STABILITY OF 2014 WYOMING SCHOOL ACCOUNTABILITY INDICATORS ACROSS YEARS

Prepared for the Wyoming Department of Education by Michael Flicek, Ed.D.
(01/27/2015)

Wyoming's school accountability model has three indicators for schools with grades 3-8. The indicators are achievement, growth and equity. Each school serving these grades is a assigned a target level on each of these three indicators. The target levels are exceeds target, meets target and below target. These target levels are entered into a decision table that determines the school's overall performance level. There are four performance levels. They are exceeds expectations, meets expectations, partially meets expectations and does not meet expectations.

Cut-points on the indicators that are established by a stakeholder group are used to place schools into the target levels on each indicator. The goal is to keep the same cut-points from year-to-year unless there is some change in the measure. This study investigated the stability of indicator scores and target level impact across three school years when the measure was held consistent. Impact was defined as the percentage of schools within each of the three target level categories for each of the three years studied. The main question of interest here was, to what extent did the percentage of school in each target level remain consistent when cut-points that were common across all three years were used for school target level assignment?

One of the three indicators, growth, actually used the same cut-points for target level assignment for two consecutive years. There was a change in the scale on the Proficiency Assessment for Wyoming Students (PAWS) in 2014, however, which resulted in setting new cut-points on the achievement and equity indicators in 2014, the second year of Wyoming school accountability. Next year, year three, the option of using previously established cut-points for target levels could be considered. This study is designed to provide some evidence relevant to the decisions to use established cut-points next year.

Achievement

The achievement indicator was the percent of achievement scores at the school that were at or above proficient. This study used Proficiency Assessment for Wyoming Students (PAWS) data from the 2010-11, 2011-12 and 2012-13 school years. The PAWS test was on the same scale for all three of these school years. There were two grade bands for achievement for the three years studied since some grades had considerably fewer proficient and above students than other grades. The percentage proficient and above within each grade band were generally consistent within content areas across grades. Grade band 1 included grades 3-6. Grade band 2 included Grades 7 and 8. When schools had grades in both grade bands, cut-points for the school were adjusted to reflect the percentage of students attending the school within each grade band. The minimum *n* used for achievement for this study was 10 for all indicators studied.

Cut-points for the target levels were established by a professional judgment panel (PJP) consisting of approximately 30 individuals representing stakeholder groups prescribed by legislation. Impact data used during the standard setting was from the 2012-13 school year. The PJP established cut-points were used to place each school into three categories on the achievement indicator for below target, meeting target or exceeding target for each of the three school years included in the study. The current analyses demonstrated the level of cross-year stability in impact when cut-points for category assignment based upon year 3 impact data were used without adjustment during years 1 and 2.

Table 1. Descriptive Statistics Percent of Students Proficient and Above for Wyoming Schools Across Three School Years for Grade Bands 1 and 2.

School Year	*n* of Schools	Mean	Standard Deviation
Grade Band 1			
2010-11	230	77.6	10.4
2011-12	251	81.1	10.9
2012-13	254	78.2	11.2
Grade Band 2			
2010-11	84	70.1	12.6
2011-12	85	70.9	14.0
2012-13	88	68.5	16.7

Table 1 findings indicate fairly consistent descriptive statistics across the years studied. Many schools had some grades in both grade bands but the results reported in Table 1 represent just the grades within the grade band at these schools.

Table 2. Percent of Schools in each Achievement Target Level for Three School Years.

School Year	*n* of Schools	Below Target	Meets Target	Exceeds Target
2010-11	260	34.2%	50.4%	15.4%
2011-12	264	25.8%	48.8%	25.4%
2012-13	266	33.2%	50.4%	16.4%

In Table 2 the percentage of schools in the exceeds target category increased about 10% in year 2 and fell back a bit in year 3. The converse of this happened in the below target category. Although there was some fluctuation in category membership across years, the findings presented here suggest that cut-points for the achievement indicator that were established in one year can reasonably be used in other years.

Next, target category agreement in category placements across adjacent years were studied. Table 3 shows the results comparing target categories for the 2010-11 year with those from the 2011-12 year.

