GUIDELINES FOR ASKING THE NATIONAL STUDENT AND PARENT SCHOOL OPINION ITEMS

Guidelines for sampling are provided to assist schools in surveying students and parents/caregivers, using the national school opinion items. These guidelines have been provided by the Australian Bureau of Statistics (ABS) at ACARA's request.

The ABS guidelines are summarised below. However, ACARA recommends that you read the full ABS advice before deciding on census versus survey approaches, and how to create statistically valid stratified random samples if desired.

SUMMARY OF ABS RECOMMENDATIONS:

1. The target student population is children in Years 5–12 in Australian schools.
2. The target parent/caregiver population is parents of children at all year levels, in Australian schools.
3. A census approach is recommended for all schools (ie ask all students the student items; ask all parents the parent items).
4. A census approach can however introduce non response bias. Where it is anticipated that 25% or more of the parent cohort will not respond, a sampling approach should be used (based on 4 below).
5. There are also other considerations that may lead to a sampling rather than census approach (eg cost).
6. If a sampling approach is chosen, the following are recommended:
   - use random sampling, stratified by year level for students
   - as a broad rule, the smaller the cohort to be sampled, the higher proportion must be sampled to generate statistically significant data
7. Methods for generating random samples are included in the full text guidelines. The attached spreadsheet gives a guide to sample sizes for schools of different sizes.
8. Where the number of responses to a particular item is five or less, that data and related data should be suppressed, when publishing reports.

Guidelines for asking the national student and parent school opinion items

Contents

1. Purpose of this document

This document provides statistical advice relevant to asking the national student and parent/caregiver school opinion items. The document gives broad advice which will enable the practitioner to conduct their own sample design. It then details the steps involved in conducting a specific (recommended) sample design for a typical school.

2. National student and parent/caregiver school opinion items

Asking the national student and parent/caregiver school opinion items will help schools and jurisdictions gauge stakeholder opinions about school performance.

Note that from now onwards 'parents' will refer to parents and caregivers.

The national student and parent school opinion items will enable students and parents to express their opinions across a number of performance criteria using a response scale from Strongly Disagree to Strongly Agree, in response to a statement of some aspect of school performance (e.g. "My school is well maintained"). An "NA" option will also be included.

The individual scores can then be aggregated in some appropriate manner to summarise the responses for each item. Some of these functions may be available within *School Survey* (a data collection tool which schools may use to collect survey response data); others can be achieved by extracting data and working with it outside the system (for example, within Excel).

This distribution of stakeholder responses could be presented at one of the following levels:

i. the individual school level
ii. year-on-year comparative scores for an individual school, and
iii. comparative scores between an individual school and the aggregate outcome of schools in a school system or jurisdiction.

These are only some of the possible comparisons that could be made. The dimensions of interest of this sort of data are typically in four categories:

i. geographically
ii. over time
iii. socio-economically, and
iv. education administrative units.

Current plans see national student and parent school opinion items being surveyed at least every two years by participating schools and jurisdictions, which will allow for data to be tracked over time. So looking at trends in the data will of course take considerable time to evolve.

Subject to agreement, other more analytical uses to which data from the surveys might be put at some point in the future could involve the use of socio-economic information as a related variable. This would allow for a fuller picture of the survey item results to be developed. ABS Population Census information is a candidate for this sort of additional information, as it is available at the small-area level and is of high quality.

3. Principles of good survey design

In order for national student and parent school opinion item response data to be used with confidence, it must be of high quality. Such quality must be designed into the data. This design must take into account the various components of good survey design.

A good framework to refer to when discussing quality is the *Australian Bureau of Statistics Data Quality Framework 1520.0* (http://goo.gl/TjJPy). This paper highlights the major dimensions of quality and serves to illustrate how good survey design can support the following quality dimensions:

- institutional environment
- relevance
- timeliness
- accuracy
- coherence
- interpretability, and
- accessibility.

