How Benchmarks Are Calculated

2022-PRESENT

CCSSE benchmarks are groups of conceptually related survey items that focus on institutional practices and student behaviors that promote student engagement—and that are positively related to student learning and persistence. The five *CCSSE* benchmarks are active and collaborative learning, academic challenge, student effort, student-faculty interaction, and support for learners. Two types of benchmark scores are included in each college's data set for each respondent: raw benchmark scores and standardized benchmark scores.

Standardized benchmark scores are useful for comparing one college to a comparison group of colleges (e.g., other colleges of a similar size) or the three-year cohort at any one point in time. Standardized benchmark scores can also be used to determine how well subgroups within the college are doing relative to other subgroups, such as developmental and non-developmental students.

Raw benchmark scores are the appropriate measures to use for colleges that wish to conduct longitudinal trend analyses. 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 three-year cohort. Raw benchmark scores, on the other hand, are not affected by fluctuations in the distribution of national responses from year to year.

The creation of both types of benchmark scores involves reverse coding items where necessary and converting all responses to the same scale. After these initial steps are taken, raw benchmark scores are computed for each respondent by averaging the rescaled scores of their related survey items. Benchmark scores are then standardized around the mean of *CCSSE* cohort respondents' scores so that benchmarks have a mean of 50, a standard deviation of 25, and are weighted by enrollment status and gender identity. A standard deviation of 25 is used to ensure that over 95% of benchmark scores fall between zero and 100, providing an understandable scale for member colleges.

Please note that individual colleges cannot compute standardized benchmark scores as this process can only be completed using the full three-year cohort data set. However, calculated standardized and raw benchmark scores for each student are included in each college's data file. College, campus, and group-level benchmarks can be calculated by computing the weighted average of the individual benchmark scores, either raw or standardized. The steps used to create the benchmark scores are explained in detail below.

Creating Benchmark Scores

1. Reverse code items where necessary.

The first step is to determine which items, if any, need to be reverse coded so that a high score on the item represents a desirable behavior. For example, the item "Come to class without completing readings or assignments" (item 4e) is originally coded such that 1= Never and 4=Very often. In this case, "Never" should have a higher positive impact on the benchmark score than coming to class unprepared "Very often". The easiest way to reverse code this item is to use the following formula that assumes, as in the case of "Come to class without completing readings or assignments," the item has four response options:

© 2025 CCCSE, The University of Texas at Austin

ReverseScore = 5 – OriginalScore.

For “Come to class without completing readings or assignments,” the reverse codes would be: CLUNPREP_Rev = 5 – CLUNPREP.

4 = 5-1 (4 becomes the value for “Never”)
3 = 5-2 (3 becomes the value for “Sometimes”)
2 = 5-3 (2 becomes the value for “Often”)
1 = 5-4 (1 becomes the value for “Very often”)

(NOTE: Item 4e is the only item that needs to be reverse coded on CCSSE.)

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

After reverse coding item 4e, the next step is to convert all items to a scale with a range of 0 to 1. The following formula is used to accomplish this conversion:

RescaledScore = (OriginalScore - 1) / (max_response_value - 1).

Using the “Come to class without completing the readings or assignments” example again, where the original variable name is CLUNPREP, and the maximum response value is 4, the formula would be:

CLUNPREP_RevRaw = (CLUNPREP_Rev – 1) / (4-1)
0.00 = (1 – 1) / (4-1)
0.33 = (2 – 1) / (4-1)
0.66 = (3 – 1) / (4-1)
1.00 = (4 – 1) / (4-1)

(NOTE: Remember, when working with the reverse-coded items, the reverse-coded variable should be used in this step.)

3. Create raw benchmark scores.

Calculation of the raw benchmark scores uses the newly-created rescaled (0 to 1) variables. The raw benchmark scores are created by calculating the average score of the re-scaled items that compose the benchmark. Using the Active and Collaborative Learning (ACTCOLL) benchmark as an example and the item numbers from the survey as variable names [the variable names are enclosed in square brackets], the formula for computing the raw benchmark score is:

ACTCOLL_Raw = (4a[clquest] + 4b[clpresen] + 4f[classgrp] + 4g[occgrp] + 4h[tutor] +
4i[particcbp] + 4r[oocideas]) / 7

The divisor has a value of 7 because there are seven survey items that make up the active and collaborative learning benchmark. In contrast, there are six items that constitute the Student-Faculty Interaction benchmark. Thus, the divisor when calculating the raw Student-Faculty Interaction benchmark score would be 6.

