DATA INTERPRETATION GUIDANCE DOCUMENT

RHODE ISLAND STUDENT SURVEY

ACADEMIC YEARS:
2012 – 2013
2013 – 2014
SELECTED RESULTS/ COMPARISONS

COMMUNITY RESEARCH AND SERVICES TEAM
UNIVERSITY OF RHODE ISLAND
80 WASHINGTON STREET,
PROVIDENCE, RHODE ISLAND

GUIDELINES FOR INTERPRETATION

HOW TO BEST UTILIZE INFORMATION CONTAINED IN STUDENT SURVEY OUTCOME REPORTS

INTERPRETING THE DATA

The purpose of this document is to provide a brief overview of best practices and guidelines for interpretation of results from the 2012 – 2013, and 2013 – 2014 administration of the Rhode Island Student Survey (RISS), as well as to set new standards for future feedback reports of other survey tools. This document is intended to help consumers of survey data understand: 1) the relationship between parental consent procedures and response rates among students completing school surveys; 2) whether response rates by grade level administration are sufficient for providing meaningful and useful feedback to key stakeholders; 3) how weighting procedures can be helpful in promoting greater confidence in estimates generated from survey results; and 4) how best to interpret differences in proportions using confidence intervals.

PARENTAL CONSENT: WHO, WHAT, WHERE, WHEN, AND WHY?

Active and **passive consent** procedures for administration of survey data within the school setting refer to two distinct methods for obtaining consent from parents or guardians. Typically, **passive consent** procedures involve sending a letter that explains the basic premise of the survey to all parents or guardians of students enrolled at a school, with specific steps for retracting permission should parents or guardians decide they do not want their children to participate.

In contrast, **active consent** procedures require that parents or guardians signify in writing that they permit their children to participate in the survey. The key difference between these methods is that **passive consent** usually assumes that parents or guardians have consented *unless some other action is taken to indicate otherwise*, whereas **active consent** procedures do not make this assumption and require parents or guardians to explicitly specify their consent in writing.

Consideration of consent procedures is important prior to survey implementation because of marked differences in response rates between these methods. One benefit of **passive consent** is that this method typically results in very high response rates, and may yield a more un-biased and representative sample of the student body within a school. Knowledge of a school's parental consent procedures during the planning phase of survey administration can be useful in determining which method is most likely to yield maximal response rates, as well as to standardize data collection procedures across multiple school settings, where applicable.

DIVING FOR DATA: DANGERS OF JUMPING HEADFIRST WITHOUT KNOWING THE DEPTH OF THE DATA POOL

Many sources recommend obtaining a minimum survey **response rate** of 60% in order to minimize bias. Here, **response rate** refers to the number of individuals who responded to or answered a survey divided by the number of people in the sample, and is often expressed in the form of a percentage. Response rates are an important component of survey research largely because as the **response rate** increases, the risk for bias decreases.

There are several types of bias that are important to minimize in order to obtain the best possible survey results. One type of bias or error that can be particularly problematic is **nonresponse bias**, which occurs when non-respondents differ in significant ways from respondents on key items of interest. The risk of **nonresponse bias** can be elevated when the topic under investigation is sensitive and respondents would prefer not to disclose information on these matters.

Undercoverage can also be problematic if some members of a population are inadequately represented in the sample. For instance, if only a small percentage of high-school seniors within a school respond to a survey administered across all grade levels due to lack of resources to attend their senior class trip, the results obtained from the survey may not be representative of those students who were absent, and results may be skewed.

Given that many of the items on the RISS and other similar assessments of adolescent and young adult functioning refer to sensitive topics such as drug and alcohol abuse, as well as mental health concerns and school satisfaction, it is also possible that students are not entirely honest when completing these measures. Many school survey measures contain one or more items assessing student honesty when completing the survey. Surveys from students who indicate that they were not at all honest when completing the measure will hence be discarded.

Given these considerations, *only data from surveys with a minimum response rate per grade of 60% will be considered for evaluation and analysis*. The purpose of these efforts is to enforce best practices in data management and encourage survey administrators to increase collection of more representative data from their constituents. Full data reports will be provided to communities that meet this standard.

It is the strong belief that adherence to a minimal **response rate** of 60% per grade level will provide a clearer representation and understanding of key variables under investigation, facilitate data-driven decision-making practices, and ultimately influence the implementation of targeted interventions that maximize efficiency and effectiveness.