Survey design consists of a series of processes which collectively deliver estimates of the desired quality. The National Statistical Service (NSS) website (nss.gov.au) details the steps that are involved in good survey design. They are as follows:

3.1. Determining aims and objectives

It is crucial to explicitly define the purposes to which the data will be put. It is not possible to create a single survey which can be used for any desired purpose. A sample that has been designed for a particular statistical purpose may be inadequate for another.

These aims and objectives will guide the development of the survey.

3.2. Establishing Data Requirements

To ensure the aims and objectives mentioned above are fulfilled, it is good practice to determine the detailed outputs from the start. The National Statistical Service guidelines for good survey design practice recommend:

“A common way to ensure all the appropriate data items are collected is to specify, as early as possible in the planning stages of the collection, the detailed [reports] to be derived. Such specifications not only help clarify the data items to be collected, but the levels or subgroups for which the data is required. Assuming an appropriate collection methodology, defining the levels of interest beforehand will help to ensure an adequate sample size is chosen to achieve the desired levels of accuracy.”

During requirements gathering for *School Survey*, a number of standard reports were identified as required across most schools and jurisdictions. Schools are also able to generate *ad hoc* reports using data extracts.

3.3. Ascertaining the target population

The target population refers to the group of people that are of interest in a survey. In the case of the national student items, the target population consists of all students in Australia between Years 5 and 12. In the case of the national student items, the target population consists of all parents of school children in Australia.

It is recommended that schools select a sample of students (from Years 5–12) and then another sample of parents (from Years K-12). Parents will sometimes be selected more than once if they have more than 1 child at the school.

This information can be used to construct weights for the parent survey. The issue of weights is discussed further below, in paragraph 3.7 (*Producing estimates*).

3.4. Weighing up the costs and benefits of a sample versus a census

A census of a school sees all students and parents in that school included. When a sample survey is used, only a subset of the school's student/parent population is selected. There are advantages and disadvantages to conducting a census of a target population or a sample survey.

Censuses generally are costlier, as all students/parents in a school need to be approached. However, this extra cost leads to better estimates of sub-populations in particular and (assuming full response) are not subject to bias (i.e. they estimate the target population quantities of interest exactly). If the cost of enumeration for a census is similar to that of a sample survey, then consideration to undertaking a census should be given. Censuses support reliable fine-level analysis, which may be of benefit when comparisons between small groups of students are desired.

However, a census can suffer from non-response bias. This occurs when certain important sub-groups of the student/parent population tend to not complete the questionnaire. This could lead to problems in using the responding students/parents to form estimates, as they do not adequately represent the student/parent population at that school. For example, if students from certain ethnic backgrounds or students disenchanted with their school do not complete the questionnaire, then the estimates from that school will not reflect the true opinions of the students from that school. The same issues may arise in the parent survey; non-response there may potentially be higher than in the student survey, as students can be organised more easily to complete the survey during class.

Sample surveys are generally more cost-effective than censuses and can deliver very good estimates if properly designed. With a smaller sample, it is more cost-effective to follow-up non-responding students/parents and to encourage them to participate in the survey. This will minimise non-response bias.

3.5. Minimising sampling and non-sampling errors

There are two types of errors that can arise when conducting a survey.

Sampling error arises due to the manner in which a sample of students/parents is selected from a school. There are two types of samples that can be selected; a probability sample or a non-probability sample.

In a probability sample, every student/parent in the school has a known, non-zero chance of being asked the national student/parent items. These selection probabilities are given to each student/parent in the school at the time of the initial sample design and are used to select the students/parents to be administered the relevant questionnaire.

The benefit of a probability sample is that sample estimates of population quantities formed from such samples are unbiased (i.e. they estimate the true population quantity of interest exactly in repeated samples). That is to say, there is no systematic tendency for the estimates to be higher or lower than the true population value.

One well-known type of probability sample is a 'random sample'. Such a sample selects each unit into the sample with the same probability of selection. One commonly used type of sample design sees random sampling combined with 'stratification'. Stratification is a method of grouping similar students together so that the overall sample size can be reduced. If stratification works well, the students in the stratum will be very similar with respect to the study variables and only a few students need to be sampled to form an estimate. In the case of the national student and parent items, the students/parents are naturally stratified by Year, which forms a natural group of students/parents from which to select a sample.

