

Supplementary appendix 1: methodology for composite educational indices

We developed new composite indices based on Early Development Instrument assessments, grade 3, 7 & 8 reading and numeracy assessments, the grade 9 index, and the grade 12 language arts and math indices. We describe here the methodology employed for these indices.

While these indices can be a useful measure of overall academic performance, composite indices can mask challenges only seen in one area of an assessment, but not strongly affecting the mean composite score. This is particularly true in the earlier years assessments such as EDI and grade 3 reading and numeracy, where a developmental vulnerability or need for ongoing help in one domain or area of assessment may not show up in the composite score but could greatly impact future educational achievement.

Mean early development score (MEDS)

The EDI assessment calculates numeric scores in the range 0-10 for five development domains: Physical Health & Well Being, Social competence, Emotional Maturity, Language and Cognitive Development, and Communication Skills & General Knowledge. For each of these scores a comparison is made to the population-level 10th percentile value as a cutoff. Scores below the cutoff for their domain indicate the child is developmentally vulnerable in that development domain [39, 50].

The five EDI assessment domain scores are used to calculate a composite score based on the methodology recommended by Eric Duku et al., at the Offord Centre. This methodology was described in an email conversation with Eric Duku, and a poster entitled *Creating and evaluating two cumulative developmental vulnerability risk measures*, by Eric Duku, Molly Pottruff & Magdalena Janus, Offord Centre for Child Studies, McMaster University, Hamilton, Ontario, Canada [51]. The technique involves normalising the numeric scores in the range 0–10 for the five development domains. While the Offord Centre normalises the domain scores at the full population level, for consistency with our other educational indices we normalise at the cohort level. However, for this first normalisation step, all cohort years are averaged together. For each of these scores, a normalised Z-score value is calculated. A mean composite score is calculated from the five Z-scores for each person in the cohort.

In our own addition to the methodology recommended by the Offord Centre, the mean composite scores are re-normalised at the cohort year level, to obtain a consistent scaling with the other indices. Without this final normalisation into a new Z-score, the SD of the mean of means would be less than 1. While far less year-to-year variation in testing methods is expected for this assessment than the others, normalising at the cohort year level correct, for any year-to-year variations. The score is only calculated if numeric scores for all five domains are available.

Compared with the educational indices based on scaled logit scores, the MEDS values tend to have a more skewed distribution, with a longer tail in the negative values.

Grade 3 reading and numeracy assessment index

The grade 3 assessments measure seven competencies, three in the area of reading, and four in numeracy [52]. When a child is assessed, one of the following codes is recorded for each competency:

- 0004 - Out of Range - below
- 0030 - Needs Ongoing Help
- 0031 - Approaching Expectations
- 0032 - Meeting Expectations

From these individual competency assessments, composite indicators are determined for reading and numeracy, for meeting all competencies in the area, and for meeting or approaching all competencies.

For the composite index, we first convert the 4 assessment result codes above to a number from 1-4, and sum up the numeric scores for the areas of reading and numeracy. To give the sums equal weight, we take of the sum for numeracy, as it is composed of four competency scores, while the reading score is composed of only three. These values are added to get a composite score in the range 6-24, which is subtracted from 25, then rounded down to an integer value (floor), to get a ranking from 1-19. This ranking is then used to calculate a scaled logit score in the same manner as for the grade 9 achievement index and the grade 12 LA & Math indices [53, 54]. The score is only calculated if valid results for all seven competencies are available.

Grade 7 numeracy assessment index

The grade 7 assessments measure several competencies in the area of number skills and number sense.[52] When a child is assessed, one of the following codes is recorded for each competency:

- 0004 - Out of Range - below
- 0016 - Not Meeting Mid-Grade 7 Level of Performance
- 0017 - Approaching Mid-Grade 7 Level of Performance
- 0018 - Meeting Mid-Grade 7 Level of Performance

From these individual competency assessments, composite indicators are determined for numeracy, for meeting all competencies in the area, and for meeting or approaching all competencies.

For the composite index, we first convert the 4 assessment result codes above to a number from 1-4, and sum up the numeric scores for the area of numeracy. These values are added to get a composite score in the range 5-20, which is subtracted from 21 to get a ranking from 1-16. This ranking is then used to calculate a scaled logit score in the same manner as for the grade 9 achievement index and the grade 12 LA & Math indices [53, 54]. The score is only calculated if valid results for all five competencies are available.

Grade 8 reading and writing assessment index

The grade 8 assessments measure six competencies in the areas of reading comprehension and expository writing [52]. When a child is assessed, one of the following codes is recorded for each competency:

- 0004 - Out of Range - below
- 0020 - Not Meeting Mid-Grade 8 Level of Performance
- 0021 - Approaching Mid-Grade 8 Level of Performance
- 0022 - Meeting Mid-Grade 8 Level of Performance

From these individual competency assessments, composite indicators are determined for reading and writing performance, for meeting all competencies in the area, and for meeting or approaching all competencies.

The composite index involves first converting the 4 assessment result codes above to a number from 1-4 and sum up the numeric scores for the areas of reading and writing. These values are added to get a composite score in the range 6-24, which is subtracted from 25 to get a ranking from 1-19. This ranking is then used to calculate a scaled logit score in the same manner as for the grade 9 achievement index and the grade 12 LA & Math indices [53, 54]. The score is only calculated if valid results for all six competencies are available.

Grade 9 achievement index

The grade 9 achievement index, also known as the grade 9 performance index, is calculated from the high school enrollment, marks and assessment data [54]. A scaled logit score, with a mean of 0 and standard deviation scaled to ± 1 , is created based on an overall ranking for each student.

This ranking takes into consideration the number of credits passed and the average grade in the expected year of grade 9 attendance. Students who earned fewer than 7 credits are assigned a lower rank category than those who earned similar marks but completed more credits. Rankings from 1-16 are determined based on the criteria detailed in the above-referenced concept. A frequency analysis on the rankings for the entire cohort year is used to calculate the logit score using a formula presented in Willms (1986), for the grade 12 LA & Math indices [53, 55]. To normalise the logit scores, we calculate the mean and standard deviation of these scores for each cohort year, scale each individual score by subtracting the mean score, and then divide by the standard deviation. This yields scores with a mean of 0 and SD of 1 for the cohort year. This produced scores comparable from year to year, despite differences over time in testing and scoring.

Grade 12 language arts and math composite index

The grade 12 LA & Math achievement indices are already calculated as scaled logit scores, with a mean of 0 and standard deviation scaled to ± 1 [53]. These two logit scores are created from ranking categories based on scores in the grade 12 language arts and mathematics standards tests. A frequency analysis of these two rankings, from 1-19, is used to calculate the two logit scores. To calculate a composite index for grade 12, we simply took the mean of the two scaled logit scores for LA and Math, giving them equal weight. As the standard deviation for the combined score differed from the mean of the previous two SDs, to maintain the same scaling, the score for each person is divided by the recalculated standard deviation of scores for all students in the same year, in the same way as the original two logit scores were scaled.