GOOD THINGS COME TO THOSE WHO WEIGHT: THE IMPORTANCE OF USING STATISTICAL WEIGHTING PROCEDURES FOR SURVEY DATA

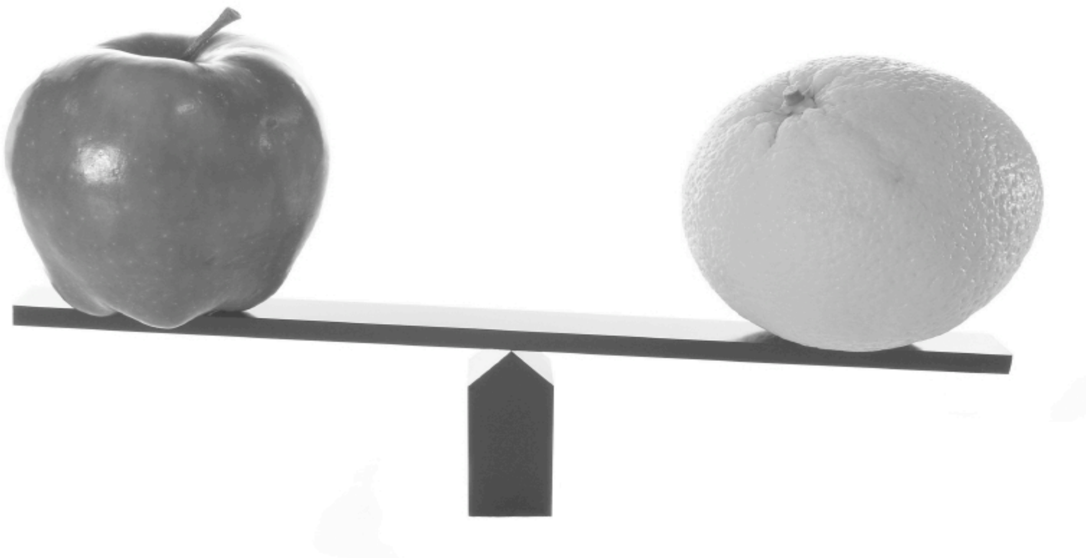
Weighting procedures for survey data can be applied when sample proportions do not match those found in the population, and accuracy of the data is of utmost importance (always!). A **data weigh**t is a multiplier that makes a given respondent's contribution larger or smaller to compensate for a variety of planned and/or unexpected disproportionate effects. Typically, numerous dimensions are taken into consideration when constructing weights. For the purpose of the RISS and other similar survey measures, weighting responses according to grade level response and population rates may one method for ensuring adequate sampling representation.

In some sense, **weights** can be thought of as the "corrective" values assigned to each one of the sample responses in a survey. Data that are un-weighted assume that each individual provides an equal contribution, whereas **weights** can be assigned to correct for under or over-representation of population groups. **Weights** are often fractions, are always positive, and non-zero. Individuals from under-represented groups get a **weight** larger than 1, whereas those from over-represented groups get a **weight** smaller than 1. This can be understood using the simple example where a **weight** of 2 means that the case counts as two identical cases in the dataset.

One argument for the importance of **weighting** can be found in the following fictional example. Suppose two people are asked to bring ten pieces of fruit to a picnic. The first person stops at a local fruit stand and buys 8 oranges for $1.00 each and two apples for $.50 each, whereas the second person stops at a large supermarket and buys 2 oranges for $1.75 each and 8 apples for $.45 each. From one perspective, the first person's purchase was more expensive since they clearly spent more money at the fruit stand ($9.00) than the second person spent at the supermarket ($7.10).

However, from another perspective, the second purchase is more expensive since the supermarket charged a much higher price for the oranges and only slightly less for the apples. In this case, the shopper may be interested in whether fruit prices are higher at the fruit stand or the grocery store, and by how much. This might be particularly relevant if the shopper were also interested in purchasing many other types of fruit.

One issue with using the total cost ($9.00 versus $7.10) or average ($.90/piece at the farm stand versus $.71/piece at the supermarket) is that the fruit stand average price gives more weight to the price of oranges because the person purchased more of them, whereas the supermarket price gives more weight to the price of apples. In this sense, we are now comparing the price of apples to oranges, rather than comparing the fruit stand to the supermarket. A solution to this dilemma might be to use a **weighting** procedure that averages the prices in the same way for each vendor, such that both averages give the same proportionate weighting to oranges. In this sense, **weighting** methods may be helpful in providing a more accurate representation of the data, particularly if more complicated methodologies are anticipated.