A big advantage with probability samples is that the selected students/parents can be said to be 'representative' of the population from which they are selected. This feature has a practical advantage. It allows for measures of precision to be calculated for the estimates from the sample. The most common types of measures of precision are the standard error (SE) and the relative standard error (RSE) of the estimate.

The standard error of an estimate is a measure of how reliable that estimate is; it is the most commonly used measure of an estimate’s precision. The standard error summarises the uncertainty in an estimate due to the use of a random sample, rather than taking a census.

To illustrate what the standard error of an estimate means in practice, assume that the proportion of students or parents who respond to a national student or parent item with a particular level of agreement (e.g. Strongly Agree) is *p*. For example, if 30% of the respondents to a particular question on *School Survey* responded as Strongly Agree, then p=0.3. The formula for the standard error of *p* is given by:

$$SE(p) = \sqrt{\left(\frac{1}{n} - \frac{1}{N}\right) p(1-p)}$$

where
$n =$ sample size of students of parents from school
N= school student or parent population size
p = estimated proportion of students or parents who responded to some level of satisfaction in the survey

Note that as the sample size increases, the SE(*p*) decreases. When a census of the school is undertaken (that is, when the sample size *n* is equal to the school population size *N*) then SE(*p*) = 0. This quantifies the intuitive notion that when a census is undertaken, the sampling error is eliminated.

The above formula for SE(*p*) also shows that the standard error of a proportion is a simple function of the proportion itself.

The Relative Standard Error (RSE) of *p* is related to the SE(*p*) by the simple relationship:

$$RSE(p) = \frac{SE(p)}{p}$$

That is, the RSE is a scaled version of the SE. In the example above, the SE(*p*) when the sample of students is 50, the school population is 800 and the estimated proportion is 0.3 is:

$$SE(p) = \sqrt{\left(\frac{1}{50} - \frac{1}{800}\right) 0.3(1-0.3)}$$

The relative standard error of an estimate (normally expressed as a percentage) succinctly shows how reliable that estimate is; it only refers to the sampling error of that estimate. For example, assume that a random sample of 50 students is selected from a school with 800 students and this sample is used to calculate the estimated proportion p of the school population which report that they Strongly Agree with a certain aspect of the School. The RSE% of the estimated proportion p is given by the formula:

$$RSE\%(p) = 100\sqrt{(\frac{1}{n} - \frac{1}{N})\frac{(1-p)}{p}}$$

Hence, once the proportion is known from the sample, it is possible to calculate the RSE% of that proportion. If the proportion of students in the above example who report being Very Satisfied is 0.3, then the RSE is 20.9%, obtained by substituting into the formula above to give:

$$RSE\%(p) = 100\sqrt{(\frac{1}{50} - \frac{1}{800})\frac{(1-0.3)}{0.3}} = 20.9\%$$

This RSE% indicates that there is considerable error in the estimate. A larger sample of students from the school would reduce this RSE. If the sample selected was 200, the RSE% would be only 9.4%. Of course, if a census of all students were taken, the RSE% would be 0.

Paragraph 3.8 below expands on this concept of sampling error and shows how the RSE can be used to construct 'confidence intervals' for the estimated proportion. These confidence intervals show how reliable an estimate of proportion is.

Another type of sample is a non-probability sample. These surveys do not select people into the survey by using probabilities of selection but by other means. The most common types of non-probability samples are:

- quota samples
- convenience samples, and
- purposive samples.

None of these sampling procedures ensure the selection of representative samples.

Quota samples work by selecting students/parents based on some characteristic of interest to the study. For example, a sample may be chosen so as to contain the same proportion of students in various age-sex categories as are contained within the whole school. In this sense, the sample of students is 'representative of the target population' but only in this narrow sense of age-sex demographics. The sample may, however, not be representative of all socio-economic groups in the school. If these characteristics are important for the study, the quota sample selected will produce biased estimates of the study variable. In addition, it will be difficult to quantify this bias and adjust the estimates for it. Furthermore, measures of precision will also be unreliable.

