

NWEA AEC Research Findings to Inform How to Best Articulate the Assessments Use

For Purposes of AEC Accountability

Prepared by Jody L. Ernst, Ph.D. for the Colorado Department of Education

Purpose

The purpose of this report is to provide a research based method upon which to 1) determine how alternative education campuses (“AECs”) can best utilize assessment data from the Northwest Education Association’s Measures of Academic Progress (“NWEA”) to set targets for student growth, and 2) make recommendation for how schools, districts, and the state can use aggregate NWEA data to inform how students attending AECs are growing toward and achieving grade level content in the core academic areas of mathematics, reading, and language use.

Data

NWEA test results, RIT scores, were obtained on 24 of the 43 alternative education campuses that use NWEA in Colorado. Three years worth of data (2007-08, 2008-09, and 2009-10) for each school, when available, was collected to increase the number of students in the overall sample, as well as to provide data on as many students as possible for each of the schools. Results for the mathematics, reading and language use assessments were analyzed separately.

The sample included test results for 7472 students for the reading assessment, 5106 for the language use assessment, and 6947 for the mathematics assessment. However, only 35 percent of students had multiple scores upon which growth could be calculated in each of the three subjects.

Analyses

In order to understand the data better, and inform how we might go about setting benchmarks for AECs, a number of exploratory analyses were done. The following is a list of the analyses conducted:

- Comparison of the students age at the time of the fall test with the students fall grade
- Frequencies on the number of weeks between test administrations
- Correlation between number of weeks between assessments and growth results
- Computation of grade equivalents, based on NWEA’s 2008 Norm Placement document, compared to the students actual grade
- Average number of years behind grade level
- Average growth by grade
- Average growth by fall grade equivalent
- Average growth by school
- Average grade equivalent compared to average student placement grade, by school

Results

For purposes of illustration, and because results across subjects were extremely similar, the following results are provided for mathematics only. Reading and language use results are referred to when it is thought to be important and are also available upon request.

Student's age

To begin with the age of the students was reviewed in a crosstab with the grade that the students were placed in. This analysis shows the distribution of the sample in terms of students who attend AECs that are in age appropriate grades and those who are over age for their grade. This is important to consider because it might influence which students are, or are not, included in the preparation of AEC benchmarks.

When comparing the students age at the time of the fall test with the students fall placement grade, it was found that at least 43 percent of the students attending the AECs in this sample were over age for their grade level, highlighted in green (see Table 1). A few students were also found to be significantly below the normal age for their grade.

Based on this analysis, it is recommended that the outliers be excluded from the computation of the metrics for the AEC framework. However, over age students are well represented in this sample and are believed to reflect the age distribution found in most AECs. Therefore, it is recommended that the over aged students remain in the sample for the computation of the NWEA metrics.

Table 1. Age of Student During the Fall Test Administration, by Placement Grade

NWEA Math	Fall Grade						Total
	7th	8th	9th	10th	11th	12th	
Age (in years) at Fall Assessment							
8	0	0	0	0	0	1	1
9	0	0	1	1	0	0	2
10	0	0	0	0	0	2	2
11	1	0	1	0	0	1	3
12	42	1	1	0	0	0	44
13	14	64	8	0	0	0	86
14	5	44	146	7	0	0	202
15	0	8	182	324	5	2	521
16	0	0	264	419	296	18	997
17	0	0	217	309	352	311	1189
18	0	0	116	145	254	313	828
19	0	0	58	41	110	182	391
20	0	0	20	20	36	84	160
21	0	0	2	1	10	15	28
24	0	0	0	0	1	0	1
27	0	0	0	0	0	1	1
Total	62	117	1016	1267	1064	930	4456

Outliers Age appropriate for grade Overage

Time between Assessments

One issue to consider when developing standardized cut-scores for growth on an assessment is the extent to which there is standardization in the timing of the test administrations. Evidence presented in Table 2 suggests that there is little to no standardization between the times that NWEA tests are administered.

Table 2. Frequencies on the Number of Weeks between Math Assessments					
Fall to Winter		Winter to Fall		Fall to Spring	
# of Weeks	Frequency	# of Weeks	Frequency	# of Weeks	Frequency
27	1	29	4	41	2
26	13	28	6	40	18
25	14	27	1	39	7
24	40	25	2	38	6
23	22	24	1	37	20
22	16	23	2	36	18
21	25	22	10	35	48
20	39	21	31	34	160
19	80	20	41	33	180
18	79	19	41	32	215
17	168	18	45	31	277
16	89	17	69	30	202
15	97	16	159	29	167
14	225	15	142	28	63
13	121	14	109	27	83
12	40	13	133	26	60
11	29	12	134	25	19
10	78	11	50	24	12
9	50	10	10	23	4
8	23	9	1	22	8
7	20	8	2	21	5
6	13	5	1	20	4
5	1			19	14
3	1				
Ave of 15		Ave of 15		Ave of 30	

However, the number of weeks between administrations has a very low, though statistically significant, correlation with the amount of growth that AEC students achieve between fall and winter (-0.09) and between winter and spring (0.13), and a non-significant relationship between fall and spring (0.002).

