EXAMINATION OF CONSISTENCY ON THE OHIO ACHIEVEMENT ASSESSMENTS AND OHIO GRADUATION TEST

A thesis submitted to
The Graduate College of
Marshall University

In partial fulfillment of
the requirements for a degree of
Education Specialist

in
School Psychology

by
Adam R. Fox

Approved by
Sandra S. Stroebel, Ph.D. Committee Chairperson
Peter N. Prewett, Ph.D.
Stephen L. O'Keefe, Ph.D.

Marshall University
August 2014

Acknowledgements

It gives me great pleasure in acknowledging all of my committee members and professors who encouraged and helped throughout this process. My sincerest gratitude is offered to the late Dr. Fred Jay Krieg for being an excellent teacher and mentor throughout the entire program. I would also deeply like to thank Dr. Sandra Stroebel for her persistent support, encouragement and guidance. An enormous amount of appreciation goes to Dr. Stephen O'Keefe for sharing his vast wealth of knowledge and experiences. Finally, I would like to thank Dr. Peter Prewett for providing valuable feedback and support for completion of this thesis.

I would be remiss if I did not include in my acknowledgements my ever encouraging wife and best friend, Jessica. I would also like to thank all of my friends, family and MUGC cohort for their support and encouragement throughout the entire program.

Table of Contents

List of Tables

ABSTRACT

This study investigated a cohort of students' performance on the Ohio Achievement Assessments (OAA) and Ohio Graduation Tests (OGT). The purpose of this study was to investigate the consistency of a cohort's reported scores on the OAA over a four-year period (5^{th}, 6^{th}, 7^{th} and 8^{th} grades) and reported scores on their OGT assessment; this was accomplished through the examination of OAA and OGT data from a rural school district located in central Ohio. The data were analyzed using correlations, regressions, and repeated measures ANOVA. A Pearson product-moment correlation coefficient was computed indicating positive correlations between all OAA and OGT assessments analyzed. A regression analysis indicated the 8^{th} grade OAA assessment is the most important predictor of the OGT assessment. The ANOVA suggested that differences between mean scores in the OAA and OGT assessments are not statistically significant indicating the scores are consistent with each other.

CHAPTER I

Review of Literature

Academic Achievement

Achievement assessments are commonplace in today's school systems. The need for accountability and measurement of students' academic performance is indeed an important part of our education system.

> If what students learned as a result of the instructional practices of teachers were predictable, then all forms of assessment would be unnecessary; student achievement could be determined simply by students is not related in any simple way to what they have been taught, assessment is a central—perhaps even the central—process in education. At the very least, assessment is integral to effective instruction. (Wiliam, 2010, p.254)

As achievement assessments have begun to take center stage and eclipse the Common Core Standards, teaching practices, and funding decisions throughout our educational system, they are in need of further examination. It is important to ensure that these assessments are, in fact, reliable tools to measure the achievement of students, and are able to withstand the generalizability that accompanies such assessments.

The No Child Left Behind (NCLB) Act of 2001 set forth a trajectory path ensuring that all children would have a fair and equal opportunity to access a high quality education; therefore at least perform proficient on state academic assessments. The act indicated that this could be accomplished through

ensuring the following: high-quality academic assessments, accountability systems, teacher preparation and training, curriculum, and correct alignment of instructional materials with state standards to allow for progress monitoring of academic achievement (No Child Left Behind [NCLB], 2001).

Current research has found mixed results both in support and opposition for the impact of the achievement assessments under the influence of NCLB (Lee & Reeves, 2012). Lee and Reeves' examination of the achievement gap and trends in US students indicate that mixed patterns have emerged when looking at both reading and math scores since NCLB was enacted. Some studies concluded significant positive effects from NCLB while other studies found insignificant mixed patterns with very small changes in student performance (Lee & Reeves, 2012). It can be concluded that individual and group opinions on current achievement assessments are unclear, yet we can agree that the intended purpose is to aide in monitoring students' achievement by obtaining acceptable levels of knowledge on standardized assessments.

Ohio's Assessments

Ohio currently uses the Ohio Achievement Assessment (OAA) and the Ohio Graduation Test (OGT) to comply with the NCLB Act of 2001. The OAA is an annual assessment used to measure students' knowledge of concepts taught in grades 3 through 8. The OGT is an assessment used to measure high school students' levels of achievement and skills by the spring of their 10th grade year. The Ohio Department of Education indicates these assessments are given to measure how well students have learned those concepts pertaining to the

content areas of English language arts, mathematics, and to applicable grades in science, writing and social studies. The OAA and OGT are based on Ohio's Academic Content Standards (Ohio, Department Ohio Education [DOE], n.d.).

