### **How Benchmarks are Calculated**

SENSE creates two types of benchmark scores based on member colleges' data: raw and standardized. Both types of benchmarks are useful, but for different purposes. The standardized benchmark scores SENSE provides are useful for comparing any given college to the cohort at one point in time. Standardized benchmark scores can also be used to compare subgroups within each college: a subgroup (e.g., full-time students) with a standardized benchmark of 52 will be more engaged on that benchmark than a subgroup (e.g., part-time students) with a standardized score of 47.

For colleges wishing to conduct longitudinal analyses of trends at their campus, the raw benchmark scores are the appropriate measures to use. The standardized benchmark scores are not appropriate for longitudinal analysis as they are recalculated every year and are based on the distribution of responses for each annual 3-year cohort. The raw benchmark scores, on the other hand, are always on a 0-1 scale and are not affected by fluctuations in the distribution of national responses from year to year.

Creating both types of benchmark scores involves several steps, including reverse coding items as necessary and converting all responses to the same scale. After these initial steps are taken, raw benchmark scores are computed for each respondent. Based on the raw benchmark scores, *SENSE* then computes standardized benchmark scores for each respondent. Please note that individual colleges cannot compute standardized benchmark scores as this process can only be completed using the complete 3-year cohort data set. However, pre-calculated standardized benchmark and raw benchmark scores for each student record are included in the college raw data file. Once the benchmark scores are created (either raw or standardized), college, campus, and group-level benchmarks can be calculated. The steps used to create these benchmark scores are explained in detail below.

## **Creating Benchmark Scores**

#### 1. Reverse coded items.

The first step is to determine which items, if any, need to be reverse coded so that a low score on the item represents a desirable behavior. For example, the item "Skip Class" is originally coded such that 1=never and 4=four or more times. In this case, we "Never" should have a higher positive impact on the benchmark score than skipping class four or more times. The easiest way to reverse code this item is to use the following formula that assumes, as in the case of Skip Class, the item has four response options:

ReverseScore = 5 - OriginalScore.

For Skip Class the reverse codes would be: SKIPCL Rev = 5 - SKIPCL.



```
4 = 5-1 (4 becomes the value for "Never")
3 = 5-2 (3 becomes the value for "Once")
2 = 5-3 (2 becomes the value for "Two or three times")
1 = 5-4 (1 becomes the value for "For or more times")
```

Several Yes/No response items are used in calculating benchmarks. These are normally coded as 1 = yes and 2 = no. The same logic presented above is used to reverse code these scores so that no = 0 and yes = 1 (ReverseScore = 2 - 0).

(NOTE: Eight SENSE items – 12a, 12b, 14, 19c, 19d, 19f, 19s, and 23 – are reverse coded.)

### 2. Convert all items to a common scale with a range of 0 - 1.

After reverse coding the items, the next step is to convert all items to a common 0 (zero) to 1 scale. The following formula is used to accomplish this conversion:

```
RescaleScore = (OriginalScore - 1) / (max_response_value - 1).
```

Using the Skip Class example again, where the original variable name is SKIPCL, this formula would be:

```
SKIPCL_RevRaw = (SKIPCL_Rev - 1) / 3

0.00 = 1 - 1 / 3

0.33 = 2 - 1 / 3

0.66 = 3 - 1 / 3

1.00 = 4 - 1 / 3
```

(NOTE: Remember when working with the reverse-coded items to use the reverse-coded variable in this step.)

#### 3. Create raw benchmark scores.

Calculation of the raw benchmark scores use the newly-created rescaled (0-1) variables. These scores are created by calculating the average score of the items that make up the benchmark. Using the Early Connections benchmark (EARLYCON) as an EXAMPLE and the item numbers from the survey as variable names, the formula for computing the raw benchmark score is:

EARLYCON\_Raw = (18a(welcome) + 18i(fainfo) + 18j(qualfa) + 18p(cstafnam) + 23(asnpers) / 5.