Based on the low, to no, correlational relationship between the number of weeks between assessments and the growth achieved by AEC students, time between test administrations is not thought to be an important component in the computation of the NWEA AEC metrics.

Student Skill Level

AECs often state that the students they serve come in to the school significantly behind grade level in core academic areas. If that is truly the case, than grade level equivalent scores on the pre-test might be a good indicator of how much growth we see students achieve throughout the year.

For this analysis, grade level equivalents were computed by using the students fall NWEA RIT score and comparing them to the median values on the NWEA Norm Placement tables. Here the median value was used at the “cut-point” for the grade level equivalent. For example, the median math RIT for 3rd grade is 192 and for 4th grade is 203. Here the 3rd grade equivalent was defined as a score above the 2nd grade median (179) up to (and including) 192. The fourth grade equivalent was then defined as above 192 up to 203. Therefore, these grade equivalents should only be seen as rough approximations, but good enough to explore the approximate skill levels (plus or minus one grade) of the students in the beginning of the academic year. These resulting “grade equivalent scores” were then crosstabulated with the actual grade the students were placed in that same fall to produce a frequency distribution of “skill level” within each grade served by AECs in the sample.

As can be seen in Table 3, very few AEC students begin the academic year at the grade level they are placed in. In fact, between 62 and 81 percent of students test at least one year behind grade level, depending on the subject area.

Fall Test Grade Equivalent	Fall Placement Grade						Total
	7th	8th	9th	10th	11th	12th	
Kindergarten	0	0	1	0	0	1	2
1st grade	0	1	3	1	1	0	6
2nd grade	6	5	11	11	9	7	49
3rd grade	10	14	51	51	45	25	196
4th grade	14	19	123	114	77	73	420
5th grade	11	18	217	218	150	120	734
6th grade	12	15	155	191	146	110	629
7th grade	2	15	153	172	179	125	646
8th grade	2	11	91	146	100	107	457
9th grade	2	6	55	81	60	54	258
10th grade	0	3	49	92	90	80	314
11th grade	0	1	15	38	49	41	144
more than 11th grade equiv	3	9	93	152	159	188	604
Total	62	117	1017	1267	1065	931	4459

This is particularly important to consider for the 12th grade students, as NWEA developers conveyed that the NWEA assessments are inappropriate to administer with 12th graders. However, they did indicate that NWEA can be used with students who test at a lower skill level (for example a 19 year old with a 9th grade math skill level), and that NWEA's growth norms could be applied in those cases.

One additional analysis was conducted to see how far students were behind as a function of their beginning of year (fall) placement grade level (Table 4). Here a pattern emerges in which students at the higher grade levels were found to be farther behind, on average, than students in lower grade levels. On average, students in all grade levels are at least 1.5 years behind at the beginning of the school year. This is true in each of the three core academic areas tested.

AEC Students in Grade	Math	N	Reading	N	Language	N
7	-2	62	-2.2	60	-1.7	50
8	-2	117	-1.8	123	-2.4	98
9	-2.4	1017	-1.5	1141	-2.1	768
10	-2.8	1267	-2.2	1147	-2.5	962
11	-3.5	1065	-3.2	1081	-3.4	824
12	-4.1	931	-3.9	918	-3.8	644

How does this play out in schools? Do some schools have students that, on average, start the year farther behind than others, making skill level more important to consider when tracking growth than placement level? The results in table 5 suggest that fall test grade level does vary by school, and that some schools show markedly larger difference between placement grade and tested skill level. As the turnover in the student body from one year to the next is so great (averaging about 55 percent) in AECs, this should not be viewed as an effect of the school, but rather the skill level of a majority of the students entering the school for the first time.

