CORRELATIONS AMONG SCHOOL PERFORMANCE LEVELS AND INDICATORS

Prepared for the Wyoming Department of Education by Michael Flicek, Ed.D.

(Draft -- 02/26/2015)

Indicators of the Wyoming school accountability model were selected to represent unique aspects of the construct of school quality. Ideally each indicator would represent a related but not identical aspect of the construct of school quality. When multiple indicators with these characteristics are aggregated to derive an overall school quality score, that score is intended to be more robust that any of the indicator scores alone. Additionally, if school size (i.e., student enrollment) is unrelated to indicator scores or target levels, this is evidence the accountability model is not biased for or against schools based upon their size.

In 2014, Wyoming had two school accountability models. There was a model for schools serving grades 3 through 8 and another model for high schools. The grade 3 through 8 model had three indicators. The state test was administered in grades 3 through 8 and each of the three school accountability indicators for schools serving these grades were based upon state test scores. There were differences among the indicators, however, in terms of which test scores were used (i.e., prior year versus current year scores), how the test scores were used, the samples of students included on the indicator and whether a static score or a change in scores over time was of interest. Specifically, the achievement indicator was based upon current year tests scores only and the sample includes all full academic year students tested. The equity indicator, in contrast, used prior year test scores to identify a sample of students with below proficient prior year test scores and then computed the average current year test score of that sample. Finally, growth scores were modeled to represent a change in test scores from prior years to the current year for students in the same grade with a similar history of test scores. High schools had a different pattern of indicators and sub-indicators which will be described in a high school section below.

On most indicators, schools have a score on a continuous scale. Cut-points were identified on these continuous scales in order to establish target levels. There were three target levels on each indicator. The target levels were below target, meets target and exceeds target. Indicator target levels were entered into a decision table to derive a SPL. The four SPLs were below expectations, partially meets expectations, meets expectations and exceeds expectations. In order to study the relationship among the indicators, school performance levels (SPLs) and enrollment, correlation coefficients were computed. Whenever at least one of the variables in a computation was a categorical variable (i.e., either a target level or the SPL) a Spearman correlation coefficient was computed. When both variables were continuous, Pearson correlation coefficients were computed.

Sample

The level of analysis was the school. Only schools that met the minimum *n* requirement of 10 on the indicators under study were included in these analyses.

Measures

The continuous score for achievement was the percent of test events that were proficient for all subjects area tests combined. The continuous score for equity was the mean standardized test score in reading and math combined for students included in the consolidated subgroup (i.e., students with a below proficient test score in reading and/or math on the prior year's state test). The growth measure

was the median student growth percentile (MGP) for reading and math combined. The indicator target levels were assigned scores from 1 for below target to 3 for exceeds target. The SPLs were assigned scores from 1 for not meeting expectations to 4 for exceeding expectations.

Findings

Schools Serving Students in Grades 3-8

Table 1 presents the correlation coefficients for grade 3 through 8 school target levels with one another and with the SPL and enrollment. The correlation coefficients presented in Table 1 show indicator target levels that were unrelated to enrollment ($p > .05$) and moderately and significantly related to one another ($p < .001$). These findings suggest the indicators were related but not identical, as expected and desired. That indicators were unrelated to enrollment was also expected and desired. The SPL was the score that represented the construct of school quality. Since it was the indicator scores that were entered into the decision table in order to derive the SPLs, a strong association of indicator target levels with the SPLs were, by definition, also expected. This expectation was confirmed by the coefficients in Table 1.

Table 1. Relationships of Grade 3 through 8 Indicator Target Levels with Each Other and with the Overall School Performance Level and with Enrollment.

	Value	Equity	Growth	SPL*	Enrollment
Achievement	*r*	0.49***	0.42***	0.88***	-0.04
	n	240	265	265	265
Equity	*r*		0.53***	0.66***	-0.02
	n		240	240	240
Growth	*r*			0.69***	0.06
	n			265	265

Note. All correlation coefficients are Spearman.
*$p < .05$. **$p < .01$. ***$p < .001$.

Next, Table 2 presents correlation coefficients for the continuous scores on the indicators with one another and with the SPLs and enrollment. The findings in Table 2 mirror those from Table 1 suggesting

Table 2. Relationships of Grade 3 through 8 Indicator Measures with Each Other and with Enrollment and the Overall School Performance Level.

