing the observers the subjects' photos in a randomised order from one subject to the next means that the observers are not at risk of getting 'trained' to see one particular type of photo first across all 23 subjects. If a pattern was able to be picked up on by the observers, then the ordering could be attributed as the reason why they gave different results from normal sleep to sleep-deprived. The random order means that any differences in the observed results can be attributed to the deliberately manipulated sleep times.	Describes that there could be a pattern observed.	Clear explanation of how the randomising of order process alleviates any issues regarding the observers noticing a pattern and is then focussed on the intervention. AND / OR If the subjects' photos were not randomised this could have an effect on the difference of the scores between the normal sleep and sleep-deprived ratings.	
(c)	Taking two photographs of each subject allows the researchers to calculate the change in attractiveness measured by each observer between the two conditions (normal sleep vs sleep-deprived). Each person then acts as their own control (baseline), because each person would have a pre-existing attractiveness to the observer (and this design feature minimises variation).	Calculation of change or difference identified.	Explanation about need to measure the change in attractiveness for each subject, since each subject is different in terms of pre-existing attractiveness to the observer.	

(d)	This is a well-designed experiment that uses randomisation with regards to the order of photos shown, hence a causal claim could be made. BUT the observed difference needs to be tested in some way to show that the result is due to more than just chance, for example whether the difference is as a result of the effect of sleep deprivation, or by random occurrence. AND / OR A randomisation test could be used to assess the significance of the observed difference.	Indicates that as an experiment a causal claim could be made.	States that because the experiment is well-designed / uses randomisation, a causal claim could be made.	States that because the experiment is well-designed / uses randomisation, a causal claim could be made. AND EITHER discussion of the use of randomisation test. OR Discussion of difference not just being due to chance. OR Discussion of potential issues in the treatment methodology that does not allow a claim to be made.
(e)	• The experiment was conducted in Sweden, which has different customs and norms compared to NZ and other countries. Applying the results to a country like NZ where sleep habits, culture, fashion, perceptions, urban / rural demographic, etc. could all potentially impact on the perception of the variables measured in this study. If conclusions for NZ were sought, then conducting the study with good experimental design principles should take place in NZ with a wide and large enough demographic. • The study was undertaken in 2010. The trends in what people find attractive may have significantly altered since then. For example, facial hair.	Describes ONE potential issue with extending the results.	Describes ONE potential issue with extending the results. AND Attempts to explain why it could limit extending the results.	Describes ONE potential issue with extending the results. AND Describes why each could limit extending the results by using specific features of the report / study.

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 j and 1 c	2 of j Or 1 j and 2 c	3 of j	1 of i	2 of i

Q	Expected Coverage	Achievement (c)	Merit (j)	Excellence (i)
THREE (a)	Largest single concern for child's health was sleep at 31%. OR Additionally, 41% of parents who had concerns for their child's health also thought it was difficult to find the necessary help.	Identifies at least one piece of supporting data, reinforced with numerical evidence.		
(b)	The question contains four aspects which are open to subjective judgement, therefore standardising results from person to person who took part in the survey would be very difficult to achieve. • Perception – this would require the survey participant recalling data across the criteria given from memory, sometimes decades ago. • Participate – some parents may interpret this at full competitive standard at a minimum level, others may take this to be exactly what it means. • Sport – some activities may be considered not a sporting venture. Did the survey question give a definition? • How much? Some parents / children may play more sports less often.	Identifies one aspect of the wording of the question.	Identifies two aspects of the wording of the question and the potential impact on results for at least one aspect.	
(c)	The insurance company funded the survey. They have a commercial interest in people buying premiums to protect themselves against unfortunate medical episodes / injuries. The report includes a quote from the CEO who said, "this particular statistic should be encouraging news for healthcare and related providers who have, or are looking to incorporate, online or digital solutions into their offer".	Identifies that the company that funded the survey is an insurance company and describes one way the survey could be used to the company's advantage.	Explains why the findings of the survey could be used to the company's advantage with reference to the survey results reported. (reference items in report)	
(d)	The 1200 parents were surveyed by a health insurance company. These people may have different perceptions of health compared to all New Zealanders, e.g. be more health conscious, and that is why is why they have health insurance. OR Self-selection bias may be present. The respondents may have chosen to take part in the study, so might not be a representative sample of all New Zealanders. *(Need a more detailed reason.)*		Describes one example of selection bias.	One possible example of selection bias is described AND a reasonable example of how it could cause bias in this survey is given.

| (e) | Comparison within a group
Poll percentage difference 1%
$2 \times \text{MoE} = 2 \times 2.9\% = 5.8\%$
Applied to poll percentage difference gives a 95% CI [–4.8% , 6.8%]
We can be fairly sure that the proportion of parents who feel their children do less sport than them at a similar age is up to 6.8% greater than those who feel their children do about the same, but it could be as much as 4.8% lower. (or similar text)
As the poll percentage difference could be either positive or negative, (or the amount spans zero) the claim parents who feel their children do less sport than them at a similar age is higher than those who feel their children do about the same is rejected.
Accept other expressions of some uncertainty with the confidence interval provided such as "It's a fairly safe bet…" or "With 95% confidence…"
95% certain → illustrates uncertainty.
Context → must mention parents and less / same amount of sport. | Confidence interval correct using MOE × 2 and difference between those parents that feel their children did less sport than they did and those that did about the same amount.
OR
Student has correctly calculated an incorrect CI and both interpretation and claim is in context. | Confidence interval correct
AND EITHER
Interpret in context.
OR
A response to the claim made by the company is stated using the context of parents that feel their children did less sport than they did and those that did about the same. | Confidence interval correct
AND
Interpret in context.
AND
A response to the claim made by the company is stated using the context of parents that feel their children did less sport than they did and those that did about the same.
Note: Both the claim and the judgement must be separate statements, and both must be in context. |

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 j and 1 c	2 of j Or 1 j and 2 c	3 of j	1 of i	2 of i

Cut Scores

Not Achieved	Achievement	Achievement with Merit	Achievement with Excellence
0 – 8	9 – 13	14 – 18	19 – 24