The Ohio Graduation Test (OGT) is an assessment established in 2001 based on recommendations by the Governor's Commission for Student Success (Ohio, DOE, n.d.). The OGT assessment is administered to all 10th grade students. Students continue to take the OGT until all sections are passed. The OGT consists of five sections: reading, writing, mathematics, science, and social studies. Students must obtain a proficient score on each individual section in order to receive a high school diploma. It was noted that alternative routes to receive a diploma are an option in specific situations. For example, students in special education can be exempted from the consequences of not passing the OGT (Ohio, DOE, n.d.).

Uses of Ohio's Achievement Assessments

As previously stated, Ohio currently uses the Ohio Achievement Assessment (OAA) to monitor and inventory the progress students have made in alignment with Ohio's academic standards. The outcomes of these assessments can result in different consequences for different students. Ohio's Department of Education indicates that these assessments could help to identify the areas where students have met proficiencies and where further help is needed (Ohio, Department Ohio Education [DOE], 2008a). When a student does not score at the proficient level on grade level assessments it could be one factor to help

evaluate the student's readiness to advance to the next grade (Ohio, DOE, n.d.). Refer to Table 1 for the OAA scaled score ranges:

Table 1

Scaled Score Ranges: Ohio Achievement Assessment (OAA)

Grade	Subject	Limited	Basic	Proficient	Accelerated	Advanced
5	Reading	< 384	384	400	441	459
	Math	< 382	382	400	424	439
	Science	< 363	363	400	417	448
6	Reading	< 380	380	400	436	456
	Math	< 378	378	400	429	448
7	Reading	< 379	379	400	432	452
	Math	< 378	378	400	436	458
8	Reading	< 378	378	400	428	451
	Math	< 379	379	400	432	459
	Science	<365	365	400	427	445

(Ohio, Department Ohio Education [DOE], 2013.)

The Ohio Graduation Test (OGT) requires students to perform at or above the proficient level in order to receive a high school diploma. The OGT allows students multiple attempts to demonstrate proficiency in each content area. Students will also be provided with interventions to improve specific areas where deficits are indicated (Ohio, Department Ohio Education [DOE], 2008b). Refer to Table 2 for the OGT scaled score ranges.

Table 2

Scaled Score Ranges: Ohio Graduation Test (OGT):

Subject	Limited	Basic	Proficient	Accelerated	Advances
Reading	255-382	383-399	400-428	429-447	448-559
Mathematics	255-382	384-399	400-424	425-443	444-566
Writing	274-377	378-399	400-429	430-475	476-577
Science	206-370	371-399	400-424	425-444	445-599
Social Studies	227-381	382-399	400-428	429-445	446-592

(Ohio, Department Ohio Education [DOE], 2005)

Students' performance on both the OAA and the OGT can be used to judge the performance of schools and districts relating to funding (Ohio, DOE, n.d.). It was also noted that when schools perform below expectations the state or district may offer additional resources like teacher training, extra instructional materials or coaching from more experienced teachers (Ohio, DOE, n.d.). When a school fails to meet performance goals for a period of time, students may be given the option to transfer to another school within their district (Ohio, DOE, n.d.). While these achievement assessments may have been implemented to motivate schools, there are resulting consequences that are counterproductive to their purpose and intent. Consequences may include unwarranted student retention, unequal district funding and students' inability to graduate.

With the importance placed on these assessments within each district and the outcomes that are concluded from these assessments, reliability is an important factor to consider. Reliability in achievement assessments is the

consistency and accuracy to which the tool measures students learning. Reliability is particularly important in assessments like the OAA and OGT due to the fact that the results are being generalized and used to make decisions about student placement, student retention, student ability to graduate, student intervention, district funding and teacher training. Reliability is also important to consider when looking at the generalizability and how the results of assessments are used. An analogy has been made that an unreliable assessment is like carpenters working with a rubber yardstick that stretches and contracts therefore, misrepresenting the true length of the board (Kaplan and Saccuzzo, 2009). The education field, as a whole needs to consider the influence these decisions may have on student placement and district funding.

Ohio's Assessment Publisher and Ohio Department of Education

The American Institute of Research (AIR) is the publisher and joint creator in part with the Ohio Department of Education (ODE) of both the OAA and the OGT. In an effort to find research articles or peer reviewed journals pertaining to the OAA and the OGT, contact was made via email and phone with both AIR and ODE. Contact was made with AIR and ODE after a computerized search in September 2013 of the OhioLINK Electronic Journal Center and EBSCO databases for the terms OAA, OGT, and consistency in the abstract and key words did not provide useful articles pertaining to this study. The information requested was research or studies that have been completed looking at reliability of these assessments. AIR responded quickly, indicating that ODE would need to be contacted in order to gain this information.