A convenience sample is a very simple type of sample design. A typical example could be posting a question on a website and then allowing anyone to respond. Such a sample usually has little credibility. There is very little (if any) information on the people who respond to such a question and even basic information such as their demographics can sometimes be unknown. In addition, some (controversial) questions tend to attract people with unusually strong or otherwise unrepresentative views. Such a data collection methodology tends to be inexpensive and easy to administer, but suffers from being biased. In addition, sampling error is difficult to quantify in some of these studies. This is not recommended for asking the national student and parent items.

Purposive samples arise by the use of expert knowledge to select a sample of potentially representative students. Such samples are often used to gain general information about a target population, without necessarily attempting to form estimates. Such a sample is subject to the biases of the researcher selecting the sample; it is generally difficult to quantify the sampling error or the biases that arise in such a selection scheme. This sampling method will not be used for national student and parent items.

From the above discussion, the most desirable type of sample design would see a probability sample of students being selected. Such a sample would be objectively chosen and would see every student having a known probability of being included in the sample. This would allow for unbiased estimation of the desired population quantities (in the case of the national student and parent items, these quantities are the various opinion measures). It would also allow for measures of precision to be calculated, which would allow users of the estimates to gauge the level of uncertainty in the data.

3.6. Using good sample design

Sample design aims at selecting the sample of students/parents which is best suited to the production of aggregated opinion measures. This is achieved by choosing a sample which is expected to produce the most precise estimates with the least possible sample. This is both cost-effective and lowers the response load on the student and parent populations.

The statistically most desirable sample would be a census of the students/parents in the school (as it has no sampling error, although non-sampling errors may be present). If a census is a cost-effective option, then this would enable high quality opinion measures for not only the whole school, but for sub-populations within the school. This would allow for better analysis of the opinion measures, as comparisons between various demographic groups within the student and parent populations could be made free of sampling error.

If a school decides to conduct a sample survey (and not a census), then it will need to consider how it selects the sample of students/parents to complete the questionnaire. There are many ways of selecting a sample of students/parents from a school. Paragraph 3.5 above gives general guidance and Paragraph 4.3 below gives a recommended approach that is simple to understand, robust for estimating proportions, and easy to implement in practice.

3.7. Producing estimates

After the data has been collected, it is necessary to summarise the data in a set of estimates.

The most likely scenario is that the collected data is imperfect in one or both of the following ways:

- complete non-response (a selected student or parent did not complete any of the questionnaire), and/or
- data does not represent the views of the student or parents (for whatever reason).

These forms of imperfect data can lead to problems with estimation. Non-response can lead to the composition of the sampled students/parents being unrepresentative of the school population. Unless steps are taken to correct this, the estimates will be biased. Weighting is often used to help correct this problem. The weights of some responding students/parents are inflated to ensure the sample is able to properly represent the frame population.

The more difficult situation that arises is when non-response cannot directly be corrected by population counts. For example, suppose student non-response status is related to some unobserved characteristic of the student population. In this case, weighting will have difficulty in correcting for the bias due to non-response, unless this unobserved variable is at least partly explained by known demographic information such as age and gender (for which school population counts may be available).

It is likely that the national student items will have a different non-response pattern from the national parent items, within each school. Students will be able to be organised relatively easily to complete the national student items, as this can be done in the classroom. For parents it will be considerably more difficult to respond to the national parent items (as parents will probably need to do so after work hours). This will lead to non-response for parents being higher than for students.

As non-response patterns for students and parents will differ, this will require separate weights to be used for each survey in order to ensure each sample properly represents the population it came from (i.e. the student and parent populations at each school).

Weighting the sample according to the number of units responding to each survey is a common method of estimation. The non-responding units will then be represented in any estimates by the responding sample.

