
Introduction

This document provides the technical details about the calculations used for Student Experience Survey (SES) data used on the ComparED website. It is intended for an audience with some technical and data background who wish to understand the statistical details of the calculations.

Data sources, variables and coverage

Data sources

Student Experience Survey (SES)

The SES is Australia's national survey of higher education coursework experience. Students in their first and later years of their course are surveyed in the second semester and asked about their experience at their higher education institution. The following SES data are used:

- *Quality of educational experience (item)*: the proportion of students who rated the quality of their overall educational experience positively;
- *Teaching quality (scale)*: the proportion of students who rated the quality of teaching they had experienced positively, as measured by the items in the teaching quality scale;
- *Learner engagement (scale)*: the proportion of students who rated learner engagement at their institution positively, as measured by the items in the learner engagement scale;
- *Learning resources (scale)*: the proportion of students who rated the learning resources provided by their institution positively, as measured by the items in the learning resources scale;
- *Student support (scale)*: the proportion of students who rated the support they received at their institution positively, as measured by the items in the student support scale; and
- *Skills development (scale)*: the proportion of students who rated the skills development they experienced through their studies positively, as measured by the items in the skills development scale.

Results are calculated separately for undergraduate and postgraduate coursework level students. Indicators are calculated from two years of pooled data. This incorporates the most recent year of published data and the immediately preceding year. For example, indicators released in association with the 2018 SES were based on results from the 2018 and 2017 surveys. In this paper these years are notated as Y1 and Y2, where Y1 is the most recent year of published data.

The variables that were used to filter the data can be found in Tables 1 and 2 below. The coverage for each variable is applied before the calculation of the indicators and the SAS code used is provided in brackets after each variable in the table. The full SAS and/or R code to create the indicators is available from the Social Research Centre (SRC) on request.

Table 1: Data coverage for the SES based indicators – undergraduate level

Variables (coverage)	Indicators					
	Quality of educational experience	Teaching quality	Learner engagement	Learning resources	Student Support	Skills Development
SES Survey data file:						
Undergraduate level students only (if crslevel = 1)	X	X	X	X	X	X
In scope including different study areas for double degree students (if analysis in (1,2))	X	X	X	X	X	X
Question regarding quality of educational experience answered (overall in (1,2,3,4))	X					
Valid Teaching Quality scale score (if teachsat in (0,1))		X				
Valid Learner Engagement scale score (if engagsat in (0,1))			X			
Valid Learning Resources scale score (if resrsat in (0,1))				X		
Valid Student Support scale score (if suppsat in (0,1))					X	
Valid Skills Development scale score (if develsat in (0,1))						X
Total minimum sample size of 25 across the 2 years of pooled data (if n ≥ 25)	X	X	X	X	X	X
Institutions with data in Y1 (if 1+ valid overall record in Y1)	X	X	X	X	X	X
SES Frame file:						
Enrolled at time of frame checking (if enr1stat = 0)	X	X	X	X	X	X
Undergraduate level students only (if crslevel = 1)	X	X	X	X	X	X
In scope for the SES – including those randomly excluded from large strata (if exclude in (0,6,7))	X	X	X	X	X	X
Not contained in extra quota (if finframe ≠ 2)	X	X	X	X	X	X
Count two records for double degree students in different study areas	X	X	X	X	X	X

X Indicates that the restriction is applied to the data before a particular indicator is calculated.

Table 2: Data coverage for the SES based indicators – postgraduate coursework level

Variables (coverage)	Indicators					
	Quality of educational experience	Teaching quality	Learner engagement	Learning resources	Student Support	Skills Development
SES Survey data file:						
Postgraduate coursework level students only (if crslevel = 2)	X	X	X	X	X	X
In scope including different study areas for double degree students (if analysis in (1,2))	X	X	X	X	X	X
Question regarding quality of educational experience answered (overall in (1,2,3,4))	X					
Valid Teaching Quality scale score (if teachsat in (0,1))		X				
Valid Learner Engagement scale score (if engagsat in (0,1))			X			
Valid Learning Resources scale score (if resrsat in (0,1))				X		
Valid Student Support scale score (if suppsat in (0,1))					X	
Valid Skills Development scale score (if develsat in (0,1))						X