Contact from ODE produced the following explanation as to why the OAA's and OGT's can claim reliability. It was indicated that reliability comes from the approval of Ohio's plan that was submitted and approved by the U.S. Department of Education under NCLB (T.W. Moore, personal communication, 2013). It was noted the approval process went through peer reviews from a team of individuals from around the country (U.S. Department of Education, 2006a). On November 15, 2006 the U.S. Department of Education approved Ohio's assessment system under the Title I of the Elementary and Secondary Act of 1965 as amended by the NCLB Act of 2001 (U.S. Department of Education, 2006b). The decision was made as a result of peer reviewers who were not internal to the U.S. Department of Education (U.S. Department of Education, 2006b). This decision approved Ohio's standards and assessments under NCLB giving Ohio the ability to claim reliability of their achievement assessments (T.W. Moore, personal communication, 2013).

With the information provided from the Ohio Department of Education in response the reliability of the OAA and OGT assessments, a closer look will be taken as to what reliability is, how it is determined, and the importance of having said reliability.

Reliability

Reliability is the extent to which a score or measure is free of measurement error. Theoretically, reliability is the ratio of true score variance to observed score variance. This ratio can be estimated using a variety of correlational methods, including *coefficient alpha, Kuder-Richardson 20, tests-*

retest and parallel forms (Kaplan & Saccuzzo, 2009). Test-retest reliability allows the estimation of error associated with administering a test at two different times (Kaplan & Saccuzzo, 2009). Typically a test or assessment is considered reliable when the same results are produced repeatedly. Examination of the OAA and OGT results for a cohort of students across grades would be helpful in determining the reliability and consistency of these measures. This examination is needed due to the lack of empirical information on the reliability of these assessments.

Need For The Study

A representative from the Ohio Department of Education indicated that the assessments designed for Ohio went through a peer review and approval process conducted by the U.S. Department of Education. Due to that process there is not supporting evidence indicating that these assessments lack reliability (T.W. Moore, personal communication, 2013). For that reason the U.S. Department of Education deem these assessments to be reliable for their intended purposes (T.W. Moore, personal communication, 2013).

It was concluded that because there have not been any studies that looked at the consistency of scores on these tests over time that these tests can be considered reliable. Of course, the lack of data concerning the consistency of OAA and OGT scores does not mean the tests are reliable. It does, however, point to the need for research to examine this issue. For that reason the examination of data from these assessments would be beneficial to support the claims of reliability.

Hypothesis

The null hypothesis states that the related population means are not equal and at least two means are significantly different. The null hypothesis indicates that no relationships exist and that the related population scores for the OAA and OGT will not be equal or consistent. The null hypothesis also indicates that the most recent OAA score will not be the best predictor of the related OGT assessment score. The alternative hypothesis states that the related population means between subjects are equal and significant differences will not occur. The alternative hypothesis indicates that relationships will exist and that the related populations scores for the OAA and OGT will be consistent or stable. The alternative hypothesis also indicates the most recent OAA score will best predict the related OGT assessment score.

CHAPTER II

Method

Subjects

The subjects in this study are 167 students, 79 male and 88 female, from a rural school district located in central Ohio. All 167 subjects participate in the regular education setting with typical peers throughout the period of time the assessments were completed. Names are not associated with scores in order to maintain confidentiality. All subjects in the study had archived scores available for the following assessments: OAA's 5th, 6th, 7th and 8th grades and OGT data from the 10th through 12th grades.

Instruments

The Ohio Achievement Assessment (OAA) is an annual assessment used to measure students' knowledge of concepts taught in grades 3 through 8, pertaining to the content areas, English language arts, mathematics, and to applicable grades in science, writing and social studies. The OAA are given each spring and are based on Ohio's Academic Content Standards, which teachers are required to use as guidelines for instruction (Ohio, Department Ohio Education [DOE], n.d.).

The Ohio Graduation Test (OGT) is an assessment administered to all 10th graders in the spring of that year. The OGT consists of five sections: reading, writing, mathematics, science, and social studies. Students must obtain a proficient score on each individual section in order to receive a high school

diploma unless alternative routes are used. Students in special education can be exempted from the consequences of not passing the OGT (Ohio, DOE, n.d.).

Students' scores on both the OAA and OGT are described using the following performance levels: limited, basic, proficient, accelerated and advanced. The specific scaled scores that coincide for the performance level descriptors for each grade of the OAA and the OGT are on pages 4 and 5 of this document.