4. Compute standardized benchmark scores.

Before explaining this step, it is important to reiterate that standardized benchmark scores cannot be computed without having the entire cohort data set (all respondents included in the 3-year cohort). As

© 2025 CCCSE, The University of Texas at Austin

such, this step is only briefly explained.

The Center uses the STANDARD procedure in SAS to create the standardized benchmark scores across the three-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. Compute group-level benchmark scores.

The processes explained above create benchmark scores (raw and standardized) for every respondent in the primary sample. The process for creating group-level (e.g., male and female) benchmark scores is the same as 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.

Colleges can choose to administer *CCSSE* and *SENSE* either in-class or online. For the colleges administering in-class, the survey is administered to a sample of students. Sampling is conducted 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. Sampling is not conducted for the colleges that administer the survey online; all eligible students are invited to participate. However, online respondents tend to be disproportionately full-time and women. To account for these known biases in the respondent populations for both types of survey administration, 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.

See "When to Use Weights" for a more detailed discussion of using weights in analyzing *CCSSE* data.

(NOTE: Standardized benchmark scores are not created for oversample respondents.)

© 2025 CCCSE, The University of Texas at Austin

Computing the Five *CCSSE* Benchmark Scores

As previously discussed in detail, 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)
3. Computing the benchmark score
4. Computing group-level benchmark scores
   a. Raw benchmark scores
   b. Standardized benchmark scores

Active and Collaborative Learning (7 items: 4a, 4b, 4f, 4g, 4h, 4i, and 4q)

The Active and Collaborative Learning benchmark 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 as described above, varying only by the number of response options for any given item. The math for converting each item is presented below.

Item 4a	(4-point scale, 1-4): clquest_raw = (clquest – 1) / 3
Item 4b	(4-point scale, 1-4): clpresen_raw = (clpresen – 1) / 3
Item 4f	(4-point scale, 1-4): classgrp_raw = (classgrp – 1) / 3
Item 4g	(4-point scale, 1-4): occgrp_raw = (occgrp – 1) / 3
Item 4h	(4-point scale, 1-4): tutor_raw = (tutor – 1) / 3
Item 4i	(4-point scale, 1-4): particcbp_raw = (particcbp - 1) /3
Item 4q	(4-point scale, 1-4): oocideas_raw = (oocideas - 1) / 3

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 seven rescaled items:

ACTCOLL = (clquest_raw + clpresen_raw + classgrp_raw + occgrp_raw + tutor_raw + particcbp_raw + oocideas_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 (ACTCOLL) for all respondents in the subgroup of interest.

In a college's data file, the raw benchmark variable for the Active and Collaborative Learning is ACTCOLL and the standardized benchmark variable is ACTCOLL_STD. Computation of a population subgroup standardized benchmark score follows the same procedure just described for the raw subgroup benchmark score substituting ACTCOLL_STD for ACTCOLL.

Student Effort (8 items: 4c, 4d, 4e, 6b, 10a, 12d1, 12e1, and 12h1)

The Student Effort benchmark contains one item (4e) that requires reverse coding. This item is reverse coded using the following process:

Q4e (4-point scale): clunprep_rev = (5 – clunprep)

© 2025 CCCSE, The University of Texas at Austin

The process for converting the original scale for each item to a 0–1 scale is the same as described above, varying only by the number of response options for any given item. The math for converting each item is presented below.

Item 4c	(4-point scale, 1-4): rewropap_raw = (rewropap – 1) / 3
Item 4d	(4-point scale, 1-4): integrat_raw = (integrat – 1) / 3
Item 4e	(4-point scale, 1-4): clunprep_revraw = (clunprep_rev – 1) / 3
Item 6b	(5-point scale, 0-4a): bkreadown _raw = (bkreadown) / 4
Item 10a	(6-point scale, 0-5a): acadpr01_raw = (acadpr01) / 5
Item 12d1	(4-point scale, 0-3a): freqtutor_raw = (freqtutor) / 3
Item 12e1	(4-point scale, 0-3a): freqlab_raw = (freqlab) / 3
Item 12h1	(4-point scale, 0-3a): freqcomlb_raw = (freqcomlb) / 3

(NOTE: If the lowest value on the original scale is zero, there is no need to subtract 1 from the original scale prior to division to create the 0-1 rescaled variable.)