Table 5. Mean Grade Equivalent on the Fall Math Assessment Compared to Mean Fall Grade, By School						
Fall School ID	Mean Grade Equiv	Std. Deviation	Mean Stated Grade	Std. Deviation	Means Diff	N
	6.23	2.052	9.42	.575	-3.192	73
4201						
5660	8.95	2.618	10.55	1.043	-1.606	508
7647	8.48	2.510	10.34	.479	-1.864	44
8602	6.62	2.282	10.52	.976	-3.902	61
11545	6.55	2.437	10.89	1.139	-4.341	299
12156	5.81	2.414	9.26	1.365	-3.452	31
14185	7.83	2.490	9.68	.474	-1.850	40
14215	9.00	2.485	10.98	.847	-1.979	47
16498	8.10	2.697	11.27	.887	-3.166	439
18253	6.60	2.365	9.88	.492	-3.274	106
18262	7.22	2.394	10.30	.756	-3.087	46
18269	7.23	2.706	8.93	1.064	-1.700	240
18270	7.21	2.528	10.11	.737	-2.902	82
18607	8.07	2.731	10.14	1.014	-2.064	109
19995	6.76	2.547	10.06	1.003	-3.303	747
20649	6.61	2.924	8.90	1.470	-2.288	59
21679	6.53	2.475	9.73	.458	-3.200	15
22176	6.20	2.562	10.45	1.011	-4.254	342
22418	7.38	2.878	10.58	1.109	-3.192	52
22953	6.56	2.178	11.19	1.029	-4.630	54
23542	6.15	2.397	10.93	.937	-4.779	95
25257	4.99	2.818	9.79	1.747	-4.799	164
25363	7.16	2.388	10.25	1.084	-3.088	545
27827	8.21	2.786	10.73	1.105	-2.513	261
Total	7.24	2.740	10.33	1.192	-3.093	4459

On average, students are approximately three years behind grade level in mathematics. However the range between schools varies from a low of 1.6 years behind to a high of 4.8 years behind. In reading the overall average was 2.6 years behind, with school averages ranging from 0.9 to 5.4 years behind. The average difference between placement grade and grade equivalent for language arts was 2.9 years behind, and schools ranged from an average of 1.3 to 5.4 years behind.

While the average placement grades varied as well, averaging at around the 10th grade, these appear to be a bit more consistent with a majority of the schools (17 out of 24) serving students whose average placement grade was between the 10th and 11th grade.

Average Growth by Placement Grade and Grade Equivalent Scores

Next, the average growth of AEC students on the NWEA assessments were analyzed by both placement grade (Table 6) and fall test grade equivalent (Table 7).

	Fall to Winter			Winter to Spring			Fall to Spring		
	Mean	Median	N	Mean	Median	N	Mean	Median	N
7th grade	1.7	0.91	32	4.1	5.3	30	4.1	4.9	45
8th grade	4.4	2.9	30	1.2	0.92	26	3.3	2.9	79
9th grade	2.3	2.3	305	-1.2	-0.6	168	1.9	2	495
10th grade	1.4	1.4	379	0.54	0.99	202	1.7	2.6	663
11th grade	1.6	1.6	328	1.2	1.8	186	2	2.4	538
12th grade	1.7	2	210	0.4	-0.4	98	1.6	1.7	366

Fall Test Grade Equivalent	Fall to Winter Growth			Winter to Spring Growth			Fall to Spring Growth		
	Mean	Median	N	Mean	Median	N	Mean	Median	N
1st grade	-	-	-	-	-	-	23.9	29.8	3
2nd grade	7	4	21	6.3	4.6	17	12.8	9.7	25
3rd grade	6.5	4.1	86	1.9	3.8	63	6.9	7.3	117
4th grade	5.8	5.6	142	0.65	0.23	81	6.3	5.6	204
5th grade	2.2	2	234	0.4	0.8	129	2.7	3.5	368
6th grade	0.95	0.99	222	1.7	2	109	1.4	2.2	335
7th grade	0.96	1.4	174	0.4	0.2	87	1.7	2.6	297
8th grade	0.5	1.5	131	-0.52	-0.39	73	0.94	1.3	232
9th grade	-1	0.05	61	2.3	2.5	28	-0.03	1.2	122
10th grade	-0.4	1.1	75	-0.3	1.2	43	-0.6	1.2	157
11th grade	-3.2	-2.4	21	-1.1	-3.1	11	-1.8	-1.7	60
greater than 11th grade equiv	-0.6	-0.4	116	-2.6	-3	69	-1.4	-0.5	265

In both cases, looking at growth by placement grade or fall test grade equivalent, students at lower grade levels tend to show higher average growth than students at higher grade levels. This finding is consistent with the pattern of growth found in the NWEA national norming sample. However, this pattern is more extreme in the fall test equivalent analysis (Table 7), which is at least in part due to regression to the mean—where students testing at either extreme (high or low) are likely to move toward the mean (or middle of the score distribution) on the next test administration.

These findings indicate that either placement grade or tested grade level equivalent, or both, need to be taken into consideration when average RIT score or RIT growth are used to determine the effectiveness

of a school. If this does not occur, schools serving higher grade levels, whether by placement or student skill level, will appear to be producing lower growth and achievement results, on average, than schools serving students in lower grade levels.