The weights for each responding student or parent will then be of the form:

$$w_i = \text{weight of responding unit i (either student or parent)} = \frac{N_h}{r_h}$$

$$\textit{where}:$$

$$h = h^{th} \text{ stratum within the school}$$

$$r_h = \text{ responding sample of students/parents in stratum h}$$

$$N_h = \text{ population of students/parents in stratum h}$$

'Stratum' in this case refers to the level within the school at which the sample is selected. In the recommended sampling scheme in Paragraph 4.3 below, the stratum is the Year cohort. Thus, the population in stratum h is the total number of students (or their parents) in Year h.

Note that the population of parents and the population of students are in general different. This is because a set of parents may have more than one child at the school. To illustrate, suppose a parent had 3 children at a school; a set of twins in Year 4 and an older child in Year 6. The parents would then contribute to two population counts; once to the set of parents of year 4 students and once to the set of parents in year 6. The twins would contribute a count of 2 to the set of students in Year 4 and the older child would contribute a count of one to the population of students in Year 6.

3.8. Analysing the outputs

The data that is collected contains objective information from which real-world conclusions can be made. Data may be used by schools and in the case of government schools, by state/territory education departments, for a number of purposes.

The first step in using any data collected from a survey is to have some metadata about the data. Metadata includes information about the collected data which summarises the various dimensions of quality of data. Some of the more important measures of quality that a user should be informed about are:

At jurisdiction level only:

- which schools were available for the survey and which were excluded (and why)

At jurisdiction and school level:

- response rate (how many of the selected students/parents responded in each school)
- how the non-response affected estimates
- what sort of methods were used to mediate the non-response
- basic demographics of the non-respondents versus the respondents
- RSE's of the estimates
- size of the sample versus size of the population, and
- a description of the sampling methodology employed.

The most common types of analyses involving survey data are done by tables and graphs. Tables usually involve cross-analysis of two classificatory variables with each other and then producing counts or proportions within each cross-cell. Tables allow analysis of two-way interactions between data items and are visually accessible and easy to interpret. They suffer from the limitation of not being able to easily identify more complex relationships (tables of three or more dimensions are very difficult to understand, except as a series of two-dimensional tables).

For analysing these more complex relationships between the survey variables, statistical tools such as linear regression are generally used. These tools analyse the relationship between the proportion of students or parents who respond a particular way to a survey question (e.g. they respond as “Strongly Disagree”) and the demographic characteristics of these students or parents. This allows a model to be developed which predicts the agreement/disagreement rating level of a student or parent with specified demographic characteristics. These models allow general statements to be made such as “students <of a certain age and gender> are satisfied with <the quality of some aspect of their school’s performance>”.

When using tables to assess differences between the response to survey items between two groups of students or parents, it is necessary to take into account the sampling error in the cells of the table. In order to do this, it is necessary to construct what is known as a 'confidence interval' about each of the estimates. A confidence interval for an estimate is a range about the estimate of which we are reasonably sure the true population value lies within. Generally, 95% confidence intervals are used (i.e. the confidence interval centred about the estimate contains the true population value with 95% certainty).

To determine whether two estimates are equal (or differ), it is necessary to take into account the confidence intervals about each estimate. Without using confidence intervals, incorrect statements can be made about whether the level of satisfaction between two jurisdictions is the same or not. Sampling error can sometimes be quite large and so it can make two estimates appear quite different even though the difference may be simply caused by sampling variability (i.e. variability due to the fact that a particular probability sample was used).

The most common (though conservative) test is to only conclude that two estimates differ if their confidence intervals do not overlap. This is an inexact test and there are more accurate tests available. However, it is a widely-used method of testing for a real-world difference in the presence of sampling error.

3.9. Presenting results

To ensure that informed decisions are made when using the national student and parent opinion item data, it is necessary to provide a user of the estimates with sufficient information regarding how the estimates were created. The information that a user needs to use the data appropriately should cover the various steps of the survey design process, the estimation methodologies used and critically the limitations of the data. Statements on the limitations of the data will include quantifications or statements about the characteristics of the students or parents who responded to the national student and parent opinion items, plus also comments on the quality of the survey frame. It is also common practice to include a copy of the items, so that users can assess what the questions were and how they were presented to respondents.

