Predictability Study of ISIP Reading and STAAR Reading: Prediction Bands

March 2014

Chalie Patarapichayatham[1], Ph.D.

William Fahle[2], Ph.D.

Tracey R. Roden[3], M.Ed.

---

[1] Research Assistant Professor in the Department of Education Policy and Leadership,
Simmons School of Education and Human Development, Southern Methodist University

[2] Senior Vice President of Development Research at Istation

[3] Vice President of Curriculum at Istation

Corresponding Author:

Chalie Patarapichayatham, 3101 Southern Methodist University, Dallas, TX 75205

Email: cpatarapichy@smu.edu

Abstract

This study provides evidence of the ability of Istation's Indicators of Progress (ISIP) Reading cut score to predict passing the State of Texas Assessments of Academic Readiness (STAAR) Reading tests. ISIP Early Reading (Grade 3), ISIP Advanced Reading (Grades 4 – 8), and STAAR Reading tests are used in this study. All data came from the Garland Independent School District, Garland, Texas, 2012-2013 school year, totaling 20,493 students. There were 3,694 third graders, 3,783 fourth graders, 3,877 fifth graders, 3,519 sixth graders, 2,973 seventh graders, and 2,647 eighth graders. Simple linear regression analysis was applied for Grades 3 – 8 data. Predictability "bands" were further computed by using the confidence interval (CI) approach to obtain the ISIP cut score to predict passing STAAR tests. In particular, we seek a range of scores on ISIP where we can say, for example, students with an ISIP score from range x to y will have probability z of passing the STAAR test. Results of a simple linear regression showed that the ISIP Reading measures are highly predictive of STAAR scores. The ISIP cut scores to predict passing STAAR Level II Phase in I were 229, 1,798, 1,844, 1,920, 1,950, and 1,949 for Grades 3 – 8, respectively. For next academic year, the ISIP cut scores will be 243, 1,898, 1,953, 2,036, 2,089, and 2,090, for Grades 3 – 8, respectively in order to pass STAAR Level II Phase in II cut score.

Introduction

Istation's Indicators of Progress, also known as ISIP, is a sophisticated Internet- and web-delivered Computer-Adaptive Testing (CAT) system that provides Continuous- Progress Monitoring (CPM) assessment in the critical domains of reading in grades Pre K – 8th. Built from Item-Response Theory (IRT) and driven by a CAT algorithm, ISIP gathers and reports frequent information about student progress in these critical domains throughout, and even across, academic years (e.g., Patarapichayatham, Fahle, & Roden, 2013).

The purpose of ISIP is to measure reading ability and identify deficits in the critical areas to provide continuous, differentiated instruction. ISIP accomplishes this by delivering short tests, at least monthly, that target critical areas to inform instruction. It is realistic to administer ISIP assessments for an entire classroom, an entire school, and even an entire district in a single day with adequate computer resources. Student results are immediately available online to teachers and administrators, illustrating each student's past and present performance and skill growth. Teachers are alerted when students are not making adequate progress so that the instructional program can be modified before a pattern of failure becomes established (e.g., Mathes, 2011). See Mathes (2011) and Mathes, Torgesen and Herron (2011) for full information about ISIP Early Reading (for Pre K – Grade 3) and ISIP Advanced Reading (for Grades 4 – 8).

The State of Texas Assessments of Academic Readiness (STAAR) replaced the Texas Assessment of Knowledge and Skills (TAKS) in spring 2012. STAAR is a rigorous testing program that includes annual assessments in reading and mathematics for Grades 3 – 8. STAAR emphasizes "readiness standards" in knowledge and skills most important for college- and career-readiness. One of the state's goals in developing STAAR is that Texas will be among the top 10 states for graduating college-ready students by the 2019 – 2020 school year.

ISIP Early Reading, developed in 2006, and ISIP Advanced Reading, developed in 2010, have been delivered to more than 3,496,345 students in more than 37 states and 6 countries. A majority of these students are from the state of Texas and have been required to take the STAAR reading assessment. Patarapichayatham, Fahle, and Roden (2013) studied the relationship between ISIP Reading and STAAR Reading by applying Pearson Product-Moment correlation analysis, multiple linear regression analysis, and multiple logistic regression for Grades $3 - 8$ data. They found that the ISIP end of the year (EOY) scores were higher than the ISIP middle of the year (MOY) scores on both overall scores and each sub-skill score across grades, indicating that students improved their reading ability by the end of the year. The correlations between ISIP Reading and STAAR Reading tests were very strong across grades, indicating that students who perform well on ISIP Reading are very likely to perform very well on STAAR. They also found that ISIP Reading measures are highly predictive of STAAR scores.