As was shown in table 3, very few students had fall test grade equivalent scores that matched their placement grade. These results showed that 79 percent of the students in this sample tested below their placement grade in math, 67 percent tested below their placement grade level in reading, and 73 percent did so in language arts. Therefore, I am recommending that AECs using NWEA MAPS assessments the use fall test grade equivalent as the bases for determining growth targets for their students and that the following tables (Table 8, 9 and 10) be used to establish the targets for these students.

These targets were determined using the differences between medians in the 2008 NWEA Norm Placement document, which also maps onto the average growth displayed in table 7, but do not allow for negative growth in target setting.

Fall RIT Range	Fall RIT Grade Equivalent	Fall to Winter Growth Target	Winter to Spring Growth Target	Fall to Spring Growth Target
up to 148	K	5 RIT	5 RIT	10 RIT
149-164	1 st	7 RIT	7 RIT	14 RIT
165-179	2 nd	7 RIT	5 RIT	12 RIT
180-192	3 rd	7 RIT	4 RIT	11 RIT
193-203	4 th	5 RIT	3 RIT	8 RIT
204-212	5 th	4 RIT	4 RIT	8 RIT
213-219	6 th	3 RIT	3 RIT	6 RIT
220-225	7 th	3 RIT	2 RIT	5 RIT
226-230	8 th	2 RIT	2 RIT	4 RIT
231-233	9 th	2 RIT	2 RIT	4 RIT
234-237	10 th	1 RIT	1 RIT	2 RIT
238-239	11 th	1 RIT	1 RIT	2 RIT
240 and above	Above 11 th	0.5 RIT	0.5 RIT	1 RIT

Fall RIT Range	Fall RIT Grade Equivalent	Fall to Winter Growth Target	Winter to Spring Growth Target	Fall to Spring Growth Target
up to 146	K	5 RIT	4 RIT	9 RIT
147-160	1 st	7 RIT	6 RIT	13 RIT
161-179	2 nd	7 RIT	4 RIT	11 RIT
180-192	3 rd	5 RIT	3 RIT	8 RIT
193-201	4 th	4 RIT	2 RIT	6 RIT
202-208	5 th	3 RIT	1 RIT	4 RIT
209-213	6 th	2 RIT	1 RIT	3 RIT
214-217	7 th	2 RIT	1 RIT	3 RIT
218-220	8 th	2 RIT	1 RIT	3 RIT
221-222	9 th	1 RIT	1 RIT	2 RIT
223-226	10 th	1 RIT	1 RIT	2 RIT
227	11 th	1 RIT	1 RIT	2 RIT
228 and above	Above 11 th	0.5 RIT	0.5 RIT	1 RIT

Fall RIT Range	Fall RIT Grade Equivalent	Fall to Winter Growth Target	Winter to Spring Growth Target	Fall to Spring Growth Target
up to 180	2 nd	8 RIT	4 RIT	12 RIT
181-193	3 rd	6 RIT	3 RIT	9 RIT
194-202	4 th	4 RIT	2 RIT	6 RIT
203-208	5 th	3 RIT	2 RIT	5 RIT
209-213	6 th	2 RIT	2 RIT	4 RIT
214-217	7 th	1 RIT	1 RIT	2 RIT
218-220	8 th	1 RIT	1 RIT	2 RIT
221	9 th	1 RIT	1 RIT	2 RIT
222-223	10 th	1 RIT	1 RIT	2 RIT
224-225	11 th	1 RIT	1 RIT	2 RIT
226 and above	Above 11 th	0.5 RIT	0.5 RIT	1 RIT

Use of NWEA for AEC Status Measure

For the status measure using NWEA there are a couple of metrics that could be used. One way would be to have schools report out on the percent of students to test at their placement grade level at the end of the year (or on their last assessment administration while at the school). For this metric, it is recommended including students that have been enrolled for at least 8 weeks consecutively. Cut-points for the rating categories could follow the 90/60/40 percent criteria, consistent with the other cut-points in the AEC SPF.

Another option for a status measure using NWEA would be to use the percent of students that moved up at least one grade level during the year, using the difference between the students first and last test administration event. Students should only be included that have been enrolled for at least 8 weeks and schools would follow the 90/60/40 percent cut-points for the rating categories.

Use of NWEA for AEC Growth Measure

I recommend that the tables above (Tables 8-10) be used to assess the percentage of students that met their growth targets, following the 90/60/40 percent cut-points. I recommend that the fall to winter and winter to spring growth targets be used when the time between assessments is at least 8 weeks, but not longer than 27 weeks (6 months), and that fall to spring growth targets be used when the time between assessments is at least 28 weeks, but no longer than 41 weeks apart (9 months).

Why Not the Percentile Distribution Method?

I do not recommend using the percentile distribution method because it does not take the students beginning skill level into account, which varies considerably within the student population sampled here, as well as varying between the schools represented in this sample. If we had several thousand more cases to conduct the distributions on, by beginning skill level, this option may be a good one. Alas, we do not.