Analyses of estimates are usually accompanied by some indication of the RSE of the estimates. The way this is done varies, but a recommended method is to annotate unreliable estimates to alert users. For example, a simple rule may see estimates being annotated if their RSE is greater than 10%. Such rule-based annotations are easy to implement and also simple for a user to access.

Quality in estimates can be quantified along two dimensions. The first dimension summarises sampling error and the second dimension summarises non-sampling error.

Sampling error is usually summarised either by standard errors or by the use of confidence intervals. Standard errors are a measure of the amount of error there exists in an estimate due to the fact that a particular random sample was selected. As different random samples give rise to different estimates, it is important to have a summary measure which can specify the amount of error in an estimate. Confidence intervals give an interval about an estimate in which the true (population) value lies. Normally 95% confidence intervals are used (i.e. the interval will contain the true population value in 95% of samples).

Paragraph 3.5 above discussed RSE's of an estimate of proportion. This example can also be used to illustrate the use of confidence intervals. The 95% confidence interval can be approximated by use of the SE of the estimate and is given by:

$$(p - 1.96 * SE(p), p + 1.96 * SE(p))$$

In the case above, this evaluates to:

$$(0.3 - 1.96 * 0.0627, 0.3 - 1.96 * 0.0627) = (0.177, 0.423)$$

This is the 95% confidence interval for the estimated proportion 0.3 (when using a sample of 50 students). That is, we are 95% certain that the true value of the proportion lies between 0.177 and 0.423. This is clearly a very wide interval and probably unsuitable for most statistical purposes. If a sample of 200 students is used, the confidence interval is narrowed to the interval (0.244, 0.356), which is more acceptable.

3.10. Maintaining confidentiality

The National Statistical Service (NSS) website www.nss.gov.au refers to confidentiality as the act of protecting the privacy of a respondent through ensuring that information collected is not revealed to a unauthorised third party. The main manner in which disclosure of confidential data about a respondent can be revealed is through the use of fine-level tables ('direct disclosure'). Direct disclosure can occur, for example, by analysing counts of a school's level of agreement scores by demographic variables which are so fine that only one or two students fall into that cell of the table. If the row and column attributes of the table define publicly identifiable characteristics (such as ethnicity) then it may be possible to derive the responses of these students or parents.

Confidentialising tables can be achieved by one of several means. The first of these is simply not releasing such tables at all ('data suppression'). Data suppression has the flow-on effect of affecting all other tables which contain data provided by the relevant respondents. In particular, higher dimension tables (i.e. tables with three or more classificatory variables) would be affected. Complex algorithms exist which allow for this sort of suppression, but they are labour-intensive.

Another approach is the 'threshold method', which specifies that no data will be reported where there are fewer than a certain number of respondents (which in the case of the national student and parent items is 5). If there are fewer than five responses, that response data would be suppressed. Although other response data from that respondent would be published, some related data might be suppressed so that the count of the original suppressed data cannot be re-derived. This prevents 'inadvertent disclosure'.

Another method that is commonly used is 'data rounding' (or 'random rounding'). In this method, there is no suppression of sensitive data. Instead, data relating to a small number of respondents is perturbed by randomly adding or subtracting a small fixed number. The random rounding is done in such a way that the total count is unaffected. Sometimes more complex patterns of rounding are used, so as to ensure a greater degree of privacy. A pleasing feature of this method is that there is consistency of totals.

A more difficult confidentiality problem occurs when data is analysed in several different ways. Most confidentiality methods are applied one report at a time but do not consider the cumulative effects of several reports. Such situations are complex and require specialist advice to resolve.

Lastly, disclosure from external information is perhaps the most challenging issue in the protection of respondent privacy. A user may have information about a respondent which is obtained from a source independent of the tables that are produced. This information may include the user's personal knowledge of some of the respondents. The use of this sort of additional information is very difficult to protect against by algorithmic means. Administrative or legal means may need to be relied upon to ensure privacy compliance.