INCREASING CONFIDENCE USING CONFIDENCE INTERVALS: INTERPRETING DIFFERENCES IN PERCENTAGES VISUALLY AND STATISTICALLY

Although visual inspection of the data may reveal what appear to be changes in percentages over time, it is important to keep in mind that unless otherwise specified, these changes are not statistically significant. As such, any significant differences at the .05 level are noted by a double asterisk, '**'.

Determination of statistical significance: Statistical significance can be determined by calculating the **confidence interval** for the difference between percentages or proportions. A **confidence interval** is a range of values that describes the uncertainty surrounding an estimate, and is indicated by its endpoints. **Confidence intervals** are one way of representing how "good" an estimate is; the larger the range of values contained in the interval, the more caution is require when using the estimate. Hence **confidence intervals** are an important reminder of the limitations of estimates.

For instance, a 95% **confidence interval** for a sample proportion can be interpreted using the following guidelines. If we were to take many samples of the same size and create a confidence interval using each sample statistic, over time 95% of our confidence intervals would contain the true population parameter. Increasing sample size often decreases the margin of error; estimates generated from small samples often have wide confidence intervals that limit the ability to make broad inferences about the data.

For instance, a 95% confidence interval for a parameter of interest (i.e. mean) suggests that if we were to take 100 samples of the same size and calculate the parameter for each sample, in 95 out of 100 samples the estimated parameter will fall within this stated range. Therefore, the true parameter has a 95% chance of falling within this range. Conversely, there is a 5% chance that the true parameter is not within this interval.

Following this logic, when the **confidence interval** for the difference between proportions contains the value of zero, we cannot assert that there is a meaningful difference between these values, as there is the distinct possibility that the true difference may be zero.

Statistical significance of the difference between two proportions can also be visually assessed by plotting the confidence interval for each proportion, and examining these intervals for overlapping values. Proportions are deemed significantly different if their confidence intervals do *not* overlap. However, the converse is not *absolutely* true under all conditions, such that it is possible that proportions with overlapping intervals *are*, in actuality, distinct. Hence, statistical tests of significance are presented in all reports of survey data, using the methods described above.

Figure 1: Visual inspection of confidence intervals for differences in proportions with no difference in proportions. (Note that confidence intervals for each proportion overlap).

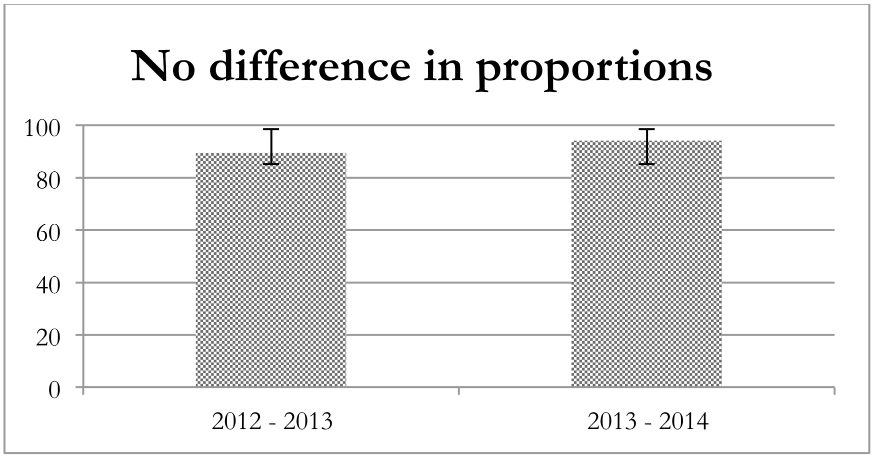


Figure 2: Visual inspection of confidence intervals for difference in proportions. Note that the confidence intervals for each proportion do not overlap, hence these values are significantly different.