Although Patarapichayatham, Fahle, and Roden (2013) found very strong correlations between ISIP Reading and STAAR Reading tests across grades, and ISIP Reading measures are highly predictive of STAAR scores, finding the ISIP cut score which determines a high probability of passing STAAR was outside the scope of their study. For this reason, this study aims to propose the ISIP cut score for this purpose. The result of simple linear regression was further used to obtain the multiple prediction bands by using the confidence interval (CI) approach for the predictor. The Middle of the Year Overall Reading (MOY_Overall) was used as the predictor and the STAAR scale score was the outcome variable in this current study. The samples were taken from students in the Garland Independent School District (GISD), for the $2012 - 2013$ school year.

Prediction bands commonly arise in regression analysis. The goal of a prediction band is to cover with a prescribed probability the values of one or more future observations from the same population from which a given data set was sampled. There are two types of prediction bands: confidence interval (CI) and prediction interval (PI). A CI is used in statistical analysis to represent the uncertainty in an estimate of a curve or function of the data. Similarly, a PI is used to represent the uncertainty about the value of a new data point on the curve, but subject to noise. The 95% confidence intervals enclose the area that we can be 95% sure contains the true curve. If we have many data points, the confidence intervals will be near the line or curve, and most of our data will lie outside the confidence intervals. The 95% prediction intervals enclose the area that we expect to enclose 95% of future data points. They are wider than confidence bands and they are much wider with large data sets.

Methods

Measures

Results from ISIP Early Reading for Grade 3, ISIP Advanced Reading for Grades 4 – 8, and STAAR Reading tests for Grades 3 – 8 for the 2012 – 2013 school year were used in this study. ISIP Reading at the Middle of the Year (MOY) of 2012 – 2013 academic year and the STAAR Reading scores from 2012 – 2013 for these same students were used. Four skills are measured under Grade 3 ISIP Early Reading and Grades 4 – 8 ISIP Advanced Reading tests: (a) Comprehension (CMP), (b) Spelling (SPL), (c) Vocabulary (VOC), and (d) Connected Text Fluency (TF). Three reporting categories (e.g., sub-skills) are measured under Grades 3 – 8 STAAR Reading tests. They are (a) Understanding across genres (RC1), (b) Understanding/analysis of literary texts (RC2), and (c) Understanding/analysis of informational texts (RC3). The scale scores are reported for both ISIP and STAAR tests.

Samples

       This sample consisted of students in Grades 3 – 8 in the Garland Independent School District, Garland, Texas, totaling 20,493 students. There were 3,694 third graders, 3,783 fourth graders, 3,877 fifth graders, 3,519 sixth graders, 2,973 seventh graders, and 2,647 eighth graders in this current study. The means of ISIP scores were 250.817, 1,959.211, 2,037.081, 2,080.394, 2,130.635, and 2,209.772 for Grades 3 – 8, respectively. The means of STAAR scores were 1,447.780, 1,526.860, 1,569.100, 1,593.200, 1,634.720, and 1,691.280 for Grades 3 – 8, respectively.

Analysis

       The purpose of this study is to provide the ISIP cut score to predict passing STAAR Reading scores, for Grades 3 – 8. Patarapichayatham, Fahle, and Roden (2013) found that the correlations between the Middle of the Year Overall Reading (MOY_Overall) and the End of the Year Overall Reading (EOY_Overall) were very high. They were .886, .876, .867, .842, .845, and .862 for Grades 3 – 8, respectively. For this reason, we decided to use only the MOY_Overall in this current study. The simple linear regression analysis was applied for each grade data by using SPSS software. Thus, the MOY_Overall was the predictor and the STAAR score was the outcome variable.

       We first ran the simple linear regression. The $\hat{y}$ for each grade data was computed (see Part I result). Then, the confidence interval (CI) and the prediction interval (PI) were further computed. The CI was chosen because it obtained reasonable results than the PI. Both CI and PI are centered at $\hat{y}$, the PI is much wider than the CI, for a given x^* and confidence level. The reason is the PI takes account of the tendency of y to fluctuate from its mean whereas the CI

needs to account for the uncertainty in estimating the mean. The results from the CI are reported in this study.