	Value	Equity	Growth	SPL*	Enrollment
Achievement	*r*	0.68***	0.52***	0.85***	-0.04
	n	240	265	265	272
Equity	*r*		0.58***	0.71***	-0.07
	n		240	240	240
Growth	*r*			0.66***	-0.04
	n			265	265
SPL	*r*				0.06
	n				265

Note. Correlation coefficients were Pearson except for those with the SPL, which were Spearman.

*p < .05. **p < .01. ***p < .001.

that the relationships among the indicators were similar when the continuous scale indicator scores were studied as when target levels were studied. The findings presented in Tables 1 and 2 suggest that 2014 Wyoming school accountability model for grades 3 through 8 had indicators that relate to one another but were not identical and that relate to the SPL and enrollment as expected. As such, school quality was defined in this model as being comprised of achievement, growth and equity measured as defined based upon state test results. Given this definition, the findings presented here suggest the grades 3 through 8 school accountability model was working as expected in the identification of school quality.

High Schools, All

High school SPLs were based upon target level designations for academic performance and overall readiness. The three target levels (i.e., exceeds, meets and below target) on academic performance and overall readiness were entered into a 3-by-3 decision table to determine the SPL of a high school. Table 3 presents the Spearman correlation coefficients for school academic performance and overall readiness with one another and with the SPL and enrollment.

Table 3. The Relationships of Overall Achievement and Overall Readiness with One Another and with the SPL and with Enrollment.

	Value	Overall Readiness	SPL	Enrollment
Overall Achievement	*r*	0.50***	0.91***	0.37***
	n	80	80	80
Overall Readiness	*r*		0.75***	0.10
	n		80	80

Note. All correlation coefficients are Spearman.

*p < .05. **p < .01. ***p < .001.

The coefficient for overall readiness with overall achievement of r = 0.50 indicates that two related but not identical aspects of school performance were being measured by these variables. The SPL was derived from overall achievement and overall readiness. As such, the strong association of these variables with the SPL was expected and confirmed by the analyses. Overall readiness was unrelated to enrollment but overall achievement was positively and significantly (i.e., p < .001) related to enrollment. Ideally, student enrollment would be unrelated to both overall achievement and overall readiness.

Of the 80 high schools with SPLs, 13 were alternative schools. The 80 high schools were dummy coded with an alternative school designation equal to 1 and a not alternative school designation equal to 0. The alternative schools had significantly low enrollment (i.e., point biserial r = -0.18, p < .001) and significantly low percent proficient and above (i.e., point biserial r = -0.65, p < .001). As such, the significantly high relationship of overall achievement with enrollment, in part, reflects the findings that alternative high schools have significantly low enrollment and significantly low performance on the tests of student achievement.

The academic performance target level was determined by entering target levels for achievement and for equity into a 3-by-3 decision table. The overall readiness target level was determined by entering the target levels for graduation and for additional readiness into a 3-by-3 decision table. Graduation included here was graduation after an improvement measure was applied. Ten schools were increased

one target level because their four year, on-time graduation rate in the current year had increased from the prior year to a point that was one third of the way closer to the higher target. Table 4 presents the correlation coefficients for school target level categories on achievement, equity, graduation and additional readiness with one another and with the overall SPL and with student enrollment at the schools.

Table 4. Relationships of High School Indicator Target Levels with Each Other and with the Overall SPL and Enrollment.

	Value	Equity	Graduation	Additional Readiness	SPL	Student Enrollment
Achievement	*r*	0.46**	0.34**	0.63***	0.87***	0.36**
	n	42	80	80	80	80
Equity	*r*		0.34*	0.37*	0.65***	0.11
	n		42	42	42	42
Graduation	*r*			0.55***	0.56***	0.05
	n			82	80	83
Additional Readiness	*r*				0.81***	0.19
	n				80	83
SPL	*r*					0.32*
	n					80

Note. All correlation coefficients are Spearman.
*$p < .05$. **$p < .01$. ***$p < .001$.

Findings presented in Table 4 show statistically significant ($p < .05$ to $p < .001$) moderate relationships among target levels on the four indicators (i.e., the coefficients range from $r = 0.34$ to $r = 0.63$). Furthermore, the four variables were significantly ($p < .001$) and positively related to the SPLs (i.e., the coefficients range from $r = 0.56$ to $r = 0.87$). Finally, student enrollment was positively and significantly related to both achievement and the SPL (i.e., $p > .01$ and .05 respectively). Ideally, school size as represented by student enrollment would not be associated with any of the indicators or with the overall SPL. As discussed above, the performance of alternative schools contributed to the positive relationship of enrollment with achievement and with the SPL.