#### 4. Compute standardized benchmark scores.

Before explaining how this step is carried out, it is important to note that standardized benchmark scores cannot be computed without having the entire cohort data set (all respondents included in the 3-year cohort). As such, this step is only briefly explained.



SENSE uses the STANDARD procedure in SAS to create the standardized benchmark scores across the 3-year cohort so the average benchmark is 50 with a standard deviation of 25 at the student record level. To account for the inherent sampling bias, this calculation includes weights, the utilization of which is explained in the next step.

### 5. Calculate group-level benchmark scores.

The process explained above creates benchmark scores (raw and standardized) for every respondent in the primary sample. (Standardized benchmark scores are not created for over-sample respondents.) The process for creating group-level benchmark scores is the same for both raw and standardized benchmarks. In most circumstances, the group-level benchmarks are created by calculating the *weighted* average of a benchmark variable for the members of the group (e.g., males and females). Note that the word "weighted" in the previous sentence is emphasized. Sampling for *SENSE* is done at the class level and, as such, full-time students are more likely to be included in the sample than part-time students because full-time students take more classes. To account for this sampling bias, *most* analyses, including the computation of group-level benchmark scores, must incorporate weights so the results are more representative of the actual distribution of students at a given college. As noted, most, but not all, analyses require the use of weights. The exception to this is when groups are formed based on enrollment status (full-time vs. part-time). Any time a group consists of only part-time or only full-time students, weighting should not be used.

## Computing the Six SENSE Benchmark Scores

Raw data sets returned to participating colleges include all core *SENSE* survey items as well as special-focus item data, if applicable. Two sets of benchmark scores are included in each college's data set: raw benchmark scores and corresponding standardized benchmark scores for each respondent. These scores can be used to calculate sub-group benchmark scores.

The process SENSE uses to calculate raw benchmark scores and how to create subgroup benchmarks using the raw and standardized benchmark scores are described in this section. (See "When to Use Standardized and Raw Benchmark Scores" for a brief discussion of when it is appropriate to use each of these types of benchmark scores.) Please note that individual colleges cannot compute individual respondent-level standardized benchmark scores as this process can only be completed with the complete 3-year cohort data set.

The standard process for calculating individual respondent-level benchmark scores involves:

- 1. Reverse coding items (where applicable)
- 2. Converting scores on benchmark items to a common scale with a range of 0-1 (zero to one).
- Computing the benchmark score



- 4. Computing group-level benchmark scores
  - a. Raw benchmark scores
  - b. Standardized benchmark scores

## **Early Connections (items 18a, 18i, 18j, 18p, and 23)**

Early Connections includes one item that requires reverse coding: ASNPERS. This is a yes/no response item so the reverse coding is:

```
Q23 (2-point scale): asnpers Rev = (2 - asnpers)
```

Early Connections does not include any items that require reverse coding, so the first step above is not applicable.

The process for converting the original scale for each item to a 0–1 scale is the same varying only by the number of response options for any given item. The math for converting each item is presented below. (Note: the lowest numeric response value for all items in the original scale is 1 and the highest value is 5 for all items except question 23, a yes/no response item, which is 2.)

```
Q18a (5-point scale): welcome_Raw = (welcome - 1) / 4
Q18i (5-point scale): fainfo_Raw = (fainfo - 1) / 4
Q18j (5-point scale): qualfa_Raw = (qualfa - 1) / 4
Q18p (5-point scale): cstafnam_Raw = (cstafnam - 1) / 4
Q23 (2-point scale): asnpers_Raw = (asnpers - 1)
Q23 (2-point scale): asnpers_RevRaw = (asnpers_Reva)
```

NOTE <sup>a</sup>: This variable is already on a 0-1 scale so asnpers\_RevRaw and asnpers\_Rev have the same value.