Procedures

The data comes from one cohort's scores on OAA and OGT over an eight-year period. The OAA data available is from 5th, 6th, 7th and 8th grades and OGT data from the 10th, 11th and 12th grades. Data pertaining to the OAA consist of reported scores from 5th grade reading, mathematics and science; 6th grade reading and mathematics; 7th grade reading and mathematics; and 8th grade reading, mathematics, and science. OAA for grades 5-8 are given one time per year and this is in the winter/spring, typically during the months of February and March. The cohort's available data from the OGT would range from spring of the 10th grade year through spring of the 12th grade year. Students are allowed to take the OGT assessment every fall, spring, and summer until passed; therefore, it was not possible to specify when the OGT assessment was taken. With multiple chances allowed to perform at or above the proficient level and the nature of the data available, it was not possible to pinpoint when each assessments was complete. Data pertaining to the OGT are available for writing, reading, mathematics, science and social studies.

Only complete sets of subject data were used. That was, only students for which OAA scores for the 5th, 6th, 7th, 8th grades and 10th grade OGT scores were included in this study. Also, students who use accommodations or modifications on these assessments, such as students in special education, were not included in this study because of the unknown effect these accommodations or modifications have on the test results. The Institutional Review Board (IRB) approval was obtained indicating that American Psychological Association ethical guidelines were observed for the study. The approval letter can be found in the Appendix.

Data Analysis

The data were analyzed with the following questions in mind. First, could OGT assessment scores be predicted from the OAA scores? The OGT scores from mathematics and reading were be used as a criterion measure against each math and reading OAA assessment score from 5th through 8th grade. This information was used to determine which OAA score was most predictive of the OGT assessment score.

Next, each individual OAA was compared to the succeeding year's OAA. This answered the question, does the preceding years OAA score predict the next years score? Third, what relationships exist between the OAA and the OGT assessment in the areas of mathematics and reading? To explore this area a correlational analysis was used to look at those relationships between the OAA scores for mathematics and reading against OGT assessment scores. This helped to answer the question as to whether relationships existed and if so what

is the strength of their relationship. Last, differences in mean scores in the OAA and OGT assessments were examined. This allowed us to determine if the cohort's performances on the OAA and OGT assessments were significantly different.

CHAPTER III

Results

Regression Analysis Mathematics

A regression analysis was computed to look at relationships between the mathematics OAA scores in comparison to the mathematics OGT assessment score. The five independent variables Ohio Achievement Assessment 5th grade Mathematics (OAA5MA), Ohio Achievement Assessment 6th grade Mathematics (OAA6MA), Ohio Achievement Assessment 7th grade Mathematics (OAA7MA), and Ohio Achievement Assessment 8th grade Mathematics (OAA8MA) in the regression model account for 66.5 % of the total variation in a given subject's score, the dependent variable Ohio Graduation Test Mathematics (OGTMA). Due to fact the test produced a significant result ($p < .001$), there is essentially no chance that the observed correlation between one or more of the independent variables and the dependent variable is due solely to random sampling. This allows the conclusion that significant differences do not exist. The standardized coefficients listed in the Beta column allow for a direct strength comparison between each independent variable, OAA score for mathematics. The OAA 8th grade assessment ($B = .381$) is the most important predictor of the OGT assessment, followed by OAA 6th grade assessment ($B = .268$). A coefficient table summarizes these results (Table 3):

Table 3

Regression OGT: Mathematics

Model	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence	
	B	Std. Error	BETA	t	Sig.	Lower Bound	Upper Bound
1(Constant)	43.163	22.350		1.931	.055	-.971	87
OAA5 MA	.066	.065	.073	1.014	.312	-.062	.194
OAA6 MA	.266	.083	.268	3.196	.002	.102	.431
OAA7 MA	.170	.084	.167	2.021	.045	.004	.336
OAA8 MA	.438	.098	.381	4.457	.000	.244	.631

Three regression analyses were computed to look at relationships between the mathematics OAA scores in comparison to the succeeding mathematics OAA score. The independent variables (preceding OAA score) in the regression model account for 50.2 to 60.9 % of the total variation in a given subjects score, the dependent variable (succeeding OAA score). Due to fact that all three tests produced a significant result ($p < .001$), there is essentially no chance that the observed correlation between the independent variable and the dependent variable is due solely to random sampling. This allows the conclusion that significant differences do not exist. The standardized coefficients listed in the Beta column allows for a direct strength comparison between each pair of OAA scores for mathematics. OAA 7th grade mathematics has the highest predictive ability for the succeeding (B = .781). Coefficient tables summarize these results (Table 4):