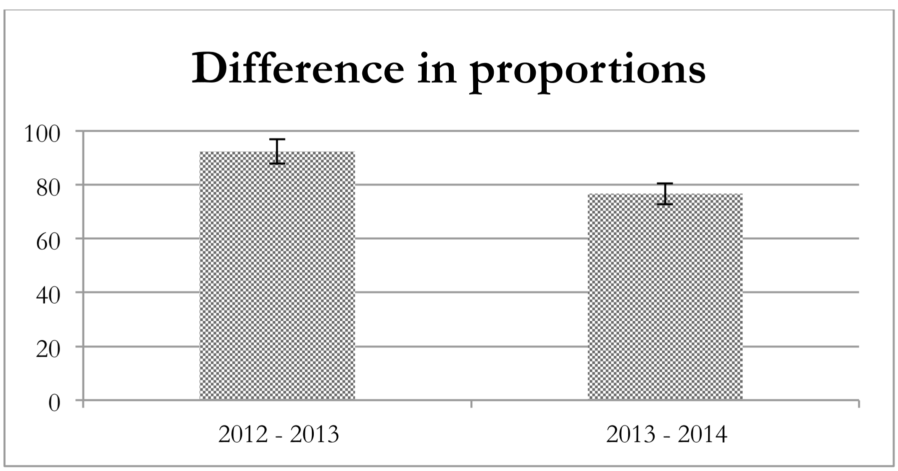
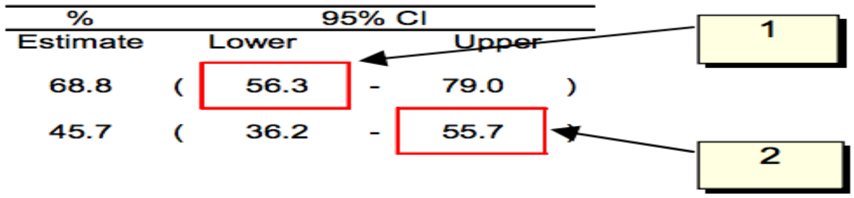
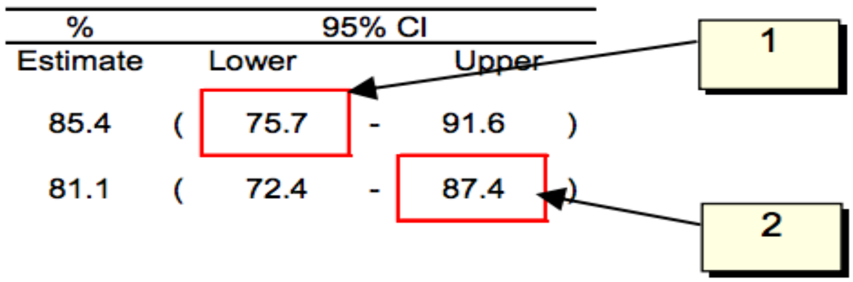


Figure 3: Visual inspection of confidence intervals. (Images derived from: https://www.health.wa.gov.au/publications/documents/Confidence_intervals_How_they_work.pdf)



In the figure above, the upper bound of the smaller estimate is less than the lower bound of the higher estimate. Hence, the intervals do not overlap and the estimates are significantly different.