The new rescaled variables can now be used to calculate the raw individual-level benchmark scores. This is simply a matter of computing the average of the five rescaled items::

```
EARLYCON = (welcome_Raw + fainfo_Raw + qualfa_Raw + cstafnam_Raw + asnpers RevRaw) / 5
```

The final step is creating the raw benchmark score for a given population subgroup. This is accomplished by computing the weighted average of the raw benchmark score (EARLYCON) for all respondents in the subgroups of interest. (IMPORTANT NOTE: If your population subgroup is based on enrollment status (Full-time vs Part-time), then the weight should not be used; instead the



unweighted average should be calculated. See "When To Use Weights" for a more detailed discussion of using weights in analyzing SENSE data.)

The raw benchmark variable for the Early Connections benchmark is EARLYCON and the standardized benchmark variable is EARLYCON\_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting EARLYCON\_STD for EARLYCON.

# High Expectations and Aspirations (items 18b, 18t, 18u, 19c, 19d, 19f, and 19s)

High Expectations and Aspirations contains four items that require reverse coding. These items are reverse coded using the following process:

```
Q19c (4-point scale): latetrn_Rev = (5 - lateturn)
Q19d (4-point scale): notturn_Rev = (5 - notturn)
Q19f (4-point scale): notcompl_Rev = (5 - notcompl)
Q19s (4-point scale): skipcl_Rev = (5 - skipcl)
```

The process for converting the original scale for each item to a 0–1 scale is the same varying only by the number of response options for any given item. The math for converting each item is presented below. (Note: the lowest numeric response value for all items in the original scale is 1 and the highest value is 5 for Q18 items and 4 for Q19 items.)

```
Q18b (5-point scale): wntsccd_Raw = (wntsccd - 1) / 4
Q18t (5-point scale): ittakes_Raw = (ittakes - 1) / 4
Q18u (5-point scale): acprprd_Raw = (acprprd - 1) / 4
Q19c (4-point scale): latetrn_RevRaw = (latetrn_Rev - 1) / 3
Q19d (4-point scale): notturn_RevRaw = (notturn_Rev - 1) / 3
Q19f (4-point scale): notcompl_RevRaw = (notcomp_Rev - 1) / 3
Q19s (4-point scale): skipcl_RevRaw = (skipcl_Rev - 1) / 3
```

The new recoded variables can now be used to calculate the raw individual-level benchmark scores. This is simply a matter of computing the average of the seven recoded items in this scale:

```
ENGAGLRN=(wntsccd_Raw + ittakes_Raw + acprprd_Raw + latetrn_ RevRaw+ notturn_ RevRaw + notcompl_ RevRaw + skipcl_ RevRaw) / 7
```

The final step is creating the raw benchmark score for a given population subgroup. This is accomplished by computing the weighted average of the raw benchmark score (HIEXPECT) for all respondents in the subgroups of interest. (IMPORTANT NOTE: If your population subgroup is based on enrollment



status (Full-time vs Part-time), then the weight should not be used; instead the unweighted average should be calculated. See "When To Use Weights" for a more detailed discussion of using weights in analyzing SENSE data.)

Your college data set will include both raw benchmark scores for each respondent as well as the standardized benchmark scores for each respondent. The raw benchmark variable for the High Expectations and Aspirations benchmark is HIEXPECT, and the standardized benchmark variable is HIEXPECT\_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting HIEXPECT\_STD for HIEXPECT.

### Clear Academic Plan and Pathway (items 18d, 18e, 18f, 18g, and 18h)

Clear Academic Plan and Pathway does not include any items that require reverse coding, so the first step above is not applicable.