Table 4:

Regression: Math 5-6 OAA

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence	
Model	B	Std. Error	BETA	t	Sig.	Lower Bound	Upper Bound
OAA5MA	.646	.050	.708	12.893	.000	--	--

Regression: Math 6-7 OAA

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence	
Model	B	Std. Error	BETA	t	Sig.	Lower Bound	Upper Bound
OAA6MA	.746	.049	.766	15.318	.000	--	--

Regression: Math 7-8 OAA

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence	
Model	B	Std. Error	BETA	t	Sig.	Lower Bound	Upper Bound
OAA7MA	.694	.043	.781	16.047	.000	--	--

Regression Analysis Reading

A regression analysis was computed to look at relationships between the reading OAA scores in comparison to the reading OGT assessment score. The five independent variables Ohio Achievement Assessment 5th grade Reading (OAA5RD), Ohio Achievement Assessment 6th grade Reading (OAA6RD), Ohio Achievement Assessment 7th grade Reading (OAA7RD), and Ohio Achievement Assessment 8th grade Reading (OAA8RD) in the regression model account for 48.5 % of the total variation in a given subject's score, the dependent variable Ohio Graduation Test Reading (OGTRD). Due to the fact the test produced a significant result ($p < .001$), there is essentially no chance that the observed correlation between one or more of the independent variable and the dependent

variable is due solely to random sampling. This allows us to determine that significant differences do not exist. The standardized coefficients listed in the Beta column allow for a direct strength comparison between each independent variable, OAA assessment score for reading. OAA 8th grade assessment (B = .299) is the most important predictor of the OGT assessment, followed by OAA 6th grade assessment (B = .250). A coefficient table summarizes these results (Table 5):

Table 5

Regression OGT: Reading

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence	
Model	B	Std. Error	BETA	t	Sig.	Lower Bound	Upper Bound
1(Constant)	129.261	25.601		5.049	.000	78.707	179.816
OAA5 RD	.106	.062	.128	1.704	.092	-.017	.228
OAA6 RD	.228	.074	.250	3.071	.002	.081	.374
OAA7 RD	.129	.090	.132	1.426	.156	-.050	.307
OAA8 RD	.249	.075	.299	3.317	.001	.101	.396

Three regression analyses were computed to look at relationships between the reading OAA scores in comparison to the succeeding reading OAA score. The independent variables (preceding OAA score) in the regression model account for 29.7 to 53.2 % of the total variation in a given subject's score, the dependent variable (succeeding OAA score). Due to fact that all three tests produced a significant result ($p < .001$), there is essentially no chance that the observed correlation between the independent variable and the dependent variable is due solely to random sampling. This allows the determination that significant differences do not exist. The standardized coefficients listed in the Beta column allows for a direct strength comparison between each pair of OAA

scores for reading. OAA 7th grade reading assessment has the highest predictive ability for the succeeding year with a (B = .730). Coefficient tables summarize these results (Table 6):

Table 6

Regression: Reading 5-6 OAA

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence	
Model	B	Std. Error	BETA	t	Sig.	Lower Bound	Upper Bound
OAA5RD	.493	.059	.545	8.344	.000	---	---

Regression: Reading 6-7 OAA

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence	
Model	B	Std. Error	BETA	t	Sig.	Lower Bound	Upper Bound
OAA6RD	.593	.056	.633	10.512	.000	---	---

Regression: Reading 7-8 OAA

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence	
Model	B	Std. Error	BETA	t	Sig.	Lower Bound	Upper Bound
OAA7RD	.854	.062	.730	13.700	.000	---	---

Correlational Analyses

A Pearson product-moment correlation coefficient was computed to assess the relationship between all the OAA and OGT mathematics scores. There are positive correlations to report between all variables for mathematics. A correlation matrix summarizes the results (Table 7). Overall, there were *strong*, positive correlations between all variables ranging from r = .708 to r = .798. The only exceptions being OAA 5th grade mathematics (OGT5MA) and OGT mathematics (OGTMA) with a *moderate* positive correlation of r = .652.