Total minimum sample size of 25 across the 2 years of pooled data (if n ≥ 25)	X	X	X	X	X	X
Institutions with data in Y1 (if 1+ valid overall record in Y1)	X	X	X	X	X	X
SES Frame file:						
Enrolled at time of frame checking (if enr1stat = 0)	X	X	X	X	X	X
Postgraduate coursework level students only (if crslevel = 2)	X	X	X	X	X	X
In scope for the SES – including those randomly excluded from large strata (if exclude in (0,6,7))	X	X	X	X	X	X
Not contained in extra quota (if finframe ≠ 2)	X	X	X	X	X	X
Count two records for double degree students in different study areas	X	X	X	X	X	X

X Indicates that the restriction is applied to the data before a particular indicator is calculated.

Data variability

Standard errors, and therefore the 90% confidence intervals, were calculated using the Finite Population Correction (FPC) to account for the relatively large samples when compared with the population. The FPC is generally used in cases where the sampling fraction, the proportion of the population sampled, exceeds 5%.

In order to calculate the standard errors for the survey estimates, no non-response bias was assumed and the Agresti-Coull method for confidence intervals for proportions was used.

The general formula used for confidence intervals for proportions was:

$$CI\ bound(\hat{p}) = \tilde{p} \pm z_{\frac{\alpha}{2}} \times FPC \times SE(\tilde{p}) = \tilde{p} \pm z_{0.05} \times \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}$$

$$\tilde{p} = \frac{\tilde{y}}{\tilde{n}}$$

$$\tilde{y} = y + \frac{z_{\frac{\alpha}{2}}^2}{2} = y + \frac{z_{0.05}^2}{2}$$

$$\tilde{n} = n + z_{\frac{\alpha}{2}}^2 = n + z_{0.05}^2$$

Where:

\hat{p} is the estimated proportion from the survey data

\tilde{p} is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

$z_{0.05}$ is the 95th quantile from the standard Normal distribution ~ N(0,1)

n is the number in the sample in the relevant strata over the pooled relevant number of years

N is the estimated population in the relevant strata over the pooled relevant number of years

y is the number with the characteristic in question in the sample in the relevant strata over the pooled relevant number of years

As mentioned above, the FPC was used to calculate the 90% confidence intervals and the standard errors, therefore an estimate of the population was required. The survey frame was used to determine the population, see Tables 1 and 2 for the restrictions applied. The coverage for each

variable is applied before the calculation of the population and the SAS code used is provided in brackets after each variable in the table.

Calculation of indicators and confidence intervals

Quality of educational experience

The quality of educational experience indicator is defined as the proportion of students who indicated that the quality of their entire educational experience was good or excellent. The indicator can be expressed as ‘the proportion of students who rated the quality of their educational experience positively’.

The quality of educational experience indicator is calculated as follows:

$$QEE_{pooled} = \frac{\text{Number of students rating the quality of their educational experience positively}_{Y2-Y1}}{\text{Number of students with a valid response}_{Y2-Y1}}$$

Where:

*Number of students rating the quality of their educational experience positively*_{Y2-Y1} is the total number of students who responded with a 3 or 4 (good or excellent) to the quality of entire educational experience item ‘Thinking of this year, overall at your <university/institution>, how would you rate the quality of your entire educational experience?’ (**overall in (3,4)**), over the two years combined, after filters are applied. It should be noted that this item is reported on a four point scale

*Number of students with a valid response*_{Y2-Y1} is the total number of students who responded to the quality of entire educational experience item (**overall in (1,2,3,4)**), over the two years combined, after filters are applied

The 90% confidence interval for the quality of educational experience indicator is calculated as follows:

$$\begin{aligned} 90\%CI_{QEE} &= \overline{QEE}_{pooled} \pm z_{\frac{\alpha}{2}} \times FPC \times SE_{\overline{QEE}} \\ &= \overline{QEE}_{pooled} \pm z_{0.05} \times \sqrt{\frac{N-n}{N-1}} \times \sqrt{\frac{\overline{QEE}_{pooled} \times (1 - \overline{QEE}_{pooled})}{\tilde{n}}} \\ \overline{QEE}_{pooled} &= \frac{\tilde{y}}{\tilde{n}} \end{aligned}$$

$$\begin{aligned} \tilde{y} &= \text{Number of students rating the quality of their educational experience positively}_{Y2-Y1} + \frac{z^2_{0.05}}{2} \\ &= \text{Number of students rating the quality of their educational experience positively}_{Y2-Y1} + \frac{1.645^2}{2} \end{aligned}$$

$$\begin{aligned} \tilde{n} &= \text{Number of students with a valid response}_{Y2-Y1} + z^2_{0.05} \\ &= \text{Number of students with a valid response}_{Y2-Y1} + 1.645^2 \end{aligned}$$