Next, Table 5 presents the correlation coefficients for the continuous scale scores on the high school indicators. Descriptions of the continuous scale scores included in the Table 5 analyses follow. The school score on achievement was the percent of subject area test scores on the grade 11 ACT that were proficient or better. The school score on equity was the mean reading and math subject areas test scores on the grade 11 ACT (on the Wyoming ACT scale) for students in the consolidated subgroup. Students with low reading and math scores in grade 10 were included in the consolidated subgroup. The school graduation score used for Table 5 analyses was the four year on-time graduation rate. This graduation score did not have the improvement feature of the score used in Table 4. The school additional readiness score was a weighted composite score composed of three additional readiness sub-indicators. Details of the additional readiness sub-indicators are provided below.

The coefficients in Table 5 were confirmatory of those in Table 4 in that they supported many of the conclusions that were supported by the Table 4 findings. Achievement and additional readiness were both positively related to enrollment (i.e., $p < .05$ and .01). Achievement was based entirely upon ACT

test results and the ACT, Plan and Explore results played a large role in the additional readiness variable. Schools with better test results tend to be larger schools. Again, alternative high schools tended to be smaller schools and to have lower performance on the testing variables.

Table 5. Relationships of High School Continuous Variable Scores with Each Other and with the Overall SPL and Enrollment.

	Value	Equity	4 Year Grad Rate	Additional Readiness	SPL	Enrollment
Achievement	*r*	0.72***	0.67***	0.78***	0.89***	0.26*
	n	42	80	80	80	80
Equity	*r*		0.43**	0.56***	0.64***	0.13
	n		42	42	42	42
4 Year Grad Rate	*r*			0.82***	0.68***	0.16
	n			82	80	83
Additional Readiness	*r*				0.78***	0.30**
	n				80	83

Note. All correlation coefficients are Pearson except for those with the SPL, which were Spearman.
*p < .05. **p < .01. ***p < .001.

Finally, Table 6 explores the relationships of the additional readiness sub-indicators with one another and with the overall additional readiness score and with enrollment. On tested readiness each student received an index score based upon their current year composite score on the grade 9 Explore, grade 10 Plan or grade 11 ACT. The school score was the average student index score across all three tested grades. The school grade 9 credit score was the percent of first year grade 9 students who earned one fourth of the credits required for graduation from the high school. Each graduate from a school received a Hathaway scholarship index score based upon their eligibility level for the scholarship. The school score for Hathaway eligibility was the mean index score. Table 6 presents correlation coefficients based upon the continuous scales of the four variables.

Table 6. Relationships of High School Additional Readiness Sub-indicator Scores with the Additional Readiness Score and Enrollment for All High Schools.

	Value	Grade 9 Credits	Hathaway Eligibility	Additional Readiness	Enrollment
Tested Readiness	*r*	0.38**	0.83***	0.92***	0.26*
	n	62	71	83	83
Grade 9 Credits	*r*		0.30*	0.76***	-0.02
	n		59	62	62
Hathaway Eligibility	*r*			0.86***	0.19
	n			71	71

Note. Correlation coefficients were Pearson.
*p < .05. **p < .01. ***p < .001.

The findings presented in Table 6 indicated the additional readiness sub-indicators were significantly related to one another and to the overall additional readiness score. Tested readiness was significantly

related to enrollment (i.e., $p < .05$) while grade 9 credits and Hathaway eligibility were not significantly related to enrollment.

High Schools, Alternative Schools Excluded

During the 2013-14 school year, 15 of 84 high schools were alternative high schools. Four of the 84 high schools were small schools that did not have enough data to be awarded a school performance level. Two of the 15 alternative schools were small schools that did not receive a school performance level. Each of the 13 alternative high schools that received a school performance level were in the not meeting expectations category. Because all alterative high schools were in the not meeting expectations category, there has been considerable conversation in Wyoming about the usefulness of the current high school accountability model for alternative high schools. This section considers the performance of Wyoming high schools when the alternative high schools were excluded from the analyses.