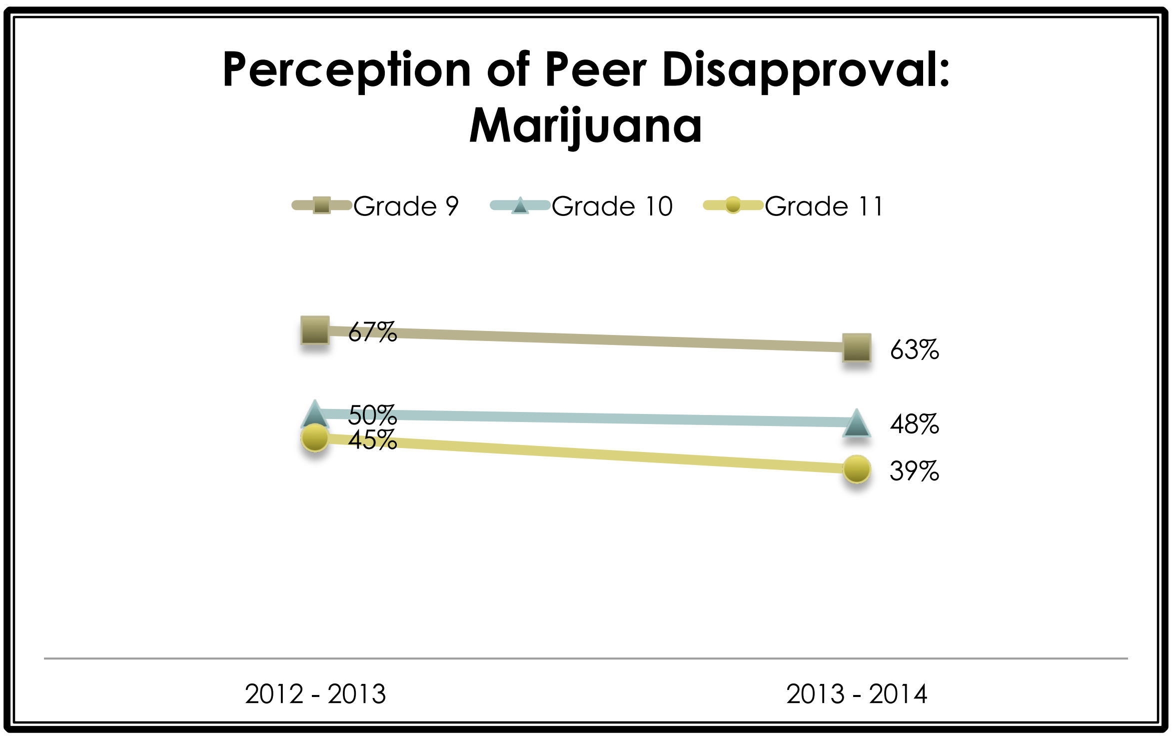


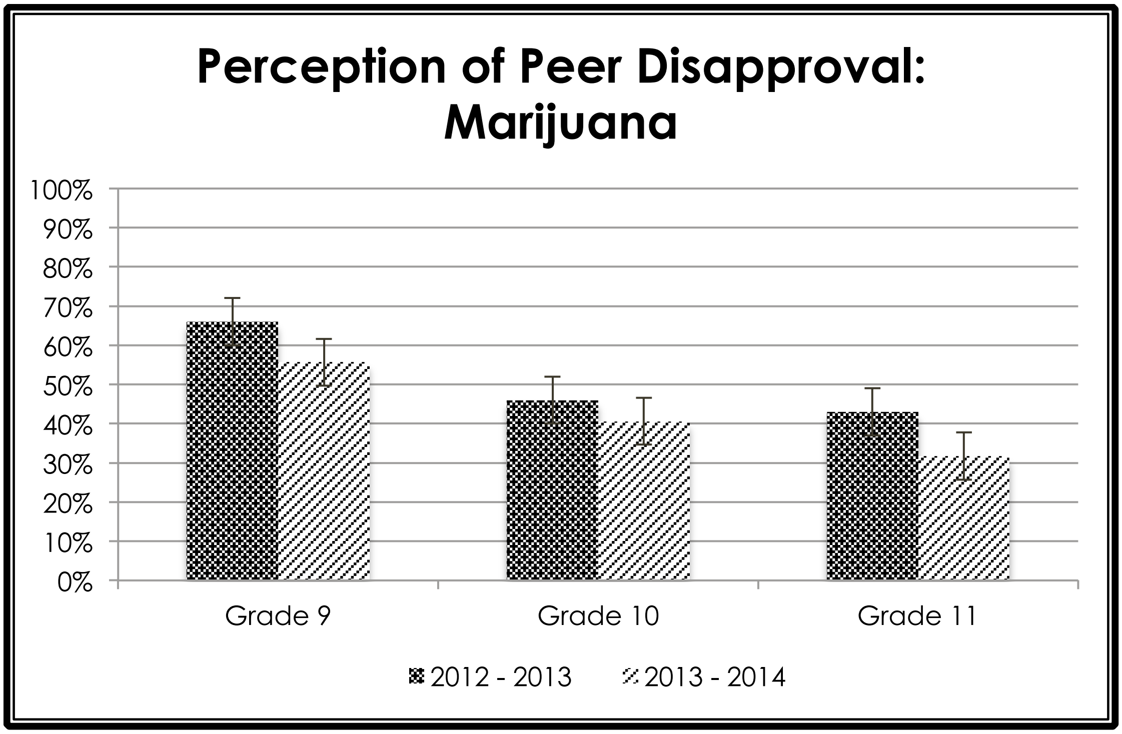
In the figure above, the upper bound of the smaller estimate is larger than the lower bound of the smaller estimate. Hence, the intervals overlap and the estimates may not be significantly different. Formal statistical testing is required to determine significance at the .05 level.

Peer Disapproval: Marijuana

Students were asked to report on their perception of peer disapproval of marijuana use. Responses indicate percentage of youth who endorsed the belief that peers would disapprove or strongly disapprove of marijuana use.

Peer Disapproval: Marijuana	*2012-2013*	*2013-2014*	*Difference*	*95% CI*
Grade 9	67%	63.4%	-3.61%	-.0609, .1298
Grade 10	50%	48.2%	-1.8%	-.0725, .1129
Grade 11	45%	38.6%	-6.4%	-.0275, .1571
Grade 12	NA	NA	NA	NA
Total	NA	NA	NA	NA
*****Significant at .05 level.***				





This marks the end of the sample report. Communities may expect to find information presented for each question of interest in the format outlined above.