The process for converting the original scale for each item to a 0–1 scale is the same varying only by the number of response options for any given item. The math for converting each item is presented below. (Note: the lowest numeric response value for all items in the original scale is 1 and the highest value is 5 for all items.)

```
Q18d (5-point scale): aacontim_Raw = (aacontim - 1) / 4
Q18e (5-point scale): aaselmaj_Raw = (aaselmaj - 1) / 4
Q18f (5-point scale): acadgoal_Raw = (acadgoal - 1) / 4
Q18g (5-point scale): acadgoal_Raw = (acadgoal - 1) / 4
Q18h (5-point scale): acadgoal_Raw = (acadgoal - 1) / 4
Q18h (5-point scale): acadgoal_Raw = (acadgoal - 1) / 4
```

The new recoded variables can now be used to calculate the raw individual-level benchmark scores. This is simply a matter of computing the average of the five recoded items in this scale:

```
ACADPLAN = (aacontim_Raw + aaselmaj_Raw + acadgoal_Raw + crsadv_Raw + oscomm_Raw) / 5
```

The final step is creating the raw benchmark score for a given population subgroup. This is accomplished by computing the weighted average of the raw benchmark score (ACADPLAN) for all respondents in the subgroups of interest. (IMPORTANT NOTE: If your population subgroup is based on enrollment status – Full-time vs Part-time – then the weight should not be used; instead the unweighted average should be calculated. See "When To Use Weights" for a more detailed discussion of using weights in analyzing SENSE data.)



Your college data set will include both raw benchmark scores for each respondent as well as the standardized benchmark scores for each respondent. The raw benchmark variable for the Clear Academic Plan and Pathway benchmark is ACADPLAN and the standardized benchmark variable is ACADPLAN\_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting ACADPLAN\_STD for ACADPLAN.

# Effective Track to College Readiness (items 12a, 12b, 14, 21a, 21b, and 21c)

Effective Track to College Readiness includes three items that require reverse coding. These are yes/no response items and are reverse coded using the following process:

```
Q12a (2-point scale): reqptest_Rev = (2 - reqptest)
Q12b (2-point scale): tkptest_Rev = (2 - tkptest)
Q14 (2-point scale): reqclass_Rev = (2 - reqclass)
```

The process for converting the original scale for each item to a 0–1 scale is the same varying only by the number of response options for any given item. The math for converting each item is presented below. (Note: the lowest numeric response value for items from question 21 is 1 and the highest value is 5; The reverse-coded scale for items from questions 12 and 14 is 0 to 1.)

```
Q12a (2-point scale): reqptest_RevRaw = (reqptest_Rev<sup>a</sup>)
Q12b (2-point scale): tkptest_RevRaw = (tkptest_Rev<sup>a</sup>)
Q14 (2-point scale): reqclass_RevRaw = (reqclass_Rev<sup>a</sup>)
Q12a (2-point scale): reqptest_Raw = (reqptest - 1)
Q12b (2-point scale): tkptest_Raw = (tkptest - 1)
Q14 (2-point scale): reqclass_Raw = (reqclass - 1)
Q21a (5-point scale): Indstudy_Raw = (Indstudy - 1) / 4
Q21c (5-point scale): Indsklls_Raw = (Indsklls - 1) / 4
```

NOTE <sup>a</sup>: This variable is already on a 0-1 scale so asnpers\_RevRaw and asnpers\_Rev have the same value.

The new recoded variables can now be used to calculate the raw individual-level benchmark scores. This is simply a matter of computing the average of the six recoded items in this scale:

```
COLLREAD = (reqptest_RevRaw + tkptest_RevRaw + reqclass_RevRaw + Indstudy Raw + Indacawk Raw + Indsklls Raw) / 6
```



The final step is creating the raw benchmark score for a given population subgroup. This is accomplished by computing the weighted average of the raw benchmark score (COLLREAD) for all respondents in the subgroups of interest. (IMPORTANT NOTE: If your population subgroup is based on enrollment status (Full-time vs Part-time), then the weight should not be used; instead the unweighted average should be calculated. See "When To Use Weights" for a more detailed discussion of using weights in analyzing SENSE data.)

Your college data set will include both raw benchmark scores for each respondent as well as the standardized benchmark scores for each respondent. The raw benchmark variable for the Clear Academic Plan and Pathway benchmark is COLLREAD and the standardized benchmark variable is COLLREAD\_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting COLLREAD\_STD for COLLREAD.