4. Recommended sample designs for schools asking the national student and parent school opinion items

The national student and parent school opinion items will be administered locally by each individual school. Hence survey design decisions will need to be made by each school as to how the items will be surveyed in their school.

There are two sets of ministerially approved school opinion items. The first set of items asks students to express their opinions on a number of questions regarding their school experience. The second set of items asks parents generally similar questions.

For both sets of items, schools are able to add additional questions to the items, thereby creating a survey, or they may create a survey which consists solely of the relevant set of items. In addition, jurisdictions may add modules to the national modules and ask schools to include these in their surveys.

A standard set of student/parent demographic items may be added to the relevant survey by a school and/or jurisdiction.

The advice below covers both sets of items. It is a strong preference to conduct a census of both students (Years 5–12) and parents (all Year levels). If a school chooses to undertake a random sample instead, the advice below should be followed when asking both the student and the parent items.

Care must be exercised to ensure that the response to both the student and parent items is as high as possible. As explained elsewhere in this document, this will ensure any non-response bias is kept to a minimum.

4.1 Census

If resources permit, the most effective way of asking the student items is via a census of all students (Years 5–12) within the school. For parent items, the recommended approach is a census of all parents (regardless of Year level). <u>This is the recommended approach for all schools</u>. Note the comments in Section 3 above regarding the potential for higher rates of non-response for parents. When non-response becomes large (say, at levels greater than 25% or so) the issue of non-response bias becomes problematic. It may thus be preferable to conduct a random survey of parents as described in paragraph 4.3 but devote some resources to following-up non-responding parents to achieve high response.

A census will eliminate sampling error (assuming full response). However, non-sampling errors will still exist. For example, it will be necessary to closely monitor, and report on, student non-response to the student items. This could include demographic information, such as non-response by age and sex of student.

A census will also allow for detailed measures of opinion within a school. It will allow for accurate measures of opinion between male and female students and also for comparisons by other information that is available for all students at the school.

4.2 Sampling

If a census of all students at a school is considered inappropriate, then a sample of students within the school will need to be carried out.

In order to carry out a sample survey of students at a school, it is necessary to consider two main parameters. The first is cost and the second is quality of estimates. These two considerations taken together will lead to the specification of a total sample size for a school that is affordable and that produce opinion measures with acceptable quality.

The **spreadsheet** that accompanies this document illustrates the association between the student sample size and the sampling error of an estimate of proportion. These sample sizes relate to a sample design where the students are selected randomly from the entire school. This sample selection method can be improved by ensuring that the students are stratified by Year and gender. This will ensure that there are students across the full range of Year groups and genders, rather than leaving this representation to chance. It will also allow for opinion measures to be produced more reliably for sub-aggregates of a school, such as by Year-cohort.

This spreadsheet will assist in choosing the appropriate sample size for a school (see Paragraph 4.3 below).

4.3 Selecting a sample

The following steps should be taken in selecting a sample of students from a school to complete the national student items:

1. Obtain a list of all the students in the school (Years 5–12).
2. Stratify the list by Year-cohort.
3. Count the number of students in each Year-cohort.
4. Decide on the total student sample size from the school, using the **spreadsheet.**
5. Allocate this sample size amongst the strata proportional to the number of students in each Year (obtained from step 2). This will result in a sample size for each Year-cohort. Round the sample size up within each year if it is fractional.
6. Randomly select the sample of students from each year-cohort. These students will receive the national student opinion items.

In selecting a sample of parents (if desired), a similar process can be followed (although bearing in mind that parents of students at all Year levels can be included in the sample).

As one parent response to the national parent items is to be completed on behalf of each selected child, it is left to the parents to agree on the survey responses. In cases where there is a complex relationship between the parents (e.g. divorced or separated parents), the parent with the best knowledge of the child's school and its conditions should answer the items. This will be taken to best represent the opinions of the joint parents.

4.4 Example sample design

It is highly recommended that all schools with fewer than 100 students undertake a census of their students and parents. Therefore the methodology outlined below is only for application in schools with more than 100 students falling within the target population (ie from Years 5–12). The example given relates to designing and selecting a student sample; similar principles apply for selecting a parent sample (bearing in mind that parents with children at any Year level may be sampled).