Table 7

Correlations Matrix Mathematics

		OAA5 MA	OAA6 MA	OAA7 MA	OAA8 MA	OGT MA
OAA5	Pearson Cor.	1	.708**	.723**	.706**	.652**
MATH	Sig. (2-tailed)		.000	.000	.000	.000
	N	167	167	167	167	167
OAA6	Pearson Cor.	.708**	1	.766**	.798**	.751**
MATH	Sig. (2-tailed)	.000		.000	.000	.000
	N	167	167	167	167	167
OAA7	Pearson Cor.	.723**	.766**	1	.781**	.722**
MATH	Sig. (2-tailed)	.000	.000		.000	.000
	N	167	167	167	167	167
OAA8	Pearson Cor.	.706**	.798**	.781**	1	.776**
MATH	Sig. (2-tailed)	.000	.000	.000		.000
	N	167	167	167	167	167
OGT	Pearson Cor.	.652**	.751**	.722**	.766**	1
MATH	Sig. (2-tailed)	.000	.000	.000	.000	
	N	167	167	167	167	167

Correlations OAA/OGT

		OAA5 MA	OAA6 MA	OAA7 MA	OAA8 MA
OGT	Pearson Correlation	.652**	.751**	.722**	.776**
MATH	Sig. (2-tailed)	.000	.000	.000	.000
	N	167	167	167	167

A Pearson product-moment correlation coefficient was also computed to assess the relationship between all the OAA and OGT reading scores. There are positive correlations to report between all variables for reading. A correlation matrix summarizes the results (Table 8). Overall, there were *moderate*, positive correlations between all variables ranging from r = .507 to r = .638. The only exceptions being OAA 6th grade reading (OAA6RD) and OAA 7th grade reading (OAA7RD) with a *strong* positive correlation of r = .730.

Table 8

Correlations Matrix Reading

		OAA5 RD	OAA6 RD	OAA7 RD	OAA8 RD	OGT RD
OAA5	Pearson Cor.	1	.545**	.638**	.528**	.507**
READING	Sig. (2-tailed)		.000	.000	.000	.000
	N	167	167	167	167	167
OAA6	Pearson Cor.	.545**	1	.633**	.677**	.607**
READING	Sig. (2-tailed)	.000		.000	.000	.000
	N	167	167	167	167	167
OAA7	Pearson Cor.	.638**	.633**	1	.730**	.591**
READING	Sig. (2-tailed)	.000	.000		.000	.000
	N	167	167	167	167	167
OAA8	Pearson Cor.	.528**	.677**	.730**	1	.633**
READING	Sig. (2-tailed)	.000	.000	.000		.000
	N	167	167	167	167	167
OGT	Pearson Cor.	.507**	.607**	.591**	.633**	1
READING	Sig. (2-tailed)	.000	.000	.000	.000	
	N	167	167	167	167	167

Correlations OAA/OGT

		OAA5 RD	OAA6 RD	OAA7 RD	OAA8 RD
OGT	Pearson Correlation	.507**	.607**	.591**	.633**
READING	Sig. (2-tailed)	.000	.000	.000	.000
	N	167	167	167	167

ANOVA

A one-way within subjects repeated measure ANOVA was conducted to compare the effect of the within subjects factor on the cohort's performance within the mathematics assessments for the OAA and OGT. The mean scores were taken to complete the ANOVA, OAA5MA = 420.24, OAA6MA = 438.98, OAA7MA = 419.71, OAA8MA = 416.83 and OGTMA = 441.44. Due to the fact that the test did not produce a significant result ($p < .001$), the sphericity assumption was not violated. For this reason we will look at the Greenhouse-Geisser. The ANOVA shows these results: $F(3.56, 592.42) = 137.98$, $p < .001$. A descriptive *statistics table* and *test of within-subject effects* table summarizes

these results (Table 9). These results suggest that differences between mean scores in the OAA and OGT assessment for mathematics are not statistically significant and not due to chance.

Table 9

ANOVA Mathematics

Descriptive Statistics

	Mean	Std. Deviation	N
OAA5MA	420.24	27.63	167
OAA6MA	438.98	25.18	167
OAA7MA	419.70	24.52	167
OAA8MA	416.82	21.80	167
OGTMA	441.43	25.03	167

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Reading Assessments	Greenhouse-Geisser	92424.53	3.56	25897.84	137.98	.000
Error (Reading Assessments)	Greenhouse-Geisser	111191.46	592.42	187.69		

A one-way within subjects repeated measure ANOVA was conducted to compare the effect of the within subjects factor on the cohort's performance within the reading assessments for the OAA and OGT. The mean scores were taken to complete the ANOVA, OAA5RD = 430.29, OAA6RD = 430.19, OAA7RD = 425.55, OAA8RD = 432.65 and OGTRD = 435.08. Due to the fact that the test did not produce a significant result ($p < .001$), the sphericity assumption was not violated. For this reason the Greenhouse-Geisser was analyzed. The ANOVA shows these results: $F(3.60, 599.14) = 12.18$, $p < .001$. A descriptive *statistics table* and *test of within-subject effects* table summarizes these results (Table 10).