The new rescaled variables can now be used to calculate the raw individual-level benchmark scores by computing the average of the eight rescaled items:

STUEFF = (rewropap_raw + integrat_raw + clunprep_revraw + bkreadown _raw + acadpr01_raw + freqtutor_raw + freqlab_raw + freqcomlb_raw) / 8

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 (STUEFF) for all respondents in the subgroup of interest.

In a college's data file, the raw benchmark variable for the Student Effort benchmark is STUEFF, and the standardized benchmark variable is STUEFF_STD. Computation of a subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup benchmark score substituting STUEFF_STD for STUEFF.

Academic Challenge (10 items: 5b, 5c, 5d, 5e, 5f, 6a, 6c, 7, 9a, and 4o)

The Academic Challenge benchmark does not include any items requiring 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 as described above, varying only by the number of response options for any given item. The math for converting each item is presented below.

Item 5b	(4-point scale, 1-4): analyze_raw = (analyze – 1) / 3
Item 5c	(4-point scale, 1-4): newideas_raw = (newideas – 1) / 3
Item 5d	(4-point scale, 1-4): evaluate_raw = (evaluate – 1) / 3
Item 5e	(4-point scale, 1-4): applying_raw = (applying – 1) / 3
Item 5f	(4-point scale, 1-4): perform_raw = (perform – 1) / 3
Item 6a	(5-point scale, 0-4a): assigread_raw = (assigread) / 4
Item 6c	(5-point scale, 0-4a): numpaprrpts_raw = (numpaprrpts) / 4
Item 7	(7-point scale, 1-7): chalngxam _raw = (chalngxam – 1) / 6
Item 9a	(4-point scale, 1-4): envschol_raw = (envschol – 1) / 3
Item 4o	(4-point scale, 1-4): workhard_raw = (workhard – 1) / 3

(NOTE: If the lowest value on the original scale is zero, there is no need to subtract 1 from the original scale prior to division to create the 0-1 rescaled variable.)

© 2025 CCCSE, The University of Texas at Austin

The new rescaled variables can now be used to calculate the raw individual-level benchmark scores by computing the average of the ten rescaled items:

ACCHALL = (analyze_raw + newideas_raw + evaluate_raw + applying_raw + perform_raw + assigread_raw + numpaprrpts_raw + chalngxam _raw + envschol_raw + workhard_raw) /10

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 (ACCHALL) for all respondents in the subgroup of interest.

In a college's data file, the raw benchmark variable for the Academic Challenge benchmark is ACCHALL and the standardized benchmark variable is ACCHALL_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting ACCHALL_STD for ACCHALL.

Student-Faculty Interaction (6 items: 4j, 4k, 4l, 4m, 4n, and 4p)

The Student-Faculty Interaction benchmark does not include any items requiring 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 as described above, varying only by the number of response options for any given item. The math for converting each item is presented below.

Item 4j	(4-point scale, 1-4): email_raw = (email - 1) / 3
Item 4k	(4-point scale, 1-4): facgrade_raw = (facgrade - 1) / 3
Item 4l	(4-point scale, 1-4): facplans_raw = (facplans - 1) / 3
Item 4m	(4-point scale, 1-4): facideas_raw = (facideas – 1) / 3
Item 4n	(4-point scale, 1-4): facfeed_raw = (facfeed – 1) / 3
Item 4p	(4-point scale, 1-4): facoth_raw = (facoth – 1) / 3

The new rescaled variables can now be used to calculate the raw individual-level benchmark scores by computing the average of the six rescaled items in this scale:

STUFAC = (email_raw + facgrade_raw + facplans_raw + facideas_raw + facfeed_raw + facoth_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 (STUFAC) for all respondents in the subgroup of interest.

In a college's data file, the raw benchmark variable for the Student-Faculty Interaction benchmark is STUFAC and the standardized benchmark variable is STUFAC_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting STUFAC_STD for STUFAC.

© 2025 CCCSE, The University of Texas at Austin

Support for Learners (7 items: 9b, 9c, 9d, 9e, 9f, 12a1, and 12b1)

The Support for Learners benchmark 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 as described above, varying only by the number of response options for any given item. The math for converting each item is presented below.