Table 3. Comparison of Achievement Target Categories for the 2010-11 Year and the 2011-12 Year (*n* = 259 schools).

	2011-12 Year		
2010-11 Year	Below Target	Meeting Target	Exceeding Target
Below Target	55	34	0
Meeting Target	12	86	31
Exceeding Target	0	4	37

Exact agreement of school category placement across years was 68.7% and exact plus adjacent agreement was 100%. The Spearman correlation coefficient for categories across years was $r = 0.72$.

Table 4. Comparison of Achievement Target Categories for the 2011-12 Year and the 2012-13 Year (*n* = 242 schools).

2011-12 Year	2012-13 Year		
	Below Target	Meeting Target	Exceeding Target
Below Target	53	14	1
Meeting Target	32	88	6
Exceeding Target	2	30	35

Exact agreement of school category placement across years was 67.4% and exact plus adjacent agreement was 99.9%. The Spearman correlation coefficient for categories across years was $r = 0.66$.

Growth

A normative growth measure is used in Wyoming (Betebenner[1], 2011). The growth model produces a student growth percentile(SGP) for each student in grades 4 through 8. The growth indicator used for schools was the median SGP for reading and math combined which is referred to as the median growth percentile (MGP).

Cut-points for the target levels for school MGPs were established by the PJP in September 2013. The cut-points were the 45th and the 60th percentile ranks. These cut-points were used to place each school into three categories for exceeds target, meets target or below target for each of the three school years included in the study. This approach was used to demonstrate the level of cross year stability in impact when cut-points for category assignment based upon year 3 impact data were used without adjustment during years 1 and 2.

Table 5. Descriptive Statistics for MGPs for Wyoming Schools Across Three School Years.

School Year	*n* of Schools	Mean	Standard Deviation
2010-11	263	50.8	10.9
2011-12	261	51.0	10.9
2012-13	265	51.5	10.1

Table 5 findings indicate that Wyoming school growth scores were quite consistent across the three years studied.

Table 6. Percent of Schools in each Growth Target Level for Three School Years.

School Year	*n* of Schools	Below Target	Meets Target	Exceeds Target
2010-11	263	27.8%	50.6%	21.7%
2011-12	251	24.5%	55.9%	19.5%
2012-13	254	25.7%	57.0%	17.4%

[1] Betebenner, D. W. (2011). *A technical overview of the student growth percentile methodology: Student growth percentiles and percentile growth.* National Center for the Improvement of Educational Assessment.

The findings in Table 6 show very consistent impact across all three years. Using cut-points established during one school year for other years is supported by these findings.

Next, growth target category agreement in category placements across adjacent years were studied. Table 7 shows the results comparing target categories for the 2010-11 year with those from the 2011-12 year.

Table 7. Comparison of Growth Target Categories for the 2010-11 Year and the 2011-12 Year (*n* = 255 schools).

	2011-12 Year		
2010-11 Year	Below Target	Meeting Target	Exceeding Target
Below Target	34	27	9
Meeting Target	23	87	20
Exceeding Target	7	29	19

Exact agreement of school category placement across years was 54.9% and exact plus adjacent agreement was 93.7%. The Spearman correlation coefficient for categories across years was $r = 0.31$.

Table 8. Comparison of Growth Target Categories for the 2011-12 Year and the 2012-13 Year (*n* = 254 schools).

	2012-13 Year		
2011-13 Year	Below Target	Meeting Target	Exceeding Target
Below Target	35	26	3
Meeting Target	27	94	24
Exceeding Target	4	28	18

Exact agreement of school category placement across years was 56.8% and exact plus adjacent agreement was 97.3%. The Spearman correlation coefficient for categories across years was $r = 0.41$.

Equity

The equity indicator intends to encourage schools to ensure that the lowest performing students are being well served. To that end it is the performance of a consolidated subgroup that is of interest. The consolidated subgroup has been defined as those students who were not proficient in reading and/or math during the prior school year.