Table 7 presents the Spearman correlation coefficients for school academic performance and overall readiness with one another and with the SPL and enrollment for all non-alternative high schools in Wyoming. Table 7 and Table 3 addressed the same high school variables. Data from alternative high schools were included for the computations presented in Table 3 and were not included for the computations presented in Table 7. Overall achievement was positively and significantly related to enrollment when alternative high schools were included in the computations but overall achievement was not significantly related to enrollment when the alternative high school data was excluded. This is the most noteworthy difference from a comparison of Table 3 with Table 7. The evidence of bias on overall achievement related to school size that was present when alternative schools were included was not present when alternative schools were excluded.

Table 7. The Relationship of Overall Achievement and Overall Readiness with One Another and with the SPL and with Enrollment for Non-Alternative High Schools.

	Value	Overall Readiness	SPL	Enrollment
Overall Achievement	*r*	0.31*	0.90***	0.17
	n	67	67	67
Overall Readiness	*r*		0.58***	-0.23
	n		67	69

Note. All correlation coefficients are Spearman.
*$p < .05$. **$p < .01$. ***$p < .001$.

Next, Table 8 examines the relationship among target levels on the four variables that were combined to derive overall achievement and overall readiness. The relationship of these variables to the SPL and to student enrollment were also examined. Tables 4 and 8 address the same high school variables for Wyoming high schools. Table 4 includes all high schools and Table 8 includes all non-alternative high schools. When alternative schools were excluded, the results in Table 8 show the SPL to be unrelated to student enrollment. This was a desired result showing the model was not biased based upon school size when the alternative schools were excluded.

An unexpected finding was the statistically significant (i.e., $p < .05$) negative relationship of graduation to student enrollment. Low enrollment was associated with a higher graduation target level. The graduation target level, therefore, was biased in favor of small schools over larger schools. The

graduation target level was also unrelated to the school achievement target level (i.e., $p > .05$). This was another unexpected finding. These findings were likely related to the graduation target level approach

Table 8. Relationships of High School Indicator Target Levels with Each Other and with the Overall SPL and Enrollment for Non-Alternative High Schools.

	Value	Equity	Graduation	Additional Readiness	SPL	Student Enrollment
Achievement	*r*	0.43**	0.17	0.52***	0.84***	0.17
	n	41	67	67	67	67
Equity	*r*		0.31*	0.35*	0.64***	0.05
	n		41	41	41	41
Graduation	*r*			0.41***	0.42***	-0.25*
	n			68	67	68
Additional Readiness	*r*				0.71***	-0.12
	n				67	69
SPL	*r*					0.06
	n					67

Note. All correlation coefficients are Spearman.
*$p < .05$. **$p < .01$. ***$p < .001$.

which allowed 10 of the schools that showed four year, on-time graduation rate improvement from the prior school year to be moved up one target level. That it was the improvement feature of this graduation target level that seemed to account for these unexpected findings was supported by findings presented in Table 9 below which shows correlation coefficients for the continuous scale versions of the school accountability indicator scores. The graduation variable included in Table 9 is the four year, on-time graduation rate. This rate had a strong positive relationship ($p < .001$) with the achievement indicator (i.e., percent of proficient test scores) and was unrelated to enrollment ($p > .05$).

Table 9. Relationships of High School Variable Scores with Each Other and with the Overall SPL and Enrollment when Alternative Schools were Excluded.

	Value	Equity	4 Year Grad Rate	Additional Readiness	SPL	Enrollment
Achievement	*r*	0.62***	0.50***	0.63***	0.86***	0.11
	n	41	67	67	67	67
Equity	*r*		-0.15	0.36*	0.62***	0.06
	n		41	41	41	41
4 Year Grad Rate	*r*			0.73***	0.51***	-0.04
	n			68	67	68
Additional Readiness	*r*				0.64***	0.14
	n				67	69

Note. All correlation coefficients are Pearson except for those with the SPL, which were Spearman.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 9 shows each of the four indicator continuous scale scores were unrelated to student enrollment ($p > .05$). Again there was a difference in the relationship to enrollment when alternative schools were excluded from the sample. Both achievement and additional readiness were positively and significantly related to enrollment when alternative schools were included and were unrelated when alternative schools were excluded. Table 9 also shows that, as expected, the four indicator continuous scale scores were significantly related to the SPL ($p < .001$). Additionally, the four variables were moderately and significantly related to one another (i.e., at least $p < .05$) with just one exception. The exception was that equity was unrelated to the four year, on-time graduation rate.