Where:

\overline{QEE}_{pooled} is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the population in the institution study area, or institution, as relevant, over the two years combined

n is the number of students with a valid response in the institution study area or institution as relevant, over the two years combined

$z_{0.05}$ is the 95th quantile from the standard Normal distribution $\sim N(0,1)$

The restrictions for this indicator can be found in Tables 1 and 2.

Teaching quality

The teaching quality indicator is defined as the proportion of students who rated the quality of teaching they had experienced positively.

The teaching quality indicator is calculated as follows:

$$TQ_{pooled} = \frac{\text{Number of students rating teaching quality positively}_{Y2-Y1}}{\text{Number of students with a valid response}_{Y2-Y1}}$$

Where:

*Number of students rating teaching quality positively*_{Y2-Y1} is the total number of students whose teaching quality scale score was at least 55 out of 100 (**teachsats=1**), over the two years combined, after filters are applied

*Number of students with a valid response*_{Y2-Y1} is the total number of students who had a valid response (**teachsats in (0,1)**), i.e. responded to at least eight of the eleven teaching quality items, over the two years combined, after filters are applied

The 90% confidence interval for the teaching quality indicator is calculated as follows:

$$\begin{aligned} 90\%CI_{TQ} &= \widehat{TQ}_{pooled} \pm z_{\frac{\alpha}{2}} \times FPC \times SE_{\widehat{TQ}} \\ &= \widehat{TQ}_{pooled} \pm z_{0.05} \times \sqrt{\frac{N-n}{N-1}} \times \sqrt{\frac{\widehat{TQ}_{pooled} \times (1 - \widehat{TQ}_{pooled})}{\tilde{n}}} \\ \widehat{TQ}_{pooled} &= \frac{\tilde{y}}{\tilde{n}} \end{aligned}$$

$$\begin{aligned} \tilde{y} &= \text{Number of students rating the quality of their teaching positively}_{Y2-Y1} + \frac{z^2_{0.05}}{2} \\ &= \text{Number of students rating the quality of their teaching positively}_{Y2-Y1} + \frac{1.645^2}{2} \end{aligned}$$

$$\begin{aligned} \tilde{n} &= \text{Number of students with a valid response}_{Y2-Y1} + z^2_{0.05} \\ &= \text{Number of students with a valid response}_{Y2-Y1} + 1.645^2 \end{aligned}$$

Where:

\widehat{TQ}_{pooled} is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the population in the institution study area, or institution, as relevant, over the two years combined

n is the number of students with a valid response in the institution study area or institution as relevant, over the two years combined

$z_{0.05}$ is the 95th quantile from the standard Normal distribution $\sim N(0,1)$

The restrictions for this indicator can be found in Tables 1 and 2.

Learner engagement

The learner engagement indicator is defined as the proportion of students who rated their engagement with learning positively. The indicator can be expressed as ‘the proportion of students who rated engagement with learning at their institution positively’.

The learner engagement indicator is calculated as follows:

$$LE_{pooled} = \frac{\text{Number of students rating learner engagement positively}_{Y2-Y1}}{\text{Number of students with a valid response}_{Y2-Y1}}$$

Where:

*Number of students rating learner engagement positively*_{Y2-Y1} is the total number of students whose learner engagement scale score was at least 55 out of 100 (**engagsat=1**), over the two years combined, after filters are applied

*Number of students with a valid response*_{Y2-Y1} is the total number of students who had a valid response (**engagsat in (0,1)**), i.e. responded to at least five of the seven learner engagement items, over the two years combined, after filters are applied