Percent of Subgroup Students with SGP Equal to or Exceeding AGP for Equity

The equity score during the pilot school year (i.e., 2013) was the percent of students with SGPs that equaled or exceeded their adequate growth percentiles (AGP) scores. AGP scores are the current year SGP score that a student needs in order to be considered to be on track to become proficient or better within three school years or by the end of grade 8. For this step in the study, regardless of whether the student was placed into the subgroup because of performance in reading, math or both reading and math, it was both reading and math scores for every student in the subgroup that were used in computing the school's equity score. The purpose of the analyses reported here was to determine the

extent that cut-points established by the PJP in 2013 resulted performed during the two preceding school years. The cut-points recommended by the PJP were 40 and 55.

Table 9. Descriptive Statistics for the Percent of SGPs Above AGPs for Consolidated Subgroup Students in Both Reading and Math Across Three School Years.

School Year	*n* of Schools	Mean	Standard Deviation
2010-11	203	59.9	12.0
2011-12	171	53.6	12.4
2012-13	169	42.5	12.1

Table 9 findings indicate that Wyoming school equity scores changed more than 17 percentage points across two of the three years studied.

Table 10. Percent of Schools in each Equity Category when Percent of SGPs Above AGPs for Consolidated Subgroup Students in Both Reading and Math Across Three School Years.

School Year	*n* of Schools	Below Target	Meets Target	Exceeds Target
2010-11	203	4.9%	23.6%	71.4%
2011-12	171	13.5%	37.4%	49.1%
2012-13	169	41.4%	43.2%	15.4%

The findings in Table 10 show dramatic shifts in impact across all three years when the cut-points established by the PJP using 2013 impact data were applied to the prior years. These findings support the conclusion that this equity score lacks the stability across years needed for it to be used as the equity indicator.

For the analyses above both reading and math scores were used in computing the equity score regardless of whether the student was placed in the subgroup for reading performance, math performance or both reading and math performance. For the analyses that follow, the equity score only included the content area(s) for each student where the student had a prior not proficient test score. Again, it was the extent that the school equity scores computed in this manner yielded reasonably stable impact across years that was of interest.

Table 11. Descriptive Statistics for Schools for the Percent of SGPs Above AGPs for Consolidated Subgroup Students in only the Content Area(s) where the Students had a Prior Below Proficient Score.

School Year	*n* of Schools	Mean	Standard Deviation
2010-11	207	57.9	15.0
2011-12	188	50.3	16.1
2012-13	181	40.5	14.0

Table 11 results are very similar to those from Table 9. Thus, considerable instability across years was present whether the equity score included or excluded scores in the content area in which the consolidated subgroup student was actually proficient.

Table 12. Percent of Schools in Each Equity Category when Equity Scores for Consolidated Subgroup Students Only Included the Content Area(s) where the Students had a Prior Below Proficient Score.

School Year	*n* of Schools	Below Target	Meets Target	Exceeds Target
2010-11	207	13.5%	28.0%	58.5%
2011-12	188	25.5%	31.9%	42.6%
2012-13	181	48.6%	35.4%	16.0%

The findings in Table 12 suggest that instability of category percentages across school years was smaller when the computation of the equity score was limited to just the content area where the consolidated subgroup student had a prior not proficient score. While the cross year shifts in proportions of schools within each category were not quite as dramatic under this condition, the shifts in impact across years remained sufficiently high to support an argument against the use of this equity score in the school accountability model.

Average Standardized Test Score of Consolidated Subgroup

During the 2013-14 school year AGPs were not available due to the implementation of a new scale on the PAWS. Therefore the equity score was the average standardized score for all students in the consolidated subgroup. This section looks at the stability of this school equity score across school years. Since using only the content area(s) where consolidated subgroup students were not proficient worked somewhat better than the use of both content areas for all consolidated subgroup students in the above analyses, that is the only approach studied for this measure. The standardized score used here had a mean of 150 and a standard deviation of 30 with the mean reflecting the grade level state-wide mean score for the content area. For these analyses, the cut-points for the three equity categories were set at the 30th and 70th percentile ranks for school scores from the 2012-13 school year. The cut-points identified in this manner were 122 and 128.

Table 13. Descriptive Statistics for Schools for the Mean Standardized Score for Consolidated Subgroup Students in only the Content Area(s) where the Students had a Prior Below Proficient Score.