# Engaged Learning (items 19a, 19b, 19e, 19g, 19h, 19l, 19j, 19k, 19L, 19m, 19n, 19o, 19q, 20d2, 20f2, and 20h2)

Engaged Learning does not include any items that require reverse coding, so the first step above is not applicable.

The process for converting the original scale for each item to a 0–1 scale is the same varying only by the number of response options for any given item. The math for converting each item is presented below. (Note: the lowest numeric response value for all items in the original scale is 1 and the highest value is 4 for all items.)

```
Q19a (4-point scale):
                       askques Raw = (askques - 1) / 3
Q19b (4-point scale):
                       prepdrft Raw = (prepdrft - 1) / 3
Q19e (4-point scale):
                       supinstr Raw = (supinstr - 1)/3
                       pinclass Raw = (pinclass - 1) / 3
Q19g (4-point scale):
Q19h (4-point scale):
                       prepoutc Raw = (prepoutc - 1) / 3
Q19i (4-point scale):
                       qrpstudy Raw = (qrpstudy - 1) / 3
                        nrgstudy Raw = (nrgstudy - 1) / 3
Q19j
       (4-point scale):
Q19k (4-point scale):
                        useintmg Raw = (useintmg -1) / 3
Q19I (4-point scale):
                        mailfac Raw = (\text{mailfac} - 1) / 3
Q19m (4-point scale):
                       facassn Raw = (facassn - 1) / 3
                       classrel Raw = (classrel - 1) / 3
Q19n (4-point scale):
Q19o (4-point scale):
                       feedback Raw = (feedback - 1) / 3
                       facidoc Raw = (facidoc - 1) / 3
Q19q (4-point scale):
Q20d2 (4-point scale): fftuse Raw = (fftuse - 1) / 3
Q20f2 (4-point scale):
                       sklabuse Raw = (sklabuse - 1) / 3
Q20h2 (4-point scale): combuse Raw = (combuse - 1) / 3
```



The new recoded variables can now be used to calculate the raw individual-level benchmark scores. This is simply a matter of computing the average of the sixteen recoded items in this scale:

```
ENGAGLRN= (askques_Raw + prepdrft_Raw + supinstr_Raw + pinclass_Raw + prepoutc_Raw + grpstudy_Raw + nrgstudy_Raw + useintmg_Raw + mailfac_Raw + facassn_Raw + classrel_Raw + feedback_Raw + facidoc_Raw + fftuse_Raw + sklabuse_Raw + comlbuse_Raw) / 16
```

The final step is creating the raw benchmark score for a given population subgroup. This is accomplished by computing the weighted average of the raw benchmark score (ENGAGLRN) for all respondents in the subgroups of interest. (IMPORTANT NOTE: If your population subgroup is based on enrollment status (Full-time vs Part-time), then the weight should not be used; instead the unweighted average should be calculated. See "When To Use Weights" for a more detailed discussion of using weights in analyzing SENSE data.)

Your college data set will include both raw benchmark scores for each respondent as well as the standardized benchmark scores for each respondent. The raw benchmark variable for the Clear Academic Plan and Pathway benchmark is ENGAGLRN and the standardized benchmark variable is ENGAGLRN\_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting ENGAGLRN\_STD for ENGAGLRN.

## Academic and Social Support Network (items 18L, 18m, 18n, 18o, 18q, 18r, 18s)

Academic and Social Support Network does not include any items that require reverse coding, so the first step above is not applicable.