The 90% confidence interval for the learner engagement indicator is calculated as follows:

$$\begin{aligned} 90\%CI_{LE} &= \bar{LE}_{pooled} \pm z_{\frac{\alpha}{2}} \times FPC \times SE_{\bar{LE}} \\ &= \bar{LE}_{pooled} \pm z_{0.05} \times \sqrt{\frac{N-n}{N-1}} \times \sqrt{\frac{\bar{LE}_{pooled} \times (1 - \bar{LE}_{pooled})}{\tilde{n}}} \\ \bar{LE}_{pooled} &= \frac{\tilde{y}}{\tilde{n}} \end{aligned}$$

$$\begin{aligned} \tilde{y} &= \text{Number of students rating their learner engagement positively}_{Y2-Y1} + \frac{z^2_{0.05}}{2} \\ &= \text{Number of students rating their learner engagement positively}_{Y2-Y1} + \frac{1.645^2}{2} \end{aligned}$$

$$\begin{aligned} \tilde{n} &= \text{Number of students with a valid response}_{Y2-Y1} + z^2_{0.05} \\ &= \text{Number of students with a valid response}_{Y2-Y1} + 1.645^2 \end{aligned}$$

Where:

\bar{LE}_{pooled} is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the population in the institution study area, or institution, as relevant, over the two years combined

n is the number of students with a valid response in the institution study area or institution as relevant, over the two years combined

$z_{0.05}$ is the 95th quantile from the standard Normal distribution ~ N(0,1)

The restrictions for this indicator can be found in Tables 1 and 2.

Learning resources

The learning resources indicator is defined as the proportion of students who rated the learning resources provided by their institution positively.

The learning resources indicator is calculated as follows:

$$LR_{pooled} = \frac{\text{Number of students rating learning resources positively}_{Y2-Y1}}{\text{Number of students with a valid response}_{Y2-Y1}}$$

Where:

*Number of students rating learning resources positively*_{Y2-Y1} is the total number of students whose learning resources scale score was at least 55 out of 100 (**resrsat = 1**), over the two years combined, after filters are applied

*Number of students with a valid response*_{Y2-Y1} is the total number of students who had a valid response (**resrsat in (0,1)**), i.e. responded to at least five of the seven learning resources items, over the two years combined, after filters are applied

The 90% confidence interval for the learning resources indicator is calculated as follows:

$$90\%CI_{LR} = \widetilde{LR}_{pooled} \pm z_{\frac{\alpha}{2}} \times FPC \times SE_{\widetilde{LR}}$$

$$= \widetilde{LR}_{pooled} \pm z_{0.05} \times \sqrt{\frac{N-n}{N-1}} \times \sqrt{\frac{\widetilde{LR}_{pooled} \times (1 - \widetilde{LR}_{pooled})}{\tilde{n}}}$$

$$\widetilde{LR}_{pooled} = \frac{\tilde{y}}{\tilde{n}}$$

$$\tilde{y} = \text{Number of students rating their learning resources positively}_{Y2-Y1} + \frac{z^2_{0.05}}{2}$$

$$= \text{Number of students rating their learning resources positively}_{Y2-Y1} + \frac{1.645^2}{2}$$

$$\tilde{n} = \text{Number of students with a valid response}_{Y2-Y1} + z^2_{0.05}$$

$$= \text{Number of students with a valid response}_{Y2-Y1} + 1.645^2$$

Where:

\widetilde{LR}_{pooled} is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the population in the institution study area, or institution, as relevant, over the two years combined

n is the number of students with a valid response in the institution study area or institution as relevant, over the two years combined

$z_{0.05}$ is the 95th quantile from the standard Normal distribution $\sim N(0,1)$

The restrictions for this indicator can be found in Tables 1 and 2.

Student support

The student support indicator is defined as the proportion of students who rated the support they received at their institution positively.

The student support indicator is calculated as follows:

$$SS_{pooled} = \frac{\text{Number of students rating student support positively}_{Y2-Y1}}{\text{Number of students with a valid response}_{Y2-Y1}}$$

Where:

*Number of students rating student support positively*_{Y2-Y1} is the total number of students whose student support scale score was at least 55 out of 100 (**suppsat = 1**), over the two years combined after filters are applied.