The confidence interval (CI) for average expected value of y for a given x^* is as follow:

$$E(y \mid x^*) = \hat{y} \pm t^*_{n-2} \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{(n-1)s_x^2}}$$

where s_y is the standard deviation of the residuals, calculated as $s_y = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-2}}$. Next, the ISIP scores around the STAAR cut point for the lower bound and the upper bound of the CI were selected. For each grade, the STAAR Phase in I Level II performance standards were used for 2013 – 2014 school year cut point. The STAAR Phase in II Level II performance standards are used for 2014 – 2015 school year cut point. The STAAR Phase in I Level II performance standards for Grades 3 – 8 are 1,331, 1,422, 1,458, 1,504, 1,556, and 1,575, respectively. The STAAR Phase in II Level II performance standards for Grades 3 – 8 are 1,400, 1,486, 1,520, 1,567, 1,615, and 1,637, respectively.

Results

Part I: Simple Linear Regression

Simple linear regression was applied for each grade's data. One predictor was studied under simple linear regression: the Middle of the Year Overall Reading (MOY_Overall). The STAAR score was the outcome variable. The results of each grade are reported as follows.

Grade 3 results are shown in Table 3. As indicated, 52.3% of the variance in STAAR can be predicted from a MOY_Overall. The equation for predicting the STAAR score is as follows:

STAAR = 1,447.780 + 5.013 (MOY_Overall) + e

The intercept was 1,447.780. The STAAR score was 1,448 as MOY_Overall was zero. The slope for MOY_Overall was 5.013 indicating that the STAAR score was 1,453 (1,448 + 5) as MOY_Overall score increases 1 unit.

Grade 4 results are shown in Table 4. In this case, 54.8% of the variance in STAAR can be predicted from a MOY_Overall score. The equation for predicting the STAAR score is as follows:

$$STAAR = 1,526.858 + 0.658 \, (MOY\_Overall) + e$$

The intercept was 1,526.858. The STAAR score was 1,527 as MOY_Overall was zero. The slope for MOY_Overall was 0. 658 indicating that the STAAR score was 1,528 (1,527 + 1) as MOY_Overall score increases 1 unit.

Grade 5 results are shown in Table 5. Here, 51.1% of the variance in STAAR can be predicted from a MOY_Overall score. The equation for predicting the STAAR is as follows:

$$STAAR = 1,569.100 + 0.551 \, (MOY\_Overall) + e$$

The intercept was 1,569.100. The STAAR score was 1,569 as MOY_Overall was zero. The slope for MOY_Overall was 0.551 indicating that the STAAR score was 1,570 (1,569 + 1) as MOY_Overall score increases 1 unit.

Grade 6 results are shown in Table 6. For grade 6, 53.1% of the variance in STAAR can be predicted from a MOY_Overall. The equation for predicting the STAAR score is as follows:

$$STAAR = 1,593.198 + 0.530 \, (MOY\_Overall) + e$$

The intercept was 1,593.198. The STAAR score was 1,593 as MOY_Overall was zero. The slope for MOY_Overall was 0.530 indicating that the STAAR score was 1,594 (1,593 + 1) as MOY_Overall score increases 1 unit.

Grade 7 results are shown in Table 7. Here, 50.4% of the variance in STAAR can be predicted from a MOY_Overall. The equation for predicting the STAAR score is as follows:

$$STAAR = 1,634.722 + 0.410\ (MOY\_Overall) + e$$

The intercept was 1,634.722. The STAAR score was 1,635 as MOY_Overall was zero. The slope for MOY_Overall was 0.410 indicating that the STAAR score was 1,636 (1,635 + 1) as MOY_Overall score increases 1 unit.

Grade 8 results are shown in Table 8. In this case 51.6% of the variance in STAAR can be predicted from a MOY_Overall. The equation for predicting the STAAR score is as follows:

$$STAAR = 1,691.279 + 0.423\ (MOY\_Overall) + e$$

The intercept was 1,691.279. The STAAR score was 1,691 as MOY_Overall was zero. The slope for MOY_Overall was 0.423 indicating that the STAAR score was 1,692 (1,691 + 1) as MOY_Overall score increases 1 unit.

In summary, our finding reported that MOY_Overall is a good predictor. The R^2 were relatively high across grades indicating that ISIP Reading measures are predictive of STAAR Reading across grades.

Part II: Prediction Bands

The confidence interval (CI) was applied to obtain the prediction band for each grade from simple linear regression results. The confidence level was set at .95. The results of each grade are reported as follows (see Table 2).