These results suggest that differences between mean scores in the OAA and OGT assessment for reading are not statistically significant and not due to chance.

Table 10

ANOVA Reading

Descriptive Statistics

	Mean	Std. Deviation	N
OAA5RD	430.29	22.65	167
OAA6RD	430.18	20.50	167
OAA7RD	425.55	19.20	167
OAA8RD	432.65	22.47	167
OGTRD	435.08	18.66	167

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Reading Assessments	Greenhouse-Geisser	8343.31	3.61	2311.62	12.18	.000
Error (Reading Assessments)	Greenhouse-Geisser	113661.09	599.14	189.70		

CHAPTER IV

Discussion

The purpose of this study was to investigate the consistency and relationships of a cohort's reported performance on the OAA over a four-year period (5^{th}, 6^{th}, 7^{th} and 8^{th} grades) and reported performance on their OGT assessment. Assessments like the OAA and OGT need evaluated to support their claims, due to the fact that the results are being generalized and used to make decisions about student placement, student retention, student ability to graduate, student intervention, district funding and teacher training.

The null hypothesis stated that the related population means would not be equal and at least two means would be significantly different. The null hypothesis indicated that no relationships would exist and that the related population scores for the OAA and OGT would not be equal or consistent. The null hypothesis also indicated that the most recent OAA score would not be the best predictor of the related OGT assessment score. With that being said the null hypotheses were rejected and the alternative hypotheses were accepted.

The alternative hypothesis stated that the related population means between subjects are equal and significant differences would not occur. The alternative hypothesis indicated that relationships would exist and that the related population scores for the OAA and OGT would be consistent or stable. Results were produced indicating that *strong* positive correlations exist between the mathematic assessments for OAA and OGT. Results were also produced indicating that *moderate* positive correlations exist between the reading

assessments for OAA and OGT. Results also indicated that differences between mean scores in the OAA and OGT for both mathematics and reading assessments are not statistically significant and not due to chance.

The alternative hypothesis also indicated the most recent OAA score will be the best predictor of the related OGT assessment score. Results were produced indicating that the 8th grade OAA for both mathematics and reading was the most important predictor of the related OGT assessment. In both areas analyzed the next most important predictor for the related OGT assessment was the 6th grade OAA assessment, not the 7th grade, which is interesting. It is also worth noting that when predictors for the OAA were examined against themselves, the 7th grade OAA assessment was most predictive of the succeeding year for both mathematics and reading. Results indicating that 7th grade OAA were the most important predictor of the related 8th grade OAA.

In conclusion, with the results produced from the analyses performed on the data, it can be said with a reasonable degree of statistical certainty that the OAA and OGT assessments are consistent and produce stable results. Therefore, the null hypotheses were rejected and the alternative hypotheses were accepted. Even though *strong* to *moderate* positive correlations were reported, predictions cannot be made from the previous year scores. This information should be interpreted with caution, even with the 8th grade OAA for mathematics and reading being the best prediction of the 10th grade OGT score. One will be unable to make predictions about passing rates or scores from previous year's scores on any one student. Other extraneous variables may

impact a student's score, which cannot be foreseen. Unfortunately a comparison between this study and other research cannot be made due to the fact that there have not been any studies that looked at the consistency of the relationships within and between the OAA and OGT assessments.

Limitations of this study include the use of only a single cohort's reported assessment scores on the OAA and OGT from a rural school district in central Ohio. Also, only complete sets of data were analyzed; therefore it may not be completely representative of a true student population performance. Identified students were also not included in the study due to unknown effects of modifications and accommodations. Future studies should include larger, more diverse populations of students. If data were available, a comparison should be made between differences in rural areas and urban areas to see if the same results, correlations, predictions and differences are produced.

References

Kaplan, R. M., Saccuzzo, D.P. (2009). *Psychology Testing: Principles, Applications, and Issues.* Belmont, CA: Wadsworth CENGAGE Learning.

Lee, J., & Reeves, T. (2012). Revisiting the Impact of NCLB High-Stakes School Accountability, Capacity, and Resources: State NAEP 1990–2009 Reading and Math Achievement Gaps and Trends. *Educational Evaluation and Policy Analysis,* 34(2), 209-231. Doi:10.3102/0162373711431604.

Moore, T.W. (2013, September 17). Personal Communication. Telephone Interview.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat 1425 (2002).

Ohio Department of Education (n.d.). Frequently Asked Questions (FAQS): Frequently Asked Questions for Families. Retrieved from http://ohio3-8.success-ode-state-oh-us.info/faqs.aspx.