*Number of students with a valid response*_{Y2-Y1} is the total number of students who had a valid response (**suppsat in (0,1)**), i.e. responded to at least six of the thirteen student support items, over the two years combined, after filters are applied

The 90% confidence interval for the student support indicator is calculated as follows:

$$\begin{aligned}
 90\%CI_{SS} &= \widetilde{SS}_{pooled} \pm z_{\frac{\alpha}{2}} \times FPC \times SE_{\widetilde{SS}} \\
 &= \widetilde{SS}_{pooled} \pm z_{0.05} \times \sqrt{\frac{N-n}{N-1}} \times \sqrt{\frac{\widetilde{SS}_{pooled} \times (1 - \widetilde{SS}_{pooled})}{\tilde{n}}} \\
 \widetilde{SS}_{pooled} &= \frac{\tilde{y}}{\tilde{n}}
 \end{aligned}$$

$$\begin{aligned}
 \tilde{y} &= \text{Number of students rating their student support experience positively}_{Y2-Y1} + \frac{z^2_{0.05}}{2} \\
 &= \text{Number of students rating their student support experience positively}_{Y2-Y1} + \frac{1.645^2}{2}
 \end{aligned}$$

$$\begin{aligned}
 \tilde{n} &= \text{Number of students with a valid response}_{Y2-Y1} + z^2_{0.05} \\
 &= \text{Number of students with a valid response}_{Y2-Y1} + 1.645^2
 \end{aligned}$$

Where:

\widetilde{SS}_{pooled} is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the population in the institution study area, or institution, as relevant, over the two years combined

n is the number of students with a valid response in the institution study area or institution as relevant, over the two years combined

$z_{0.05}$ is the 95th quantile from the standard Normal distribution $\sim N(0,1)$

The restrictions for this indicator can be found in Tables 1 and 2.

Skills development

The skills development indicator is defined as the proportion of students who rated the skills development they experienced through their studies positively.

The skills development indicator is calculated as follows:

$$SD_{pooled} = \frac{\text{Number of students rating skills development positively}_{Y2-Y1}}{\text{Number of students with a valid response}_{Y2-Y1}}$$

Where:

*Number of students rating skills development positively*_{Y2-Y1} is the total number of students whose skills development scale score was at least 55 out of 100 (**devel_sat = 1**), over the two years combined, after filters are applied.

*Number of students with a valid response*_{Y2-Y1} is the total number of students who had a valid response (**devel_sat in (0,1)**), i.e. responded to at least six of the eight skills development items, over the two years combined, after filters are applied

The 90% confidence interval for the skills development indicator is calculated as follows:

$$\begin{aligned}
 90\%CI_{SD} &= \widetilde{SD}_{pooled} \pm z_{\frac{\alpha}{2}} \times FPC \times SE_{\widetilde{SD}} \\
 &= \widetilde{SD}_{pooled} \pm z_{0.05} \times \sqrt{\frac{N-n}{N-1}} \times \sqrt{\frac{\widetilde{SD}_{pooled} \times (1 - \widetilde{SD}_{pooled})}{\tilde{n}}} \\
 \widetilde{SD}_{pooled} &= \frac{\tilde{y}}{\tilde{n}}
 \end{aligned}$$

$$\begin{aligned} \tilde{y} &= \text{Number of students rating their skills development positively}_{Y2-Y1} + \frac{z^2_{0.05}}{2} \\ &= \text{Number of students rating their skills development positively}_{Y2-Y1} + \frac{1.645^2}{2} \end{aligned}$$

$$\begin{aligned} \tilde{n} &= \text{Number of students with a valid response}_{Y2-Y1} + z^2_{0.05} \\ &= \text{Number of students with a valid response}_{Y2-Y1} + 1.645^2 \end{aligned}$$

Where:

\tilde{D}_{pooled} is an adjusted estimated proportion used only in confidence interval calculations

FPC is the finite population correction which adjusts the estimated survey standard error based on how large the sample was compared with the population

N is the population in the institution study area, or institution, as relevant, over the two years combined

n is the number of students with a valid response in the institution study area or institution as relevant, over the two years combined

$z_{0.05}$ is the 95th quantile from the standard Normal distribution $\sim N(0,1)$

The restrictions for this indicator can be found in Tables 1 and 2.