School Year	*n* of Schools	Mean	Standard Deviation
2010-11	207	125.3	6.8
2011-12	188	119.7	5.7
2012-13	181	119.0	5.6

From Table 13 it can be seen that two of the school years had reasonably consistent mean standardized scores for the consolidated subgroups. This equity score is the mean score for a subgroup defined based upon low performance on the prior year's test. Because of this the range of scores is quite restricted which largely explains the small standard deviations. This restricted range explains why the cut-points established at the 30th and 70th percentile ranks based upon the 2012-13 data that were just six points apart. Year 1 in Table 13 was more than 5 points higher than year 3. Given the restricted range, 5 points is a fairly sizable difference in means.

Table 14. Percent of Schools in Each Equity Category based upon a Mean Standardized Score for Consolidated Subgroup Students.

School Year	*n* of Schools	Below Target	Meets Target	Exceeds Target
2010-11	207	31.4%	37.7%	30.9%
2011-12	188	67.6%	25.5%	6.9%
2012-13	181	74.0%	21.0%	5.0%

Two of the three years had a fairly consistent impact across the years, however, both of these years were quite different in impact compared to the year that was used for setting the equity target level cut-points. The outlier year had about 24% fewer schools in the below target category and 24% more schools in the exceeds target category.

Conditional Growth Approach for Equity

The conditional growth approach mirrors an approach that uses school median SGPs and school median AGPs that is illustrated in Figure 5 of the Marion and Domaleski[2] report. For the example studied here it was median SGPs and AGPs of the consolidated subgroup that were of interest. First, median SGPs and median AGPs were computed for the consolidated subgroup at each school that met the minimum *n* of 10 for the consolidated subgroup. Next, each school was placed into one of two categories based upon whether their median SGP was below the median AGP or whether their median SGP equaled or exceeded their median AGP. For schools in the former condition, the cut-points for growth established by the PJP were applied to the consolidated subgroup median SGP. For schools in the latter category, 5 points were subtracted from the PJP recommended cut-points. This resulted in the schools that met this condition having an increased possibility of falling within a higher target level category. Marion and Domaleski indicated that this condition should encourage schools to focus on getting consolidated subgroup students on-track for becoming proficient. Table 15 presents the cut-points used for the simulation presented here.

Table 15. Conditional Target Level Decision Table for Equity using School Consolidated Subgroup Median SGPs.

	Median SGP ≥ Median AGP	
	YES	NO
Exceed Target	55-99	60-99
Meet Target	40-54	45-59
Below Target	1-39	1-44

This simulation included students in the consolidated subgroup only for content areas in which the student had a prior year not proficient score. Prior to creating categories for equity, descriptive statistics were computed for school SGPs and AGPs for the consolidated subgroups.

[2] Marion & Domaleski (2012). *The Wyoming comprehensive accountability framewowrk: Phase I.* Produced for the: Wyoming Select Committee on Statewide Education Accountability.

Table 16. Descriptive Statistics for School Median SGPs and School Median AGPs for the School Consolidated Subgroup.

School Year	*n* of Schools	Median	Mean	Standard Deviation
Median SGP				
2010-11	207	50.0	51.9	11.7
2011-12	188	52.0	50.7	11.1
2012-13	181	50.5	52.1	11.3
Median AGP				
2010-11	207	41	44.0	10.4
2011-12	188	48	51.3	11.6
2012-13	181	61	63.2	8.2

The median, mean and standard deviation for school median SGPs were very consistent across all three years. There was more variability in school median AGP statistics.

Next, Table 17 presents the percent of schools within each target level category when the equity categories were based upon the conditional growth approach that utilized the conditional cut-points presented in Table 15.

Table 17. Percent of Schools in each Equity Category when Conditional Growth Approach was Applied.

School Year	*n* of Schools	Below Target	Meets Target	Exceeds Target
2010-11	207	21.3%	43.0%	35.7%
2011-12	188	25.0%	42.6%	32.4%
2012-13	181	24.3%	47.0%	28.7%

The findings in Table 17 show a fairly consistent impact in terms of the percentage of schools within each equity target level. As such, use of a conditional growth approach like that illustrated here is likely to have sufficient stability for use in the school accountability model.