Ohio Department of Education (2013). Ohio Achievement Assessment-statistical summary: Table 8 – Cut scores Points for Basic, Proficient, Accelerated, and Advanced Standards. Retrieved from http://education.ohio.gov/getattachment/Topics/Testing/Statistical-Summaries-and-Item-Analysis-Reports/May-2013-OAA-Statistical-Summary-2.pdf.aspx.

Ohio Department of Education (2005). Statistical Summary of the Ohio Graduation Tests: Table 2 – Scores ranges for all subjects and all performance levels. Retrieved from https://education.ohio.gov/getattachment/Topics/Testing/Testing-Analysis-and-Statistics/Statistical-Summaries-and-Item-Analysis-Reports/March-2005-OGT-Statistical-Summary.pdf.aspx.

Ohio Department of Education (2008a). Validity Evidence Based on Internal Structure: Examination of the Factor Structure of the Ohio Achievement Tests. Retrieved from http://www.ohiodocs.org/Technical%20Docs/TR%202008%20-%203%20OAT%20Validity%20Study%202007.pdf.

Ohio Department of Education (2008b). Validity study: A collection of evidence about the Ohio Graduation Tests. Retrieved from http://www.ohiodocs.org/Technical%20Docs/TR%202008%20-%201%20OGT%20Validity%20Study%202006.pdf.

U.S. Department of Education (2003). State of Ohio Consolidated Application Accountability Workbook. Retrieved from http://www2.ed.gov/admins/lead/account/stateplans03/ohcsa.pdf.

U.S. Department of Education (2006a). LEAD & MANAGE MY SCHOOL, Ohio Assessment Letter, 5/27/2006. Retrieved from http://www2.ed.gov/admins/lead/account/nclbfinalassess/oh.html.

U.S. Department of Education (2006b). LEAD & MANAGE MY SCHOOL, Ohio Assessment Letter, 11/15/2006. Retrieved from http://www2.ed.gov/admins/lead/account/nclbfinalassess/oh2.html.

Wiliam, D. (2010). Chapter 8 What Counts as Evidence of Educational Achievement? The Role of Constructs in the Pursuit of Equity in Assessment. *Review of Research in Education*, *34*(1), 254-284. doi:10.3102/0091732X09351544.

MARSHALL UNIVERSITY
www.marshall.edu

Office of Research Integrity

June 17, 2014

Sandra S. Stroebel, PhD, NCSP
Associate Dean / Program Director / Professor
Marshall University College of Education and Professional Development
School Psychology Program

Dear Dr. Stroebel:

This letter is in response to the submitted abstract entitled *"Examination of Consistency on the Ohio Achievement Assessments and Ohio Graduation Test."* After assessing the abstract it has been deemed not to be human subject research and therefore exempt from oversight of the Marshall University Institutional Review Board (IRB). The Code of Federal Regulations (45CFR46) has set forth the criteria utilized in making this determination. The information in this study does not meet the federal definition of human subject research and is therefore not subject to Common Rule oversight. If there are any changes to the abstract you provided then you will need to resubmit that information for review and determination.

I appreciate your willingness to submit the abstract for determination. Please feel free to contact the Office of Research Integrity if you have any questions regarding future protocols that may require IRB review.

Sincerely,

Bruce F. Day

Bruce F. Day, ThD, CIP
Director

Adam R. Fox
450 Forest St. Columbus, Ohio 43206. fox98@marshall.edu

Education

Education Specialist Degree in School Psychology and Masters in Education from Marshall University, anticipated graduation Summer 2014

Bachelors of Arts Degree in Psychology from Ohio University, graduated June 2009

Professional Experience

- 2012/13 Internship/practicum with Southwest Licking School District, experience working with students' pre-k through 12th grade.
- Mentorship at Heath City Schools, August 2010 - May 2012.
- Administer, score and report numerous assessments including cognitive, academic, behavior, and executive functioning.
- Coordinate and lead Evaluation Team Report (ETR) meetings and Intervention Assistant Team (IAT) meetings.
- Coordinate consultations with parents, teachers, and school administrators to define academic and behavior issues of students.
- Create and implement Curriculum Based Measures (CBMs) to assess student progress.
- Implement Evidence Based Interventions (EBIs) for general and special education classrooms.
- Conduct Functional Behavioral Analysis (FBA) to help define and gain a deeper understanding of a student's problematic behavior.
- Possess the knowledge and skills necessary to implement the Three Tiered Model of Response to Intervention (RTI).
- Perform individual and group counseling sessions with middle school students.

Professional Membership

- Ohio School Psychologists Association (OSPA)
- West Virginia School Psychologist Association (WVSPA)
- National Association School Psychologists (NASP)