Assume that Happy Valley High School has 340 students in Years 7–12. Also, it has been decided that the school will undertake a sample survey of students and not a census due to cost issues.

After working out what sort of sample size is affordable with the resources available at the school, let's say that the school decides to sample 100 students from the school. The spreadsheet then gives the sorts of precision the estimates from the survey will have. The closest table in the spreadsheet to 340 is for schools of size 300; hence this tab in the spreadsheet is used.

Using this tab of the spreadsheet, estimates of proportions of around 0.5 will result in a confidence interval of (0.42, 0.58). That is to say, there is a 95% chance that the true value of the proportion will lie between (0.42, 0.58) if the estimated proportion is 0.5. The school will need to decide if this is sufficiently precise for their needs.

The next step is to obtain a list of students and to order them (i.e. stratify them) by Year-cohort (from Year 7 through to Year 12). The purpose of this will be to determine how many students should be selected into the sample from each year cohort. After counting the students by year, the following student counts were calculated:

Table: Happy Valley High School (sample of 100 students)

Year	Students	Proportion of students in each Year	Sample of students in each Year	Final (rounded) student sample size
7	50	50/340 = 0.147	100*0.147 = 14.7	15
8	55	55/340 = 0.162	100*0.162 = 16.2	16
9	45	45/340 = 0.132	100*0.132 = 13.2	13
10	50	50/340 = 0.147	100*0.147 = 14.7	15
11	60	65/340 = 0.191	100*0.176 = 19.1	19
12	80	75/340 = 0.221	100*0.235 = 22.1	22
Total	340			100

4.5 Selecting a random sample

The next step is to select the required sample of students from each year. The discussion in this document has shown that there are compelling reasons for selecting a random sample of students and hence this is very strongly recommended.

Selecting a random sample using *School Survey*

School Survey provides a randomisation tool which schools may use to generate a random sample from a cohort uploaded into the system. Schools can use this tool by selecting the option “Do you want to send this survey to a random set of recipients?” If this option is selected, the user will be prompted to indicate the percentage of users to be surveyed.

Student samples can be stratified by uploading recipients by year level and selecting a percentage for each year level.

Selecting a random sample using a spreadsheet

The selection of a random sample can be facilitated by the use of commonly available spreadsheets, such as Excel. The easiest way to select the random sample of students in the case of the Happy Valley High School example above is to put the stratified list of students into Excel. The list will contain the identifier (e.g. name) of all 340 students from Happy Valley High School, sorted by Year. Using the RAND() function in Excel, generate a random number between 0 and 1 for each student. Then sort the list of students within each Year cohort by this random number.

To select the required sample of 15 students in Year 7 in Happy Valley High School, it is a simple matter of then selecting the first 15 students as ordered by the random number. Repeat this random-sorting and selection within each Year cohort to obtain the sample of students within the school.

Selecting a random sample if a spreadsheet is not available

A standard method to use to select a sample when a spreadsheet is not available is to generate one random number (using random number tables) and then apply a skip within the Year to select the required students. For example, suppose that a Year within a school has 100 students in it and a sample of 20 students is required. Thus a skip of 100/20=5 will be required to apply through the list of students.

The next step is to use random number tables to randomly select the starting point (i.e. first student to select), which will lie between 1 and the skip (in this case, 5); say that random number is 3. Every 5th student after the 3rd student will be selected. That is the 3rd, 8th, 13th, 18th etc student is selected until a sample of 20 students is selected.

The national parent items will then be administered to the parents of the selected students. In the case where a family has multiple students at a school, it is important that the parent of the selected student(s) respond only on behalf of the selected student(s). If the parent(s) have three children at the selected school and two of these students are selected into sample, then the parent(s) should respond ONLY on behalf of these two selected children. In this case, there will be two sets of parent responses (as there are student-specific items which require a separate response to be completed by parents on behalf of each selected child).