The process for converting the original scale for each item to a 0–1 scale is the same varying only by the number of response options for any given item. The math for converting each item is presented below. (Note: the lowest numeric response value for all items in the original scale is 1 and the highest value is 2 for Q12 and Q14 items and 5 for the Q21 items.)

```
Q18I (5-point scale): resource_Raw = (resource -1) / 4
Q18m (5-point scale): gradepol_Raw = (gradepol -1) / 4
Q18n (5-point scale): syllabi_Raw = (syllabi -1) / 4
Q18o (5-point scale): facmeet_Raw = (facmeet -1) / 4
Q18r (5-point scale): facnam_Raw = (facnam -1) / 4
Q18s (5-point scale): stunam Raw = (stunam -1) / 4
```



The new recoded variables can now be used to calculate the raw individual-level benchmark scores. This is simply a matter of computing the average of the seven recoded items in this scale:

ACSOCSUP= (resource\_Raw + gradepol\_Raw + syllabi\_Raw + facmeet\_Raw + ostudnam\_Raw + facnam\_Raw + stunam\_Raw) / 7

The final step is creating the raw benchmark score for a given population subgroup. This is accomplished by computing the weighted average of the raw benchmark score (ACSOCSUP) for all respondents in the subgroups of interest. (IMPORTANT NOTE: If your population subgroup is based on enrollment status (Full-time vs Part-time), then the weight should not be used; instead the unweighted average should be calculated. See "When To Use Weights" for a more detailed discussion of using weights in analyzing SENSE data.)

Your college data set will include both raw benchmark scores for each respondent as well as the standardized benchmark scores for each respondent. The raw benchmark variable for the Clear Academic Plan and Pathway benchmark is ACSOCSUP, and the standardized benchmark variable is ACSOCSUP\_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting ACSOCSUP\_STD for ACSOCSUP.

### When to use Standardized and Raw Benchmark Scores

SENSE creates two types of benchmark scores, raw and standardized, for the release of the SENSE survey results. Both types of benchmarks are useful, but for different purposes. The standardized benchmark scores SENSE provides are useful for comparing any given college to the cohort at one point in time. You can also use standardized benchmarks to compare subgroups within your college or campus: a subgroup, e.g. full-time students, with a standardized benchmark of 52 will be more engaged on that benchmark than a subgroup, e.g. part-time students, with a standardized score of 47.

For colleges wishing to conduct longitudinal analyses of trends at their campus, the raw benchmark scores are the appropriate measures to use. The standardized benchmark scores are not appropriate for longitudinal analysis as they are recalculated every year and are based on the distribution of responses for each annual 3-year cohort. The raw benchmark scores, on the other hand, are always on a 0-1 scale and are not affected by fluctuations in the distribution of national responses from year to year.

Back to top



## When to Use Weights

In the paper-and-pencil administration, full-time students, who by definition are enrolled in more classes than part-time students, are more likely to be sampled. In the online administration, more full-time students tend to respond. To adjust for this, *SENSE* results are weighted using the most recently available IPEDS data. College data sets include a variable called IWEIGHT that includes the appropriate weight for each respondent. This variable is also used in the *SENSE* online reporting feature.

Because weights are based on enrollment status, analysis of *SENSE* results by enrollment status in which part-time students are in one group and full-time students are in a different group, should not employ weights. When comparing subgroups broken out by enrollment status (e.g., part-time male with part-time female students), weights should not be used. When comparing all members of one subgroup (with members of a different subgroup (e.g., all developmental students with non-developmental students in which both part-time and full-time students are included in the analysis), weights should be used.

When reporting simple demographics (e.g., the number of male and female students, number of respondents by race/ethnicity), weights should not be used.

As noted above, weights are determined using the most recent publicly available IPEDS data. As IPEDS data are approximately 3 years old, they may not accurately represent a college's current student population. For example, in the case that a college has experienced a significant change in enrollment characteristics during the three years prior to administering *SENSE*, the college's institutional research department may want to consider whether the weights based on IPEDS numbers are completely appropriate.

A final example of when to consider not using weights is when the vast majority of students at a given college are either full-time or part-time. As an example, if a college has 92% full-time students, that college may want to look at the unweighted results for full-time students to guide many campus decisions.

SENSE encourages each member college to carefully compare the student characteristics of its SENSE sample with the characteristics of its student population from which the sample was drawn in order to evaluate the effect of a possible sampling bias.