Next, Table 18 shows the number of schools that were in each of the two conditions and within each target level for each of the three years studied. Table 18 also presents the percent of schools within each condition (i.e., column) that were within each target level for each school year.

Table 18. Number and Percent (within Column and School Year) of Schools in each Equity Category by Condition.

Target Level	Subgroup Median SGP ≥ Median AGP (Count of Schools)		Subgroup Median SGP ≥ Median AGP (Percent of Schools within Column)	
	Yes	No	Yes	No
2010-13				
Below	7	37	4.7%	64.9%
Meets	69	20	46.0%	35.1%
Exceeds	74	0	49.3%	0.0%
2011-12				
Below	2	45	2.0%	51.1%
Meets	41	39	41.0%	44.3%
Exceeds	57	4	57.0%	4.5%
2012-13				
Below	0	44	0.0%	31.4%
Meets	2	83	4.9%	59.3%
Exceeds	39	13	95.1%	9.3%

From Table 18 it is evident that there are many schools in both conditions in all three years. The number of schools within each condition did vary across year. There is also support for the theory of action for this indicator in that schools in the "yes" condition were much more likely to exceed target and schools in the "no" condition were much more likely to be below target. This should encourage a focus upon getting the consolidated subgroup students on track for becoming proficient.

Next, Tables 19 and 20 show the amount of movement between categories from one year to the next when the conditional growth approach was implemented.

Table 19. Comparison of Growth Target Categories for the 2010-11 Year and the 2011-12 Year (*n* = 182 schools).

2010-11 Year	2011-12 Year		
	Below Target	Meeting Target	Exceeding Target
Below Target	11	19	9
Meeting Target	22	37	22
Exceeding Target	11	22	29

Exact agreement of school category placement across years was 42.3% and exact plus adjacent agreement was 89.0%. The Spearman correlation coefficient for categories across years was $r = 0.41$.

Table 20. Comparison of Growth Target Categories for the 2011-12 Year and the 2012-13 Year (*n* = 175 schools).

2010-11 Year	2011-12 Year		
	Below Target	Meeting Target	Exceeding Target
Below Target	17	17	9
Meeting Target	18	38	19
Exceeding Target	8	26	18

Exact agreement of school category placement across years was 42.9% and exact plus adjacent agreement was 90.0%. The Spearman correlation coefficient for categories across years was *r* = 0.19. Tables 19 and 20 do show movement across categories from one year to the next. About 10% of schools moved more than one category in a year and conversely about 90% remained in the same category or moved just one category across years.

Correlation Across Achievement, Growth and Equity

Finally, Spearman correlation were computed for the target levels on the achievement, growth and conditional equity indicators. For all indicators, the minimum *n* rule applied prior to computing the coefficients was 10. The achievement indicator used was the percent of test scores in reading, math, science and writing that were proficient or above. The growth indicator was the school median SGP for reading and math combined. The equity indicator was the conditional growth approach which involved use of the cut-points in Table 15. The coefficients are presented in Table 21.

Table 21. Spearman Correlation Coefficients for Grades 3-8 Accountability Indicator Target Levels when a Conditional Growth Method was used for the Equity indicator.

	Growth	Equity (Conditional)
2010-11		
Achievement	0.41 (*n* = 250)	0.40 (*n* = 207)
Growth		0.58 (*n* = 207)
2011-12		
Achievement	0.39 (*n* = 253)	0.23 (*n* = 188)
Growth		0.48 (*n* = 188)
2012-13		
Achievement	0.43 (*n* = 256)	0.36 (*n* = 181)
Growth		0.50 (*n* = 181)

The coefficients in Table 21 are all in a moderate range. In general, moderate coefficients among different indicators for a school accountability system is a desired condition. If the coefficients were too high it would indicate that the different indicators were measuring overlapping aspects of school quality. The moderate coefficients in Table 21 provide evidence that the three indicators included in the analyses were measuring non-overlapping aspects of school quality.