Finally, Table 10 explores the relationships of the additional readiness sub-indicators with one another and with the overall additional readiness score and with enrollment for the sample of schools with alternative schools excluded. When alternative schools were included as presented in Table 6 tested readiness was significantly and positively related to enrollment ($p < .05$). When alternative schools were excluded, however, as presented in Table 10, all three readiness sub-indicators were unrelated to enrollment ($p > .05$). The three sub-indicators were positively and significantly related to the overall additional readiness score ($p < .001$). Lastly, tested readiness was related to grade 9 credits and Hathaway eligibility. Hathaway eligibility and grade 9 credits were not significantly related to one another.

Table 10. Relationships of High School Additional Readiness Sub-indicator Measures with the Additional Readiness Measure and Enrollment All Non Alternative High Schools.

	Value	Grade 9 Credits	Hathaway Eligibility	Additional Readiness	Enrollment
Tested Readiness	*r*	0.31*	0.70***	0.87***	0.08
	n	61	61	69	69
Grade 9 Credits	*r*		0.23	0.74***	-0.04
	n		58	61	61
Hathaway Eligibility	*r*			0.74***	0.02
	n			61	61

Note. Correlation coefficients were Pearson.
*$p < .05$. **$p < .01$. ***$p < .001$.

The definition of school quality for high schools that is defined by the 2014 high school accountability model is more complicated than the model for schools serving grades 3 through 8. The model is of questionable utility as an accountability model for alternative high schools. It places them all into the single, lowest SPL. This may well mask differences in quality among the alternative schools that could be useful for motivating and informing improved practices. The development and implementation of a different model for alternative school accountability may be useful if the goal of motivating and informing improved practices at these schools is to be achieved. With the exception of the graduation indicator which included the improvement feature, the relationships among the indicators, the SPLs and enrollment when alternative schools were removed from the sample of high schools were generally as expected and provided evidence the model worked as expected.

Conclusion

Wyoming had two school accountability models in place in 2014. One for schools serving grades 3 through 8 and one for high schools. SPLs for both accountability models represent particular definitions of the construct of school quality. For schools serving grades 3 through 8, that definition was based upon the indicators for achievement, growth and equity. Given that definition, the analyses presented above suggest the three indicators relate to one another and to the SPL as expected. The SPL further related to enrollment as expected. In summary, the grade 3 through 8 model was working as expected.

The 2014 Wyoming high school accountability model represented a more complicated model of the construct of school quality. Achievement and equity in the high school model were based upon the ACT subject area tests scores but beyond that they were very similar to the achievement and equity indicators in the model for schools serving grades 3 through 8. The high school model did not have a growth indicator. Instead, the high school model had four readiness sub-indicators. The sub-indicators were graduation, tested readiness, eligibility level for Hathaway scholarship and grade 9 credits earned. When the alternative high schools were not included in the sample, given the definition of school quality represented by the indicators in the model, the high school model generally worked as expected. The one exception was on the graduation indicator target level which included an improvement feature. There was some evidence that the target level on this indicator was biased in favor of small schools over larger schools. In contrast, the four year, on-time graduation rate was unrelated to enrollment and, as such, it did not show a bias associated with school size. The extended graduation rate was not included in this study but it differed very little from the four year, on-time graduation rate. Basing graduation target levels on either the four year, on-time graduation rate or the extended graduation rate, without applying the improvement feature, would very likely produce graduation indicator unbiased by school size.

When alternative schools were included in the sample of high schools, the findings presented here suggested the model was biased against small schools. In reality the model is a bad fit for alternative schools which had significantly low test performance and significantly low student enrollment. The accountability model placed all alternative schools into the lowest category. As a result, the alternative schools were given little information by the model to inform or motivate school improvement. A different conceptualization of school quality, as represented by different target levels and/or indicators, could be developed for the alternative high schools with the intention of creating an accountability model for them that would both inform and motivate school improvement.

Finally, an indicator like the grade 3 through 8 growth indicator should be considered for all high schools, but would likely be especially helpful for alternative high schools. This would also permit the use of a better, growth based definition for equity (see Flicek, 2015[1]). An assessment system that includes assessments with characteristics that lend themselves to the measurement of growth would be needed in order for growth to be added to the high school model.

[1] Flicek, M. (2015). *Stability of 2014 Wyoming School Accountability Indicators Across Years.* Prepared for t he Wyoming Department of Education.