Item 9b	(4-point scale, 1-4): envsuprt_raw = (envsuprt – 1) / 3
Item 9c(	4-point scale, 1-4): envdivrs_raw = (envdivrs – 1) / 3
Item 9d	(4-point scale, 1-4): envnacad_raw = (envnacad –1) / 3
Item 9e	(4-point scale, 1-4): envsocal_raw = (envsocal – 1) / 3
Item 9f	(4-point scale, 1-4): finsupp_raw = (finsupp – 1) / 3
Item 12a1	(4-point scale, 0-3a): freqacad_raw = (freqacad) / 3
Item 12b1	(4-point scale, 0-3a): freqcacou_raw = (freqcacou) / 3

(NOTE: If the lowest value on the original scale is zero, there is no need to subtract 1 from the original scale prior to division to create the 0-1 rescaled variable.)

The new rescaled variables can now be used to calculate the raw individual-level benchmark scores by computing the average of the seven rescaled items:

SUPPORT = (envsuprt_raw + endivrs_raw + envnacad_raw + envsocal_raw + finsupp_raw + freqacad_raw + freqcacou_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 (SUPPORT) for all respondents in the subgroup of interest.

In a college's data file, the raw benchmark variable for the Support for Learners benchmark is SUPPORT and the standardized benchmark variable is SUPPORT_STD. Computation of a population subgroup standardized benchmark score follows the same procedure as just described for the raw subgroup population benchmark score substituting SUPPORT_STD for SUPPORT.

© 2025 CCCSE, The University of Texas at Austin

When to Use Weights

Prior to *CCSSE* 2022 reporting, weights were based on enrollment status alone to address bias introduced by the sampling process (i.e., full-time students who, by definition, are enrolled in more classes than part-time students, are more likely to be sampled). On the other hand, respondents to online surveys tend to be disproportionately full-time and women. Therefore, for *CCSSE* and *SENSE* reporting moving forward, we use new post-stratification weights based on both enrollment status and gender. Weights are calculated for each institution and are based on the most recent publicly available IPEDS enrollment figures. These new weights will be applied to both the paper-and-pencil and online survey results. The use of the new weights will not negatively impact the analysis of the paper-and-pencil surveys. For example, if the proportion of men and women among your survey respondents exactly matches the proportion in your population, results based on the new weights would be the same as results based on enrollment status alone. One detail to note is that IPEDS only reports male and female percentages while *CCSSE*'s gender item includes man, woman, other, and I prefer not to respond. Students who selected "I prefer not to respond" are treated as missing and their weight value is missing. Students who responded "Other" are assigned a weight of 1 (one). Without this assignment, students who reported their gender as "Other" would not have a weight and would be lost to all weighted analyses.

The new weight variable is called IWEIGHT and it is included in all college raw data sets. In case institutional researchers are interested in weighting analyses with just the enrollment status, the raw data files also include the variable IWEIGHT_ENRL. Please note, however, that this variable cannot be used when running benchmark analyses with the standardized benchmark scores because the weight variable based on both enrollment status and gender was used to create the standardized benchmark scores.

Because weights are based on both enrollment status and gender, analyses of *CCSSE* results in which either enrollment status or gender are used as the breakout variable should not use weights (except for analysis of standardized benchmark scores). Further, when reporting simple demographics (e.g., the distribution of age or race/ethnicity), weights should not be used. When comparing all members of one subgroup with members of another subgroup (e.g., developmental students compared to non-developmental students or traditional-age and nontraditional-age students), weights should be used.

As noted above, weights are determined using the most recent publicly available IPEDS data. As the publicly available IPEDS data at the time the *CCSSE* data set is created are approximately two years old, they may not accurately reflect a college's current student population. For example, in the case that a college has experienced a significant change in enrollment characteristics during the two years prior to administering *CCSSE*, the college's institutional researchers may want to consider whether the weights based on IPEDS numbers are completely appropriate.

Another example of when to consider not using weights is when the vast majority of students at the college are either full-time or part-time or predominantly men or women. As an example, if 92% of students are full-time or 95% men, a college may want to look at the unweighted results for full-time students or men, respectively, to guide many campus decisions.

© 2025 CCCSE, The University of Texas at Austin