For Grade 3, the prediction band ranges from 226 to 229 for the 2013 – 2014 school year. The ISIP AR score at 229 is the cut score above which we are confident students will pass STAAR. It is 95% confident that a group of students who have ISIP ER score at 229 will be scored higher than STAAR Level II Phase in I cut score at 1,331. In other words, a group of 3rd-

grade students who have scored 229 or higher on ISIP ER will almost certainly pass STAAR. For next academic year (2014 – 2015), by using STAAR Level II Phase in II cut score at 1,400, the prediction band ranges from 241 to 243. The ISIP cut score will be 243 in order to pass STAAR Level II Phase in II cut score.

For Grade 4, the prediction band ranges from 1,782 to 1,798 for the 2013 – 2014 school year. The ISIP AR score at 1,798 is the cut score above which we are confident students will pass STAAR. It is 95% sure that a group of students who have ISIP AR score at 1,798 will be scored higher than STAAR Level II Phase in I cut score at 1,422. In other words, a group of 4th-grade students who have scored 1,798 or higher on ISIP AR will almost certainly pass STAAR. For the next academic year (2014 – 2015), by using STAAR Level II Phase in II cut score at 1,486, the prediction band ranges from 1,888 to 1,898. The ISIP AR cut score will be 1,898 in order to pass STAAR Level II Phase in II cut score.

For Grade 5, the prediction band ranges from 1,828 to 1,844 for the 2013 – 2014 school year. The ISIP AR score at 1,844 is the cut score above which we are confident students will pass STAAR. It is 95% confident that a group of students who have ISIP AR score at 1,844 will be scored higher than STAAR Level II Phase in I cut score at 1,458. In other words, a group of 5th-grade students who have scored 1,844 or higher on ISIP AR will almost certainly pass STAAR. For the next academic year (2014 – 2015), by using STAAR Level II Phase in II cut score at 1,520, the prediction band ranges from 1,943 to 1,953. The ISIP AR cut score will be 1,953 in order to pass STAAR Level II Phase in II cut score.

For Grade 6, the prediction band ranges from 1,904 to 1,920 for the 2013 – 2014 school year. The ISIP AR score at 1,920 is the cut score above which we are confident students will pass STAAR. It is 95% confident that a group of students who have ISIP AR score at 1,920 will

be scored higher than STAAR Level II Phase in I cut score at 1,504. In other words, a group of 6th-grade students who have scored 1,920 or higher on ISIP AR will almost certainly pass STAAR. For the next academic year (2014 – 2015), by using STAAR Level II Phase in II cute score at 1,567, the prediction band ranges from 2,024 to 2,036. The ISIP AR cut score will be 2,036 in order to pass STAAR Level II Phase in II cut score.

For Grade 7, the prediction band ranges from 1,930 to 1,950 for the 2013 – 2014 school year. The ISIP AR score at 1,950 is the cut score above which we are confident students will pass STAAR. It is 95% sure that a group of students who have ISIP AR score at 1,950 will score higher than STAAR Level II Phase in I cut score at 1,556. In other words, a group of 7th-grade students who have scored 1,950 or higher on ISIP AR will almost certainly pass STAAR. For the next academic year (2014 – 2015), by using STAAR Level II Phase in II cut score at 1,615, the prediction band ranges from 2,076 to 2,089. The ISIP AR cut score will be 2,089 in order to pass STAAR Level II Phase in II cut score.

For Grade 8, the prediction band ranges from 1,923 to 1,949 for the 2013 – 2014 school year. The ISIP AR score at 1,949 is the cut score above which we are confident students will pass STAAR. It is 95% confident that a group of students who have ISIP AR score at 1,949 will be scored higher than STAAR Level II Phase in I cut score at 1,575. In other words, a group of 8th-grade students who have scored 1,949 or higher on ISIP AR will pass STAAR. For the next academic year (2014 – 2015), by using STAAR Level II Phase in II cut score at 1,637, the prediction band ranges from 2,073 to 2,090. The ISIP AR cut score will be 2,090 in order to pass STAAR Level II Phase in II cut score.