Conclusions

The analyses included here suggest that the achievement and growth indicators being used in Wyoming school accountability model have adequate stability for their continued use in the model. Furthermore, the previously used equity indicators showed a lack the stability needed for their continued use in the model if a common cut-score is to be used across school years. However, the use of an equity score that is based upon a conditional growth approach was supported as an alternative method for computing an equity score.

ADDENDUM: HIGH SCHOOL EQUITY OPTIONS

Conditional Growth Approach

The analyses reported above for the grades 3-8 equity indicator are relevant to the high school equity indicator. There is presently no growth measure for high schools in Wyoming. Because of this the conditional growth approach to equity, which was the most stable equity measure identified for grades 3-8, is not available for use in Wyoming high schools in 2014-15.

Average Standardized Test Score of Consolidated Subgroup

One equity metric that has been used did not require the measurement of growth. The school equity score for this metric was the average student standardized score for consolidated subgroup students. Two issues were identified above that made this metric problematic as a equity score.

First, since these scores were the average score for a group of students defined by low prior achievement, the within year standard deviation for school scores were quite low. Table 13 reported standard deviations that ranged from 5.6 to 6.8. As a result, many schools fell at each rounded score point which resulted in a very course distribution of equity scores. Moving the target level cut-points by just one place results in large shifts in the number of schools within each target category. The impact options available to the standard setting panels lack nuance as a result.

Second, the school scores on this metric produced unstable impact across school years. Specifically, when cut-points for equity target levels established in one year were applied to subsequent years, the percentage of schools within each category changed considerably. These findings were reported in Table 14 above. When cut-points for target levels were identified based upon impact from the 2010-11 school year 31% of schools fell within the below target category that year. When the same target level cut-points were used two years later 74% of schools were within the below target category. This represented a 43% increase in the percent of schools that were in the below target category. The credibility of an indicator can be called into question when the out-year impact associated with cut-points differs markedly from the impact achieved during the standard setting.

A third issue, not mentioned above, is confusion that can be associated with student standardized scores. They actually become the fourth type of student score reported to districts on the state test. The other scores include the scaled score, the performance level score and the dichotomous score of proficient or above. School staff may have difficulty understanding yet another score on the state test.

Percent of Current Consolidated Subgroup Scores Proficient or Above

For this simulation, students were placed into the consolidated subgroup in either math, reading or both reading and math based upon whether their prior year test score in the content area was below

proficient. This means all prior year test scores for consolidated subgroup students were below proficient. Put another way, the percent of prior year consolidated subgroup math and reading scores that were proficient or above was zero for all schools.

Many students in the consolidated subgroup, however, did have scores from the current school year that were proficient or above. Therefore, the percent of current year math and reading scores for the consolidated subgroup that were proficient or above was computed to serve as the equity indicator. This simulation sought to determine if this metric for schools would yield stable impact across school years. The PAWS scores from grades 3-8 were used for this purpose since there were 3 years of test scores available for these grades. School equity scores were computed on this metric for each of the 3 school years. Only those schools with at least 10 students in the consolidated subgroup were included in these analyses.

Table 22. Descriptive Statistics for Schools for the Percent of Consolidated Subgroup Current Year Scores that were Proficient and Above.

School Year	*n* of Schools	Mean	Standard Deviation
2010-11	211	48.8%	14.3%
2011-12	194	41.5%	14.6%
2012-13	193	33.4%	13.1%

The findings presented in Table 22 show considerable within school year variation in school scores on this equity measure for all school years studied. The standard deviations ranged from 13.1 to 14.3 which compares with standard deviations on the mean standardized score equity indicator studied above that were from 5.6 to 6.8. Higher standard deviations means there will be fewer schools at each score point so there could be more nuance applied when in establishing cut-points. This is helpful.

There was also considerable variation in school equity scores across school years. This finding is problematic. Specifically, the school mean percent proficient for the consolidated subgroup in 2010-11 was 50% of a standard deviation higher than that from 2011-12 which was 58% of a standard deviation higher than that from 2012-13. Over two years, from 2010-11 to 2012-13, the school mean percent proficient for the consolidated subgroup differed by 112% of a standard deviation.