Conclusions

The results of simple linear regression add to the evidence that ISIP Reading measures are predictive of STAAR Reading across grades. The ISIP tests can be used as a prediction of how a student will score on STAAR. The ISIP cut scores to predict passing STAAR are very useful for teachers, school administrators, and parents to expect how students will perform on STAAR. Since the STAAR performance standards will be changed for the next academic year (2014 – 2015), this study obtained very useful information for schools to be ready for the new STAAR performance standards. Generally speaking, Grade 4 – 8 students need approximately 100 more points on ISIP AR to pass the STAAR Level II Phase in II performance standards.

References

Patarapichayatham, C., Fahle, W., & Roden, T. R. (2013). ISIP reading versus STAAR reading: the predictability study.

Mathes, P. (2011). Technical manual: Istation's indicators of progress, advanced reading: Computer adaptive testing system for continuous progress monitoring of reading growth for students grade 4 to grade 8.

Mathes, P., Torgesen, J., & Herron, J. (2011). Technical manual: Istation's indicators of progress, early reading: computer adaptive testing system for continuous progress monitoring of reading growth for students Pre-K to grade 3.

Table 1

Descriptive Statistics of ISIP and STAAR scale scores

Grade	n	ISIP				STAAR			
		Minimum	Maximum	Mean	SD	Minimum	Maximum	Mean	SD
3	3,694	186.408	368.649	250.817	17.984	1,009	1,909	1,447.780	124.700
4	3,783	1,313.497	2,798.810	1,959.211	149.569	811	1,995	1,526.860	132.834
5	3,877	1,316.886	2,827.037	2,037.081	153.158	1,216	2,021	1,569.100	118.154
6	3,519	1,040.573	2,806.967	2,080.394	185.083	876	2,081	1,593.200	134.567
7	2,973	1,448.819	2,919.626	2,130.635	185.723	1,345	2,141	1,634.72	107.194
8	2,647	1,346.012	2,783.832	2,209.772	191.576	1,337	2,186	1,691.280	112.764

Table 2 ISIP Cut Score to Predict Passing STAAR

Grade	Cut Score	
	STAAR Level II Phase in I	STAAR Level II Phase in II
3	229	243
4	1,798	1,898
5	1,844	1,953
6	1,920	2,036
7	1,950	2,089
8	1,949	2,090

Table 3

Simple Linear Regression Analysis for Grade 3

		Model Summary		
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.723	.523	.522	86.174

	ANOVA					
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	30009558.525	1	30009558.525	4041.173	.000
	Residual	27416613.221	3692	7425.952		
	Total	57426171.747	3693			

	Coefficients					
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	1447.780	1.418		1021.116	.000
	MOY_Overall_Score	5.013	.079	.723	63.570	.000

Table 4

Simple Linear Regression Analysis for Grade 4

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.740	.548	.548	89.288

ANOVA						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	36589153.500	1	36589153.500	4589.482	.000
	Residual	30143614.556	3781	7972.392		
	Total	66732768.056	3782			

Coefficients						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	1526.858	1.452		1051.774	.000
	MOY_Overall_Score	.658	.010	.740	67.746	.000

Table 5

Simple Linear Regression Analysis for Grade 5

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.715	.511	.511	82.635

ANOVA						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	27650111.148	1	27650111.148	4049.219	.000
	Residual	26460455.620	3875	6828.505		
	Total	54110566.769	3876			

Coefficients						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	1569.100	1.327		1182.322	.000
	MOY_overall_Score	.551	.009	.715	63.633	.000

Table 6

Simple Linear Regression Analysis for Grade 6

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.729	.531	.531	92.148

ANOVA						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	33840874.982	1	33840874.982	3985.385	.000
	Residual	29863705.965	3517	8491.244		
	Total	63704580.947	3518			

Coefficients						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	1593.198	1.553		1025.637	.000
	MOY_Overall_Score	.530	.008	.729	63.130	.000

Table 7

Simple Linear Regression Analysis for Grade 7

		Model Summary		
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.710	.504	.504	75.492

			ANOVA			
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	17218377.609	1	17218377.609	3021.277	.000
	Residual	16931844.900	2971	5699.039		
	Total	34150222.509	2972			

		Coefficients				
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	1634.722	1.385		1180.703	.000
	MOY_Overall_Score	.410	.007	.710	54.966	.000

Table 8

Simple Linear Regression Analysis for Grade 8

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.718	.516	.516	78.458

ANOVA						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	17364081.965	1	17364081.965	2820.841	.000
	Residual	16281670.276	2645	6155.641		
	Total	33645752.241	2646			

Coefficients						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	1691.279	1.525		1109.061	.000
	MOY_Overall_Score	.423	.008	.718	53.112	.000