Next, Table 23 shows cross-year comparisons in the impact associated with school equity scores based upon the current year percent proficient metric. For this simulation, the cut-points were established at the 30th and 70th percentile ranks for the 2010-11 school equity score. These cut-points were then applied for all three school years. The actual cut-points used were 42% for the cut-point between below target and meets target and 55% for the cut-point between meets target and exceeds target.

Table 23. Percent of Schools in Each Equity Category based upon Consolidated Subgroup Current Year Percent Proficient and Above School Equity Scores.

School Year	*n* of Schools	Below Target	Meets Target	Exceeds Target
2010-11	211	29.2%	36.6%	34.1%
2011-12	194	49.8%	25.9%	24.3%
2012-13	193	72.3%	15.5%	12.2%

The findings presented in Table 23 show dramatic shifts in category membership across school years when cut-points established based upon impact data from one school year were applied to place schools into target categories in other school years. For example, the percent of schools in the below target category increased from 29% in the year the cut-points were established to 72% two years later. There were 43% more below target schools two years after the cut-points were established.

Conclusion

Two potential school equity scores were discussed in this addendum. One was the mean student standardized score for the consolidated subgroup and the other was the current year percent of proficient scores in reading and math for the consolidated subgroup. Both of these school equity scores had unstable impact across school years when cut-points established in one year were applied to other years.

Given the poor performance of the two equity scores addressed in the addendum, a long term solution to development of an equity score for use in high schools would involve ensuring that the high school assessment system in Wyoming permits the use of an achievement growth measure. The publishers of the assessments that are presently being used in grades 9 and 10 in Wyoming are being discontinued by the publisher. This presents an opportunity to select assessments for Wyoming high schools that lend themselves to the measurement of growth. Once a growth measure is available, the conditional growth approach described above the addendum could be used to compute high school equity scores.

A correlation study that is presently underway suggests that the current high school performance levels are largely determined by achievement status. The same was not true for schools with grades 3-8. The correlation study results for schools with grades 3-8 were similar to those in Table 21 above that included simulated indicators for the 3-8 schools. Having a growth indicator and a conditional growth metric for equity for high schools would result in performance levels that were based upon richer and more multidimensional evidence of school quality.

Since use of a growth measure is not available for high schools for the 2014-15 school year, however, an alternative approach is still needed. One option is to use the current year percent of proficient and above test scores for the consolidated subgroup. This would be preferable to using a mean student standardized score for consolidated subgroup students. This preference is based upon simplicity and understandability of the score and amount of variance in the school scores. The percent proficient and advanced is an easily understood and widely used metric in education today. In addition, this score, when compared with the mean student standardized score metric, produced more within school year variation in school scores.

The use of this percent proficient equity measure requires that steps be taken to deal with the lack of cross year stability associated with the metric. There are a couple of possible approaches for addressing this problem. One option would be for a standard setting to identify cut-points to occur each year this metric is used. This would guard against unreasonable impact. In the long term, however, annual standard setting is not a practical.

An alternative to annual standard setting was suggested by Marion and Domaleski (2012)[3]. The alternative would have a standard setting one year where the standard set would be the percentage of school expected to be in the below target category and the percentage of schools expected to be in the exceeds target category. For example, the standard setting panel might recommend that schools in the bottom 30% of schools on this equity metric would be in the below target category and that schools in the top 20% of schools on this equity metric would be in the exceeds target category. The remaining 50% of schools would be in the meets target category. Once these percentages were established by standard setting they could be used in subsequent years. This approach would ensure that the lowest ranked schools were in the below target category each year and that the highest ranked schools were within the exceeds target category each year. The drawback of this approach, however, is that it is zero sum; as one school moves up another school would by definition be required to move down. In reality, however, the annual standard setting approach would also likely have a zero sum feature.

Finally, it is possible that some other unstudied, non growth measure could be identified that would have the needed cross year stability for use as an equity indicator.

[3] see page 43 of Marion, S & Domaleski, C. (2012). *The Wyoming comprehensive accountability framework: Phase I.* Produced for the Wyoming Select Committee on Statewide Education Accountability.