When do Seventh Grade Students Perform Best on Reading and Mathematics Assessments?

Amanda A. Wolkowitz, Assessment Technologies Institute

Jonathan Templin, University of Georgia

Neal Kingston, University of Kansas

When do Seventh Grade Students Perform Best on Reading and Mathematics Assessments

When is the best time for a student to complete a state assessment? Is Monday a good day because the students have had the weekend to relax? Is Friday a better choice because the weekend is coming? Perhaps the middle of the week is the best day because students are more focused. Is it better to test students first thing in the morning or do they perform better after lunch? These are not new questions, but they are questions that researchers have left largely unanswered. Moreover, these are questions asked, and despite a lack of data, acted upon by all schools participating in high stakes accountability testing. The purpose of this study is to address these questions and to offer some insight into the best time to test students. The benefit of such a study is straightforward: schools that are held accountable for student performance on standardized state assessments may benefit from information available that suggests when their students perform best.

The first question addressed in this research is "is there a significant change in performance on a state assessment between the first and second sections of the test?" Since there is no overlap in questions between the sections, non-academic reasons may account for any possible performance changes. Although taking a second section of an assessment is quite different than retaking an assessment, reasons for a score change on a second section of an assessment may be similar to the reasons for a score change on a retake of an assessment. These reasons include measurement error, a gain in knowledge from the first testing experience, a reduction in test anxiety or unfamiliarity with the testing environment from the first testing experience, or other changes due to non-academic reasons (Lievens, Buyse, and Sackett, 2005).

The second question examined in this research is "when is the best day of the week to test students?" Laird (1925) studied how well 112 college students performed on nine different exams. The students were divided into groups of 16 so that each test could be given at different times of the day and different days of the week. Each group of 16 took the same sequence of 9 exams, but the time of day and day of the week each group completed the exams differed. Laird's research indicated that student performance gradually increased during the beginning of a week, peaked on Wednesdays, and sharply declined during the end of the week. Few other studies have investigated this question.

The final question addressed in this research is "when is the best time to test students during the day?" Although personal learning style accounts for individual performance preferences, some research found that students perform better in the morning than in the afternoon (Callan, 1999; Dunn, 1985). For example, Callan (1999) administered an algebra test from the State of New York Regents Sequential I Mathematics Examination to 281 ninth grade students in New York City. He found that students performed significantly better when they completed the assessment in the morning than in the afternoon. Dunn's (1985) research on elementary students supported Callan's findings. Dunn reported that approximately one-third of elementary students perform best before 10:00 a.m., one third between 10:00 and 10:30 a.m., and the remaining third after 10:30 a.m. Other research, however, concluded that individual preferences and learning style have the greatest affect on the optimal time to test (Ammons et al, 1995; Anderson, Petros, Beckwith, Mitchell, and Fritz, 1991).

Although little to no research suggests that afternoon testing produces higher test scores overall than morning testing, Ammons et al. (1995) and Anderson et al. (1991) suggest that the best time of day to test depends on each student's most alert time. They found that an

individual's ability to recall information from their memory was related to the time of day that person preferred. Ammons et al. also noted that some students' best time is a reflection of their teacher's best time of day for instructing.

The objective of this study is to investigate the patterns of student performance based on the day of the week and time of day students took a Midwestern state's large-scale computerized assessment. The computerized assessment was administered in three sections. The research questions for this study are the following:

1) Is there a significant difference in performance on Section 2 of the mathematics (reading) assessment compared to Section 1 regardless of the day of the week and time of day the sections were completed?
2) After controlling for the day of the week that a student completes Section 1 of the mathematics (reading) assessment, is there a significant difference in a student's performance on Section 2 of the assessment based on the day of the week this section was completed?
3) Is there a significant difference in performance on Section 2 compared to Section 1 of the mathematics (reading) assessment based on the time of day each section was completed?

It was hypothesized that no significant differences in test performance would be detected for each of the three research questions.

Methods

Participants

The sample consisted of two (highly overlapping) samples of seventh grade students who took a computerized assessment in a Midwestern state between February 18th and April 14th, 2009: those who completed all three sections of the mathematics assessment and those who completed all three sections of the reading assessment.

In the state, 25,985 seventh grade students completed all three sections of the mathematics assessment. Of these students, 19,943 completed all three sections of the assessment during the same week. A total of 17,030 students completed the test in the intended order; i.e., completed the Section 1 first, Section 2 second, and Section 3 third.

The reading portion of the assessment was also divided into three sections. Unlike the mathematics test in which each form had the same number of questions per section, the number of questions per section on the reading test varied. Therefore, to avoid any potential confounding due to minor differences in section length only one form of the reading test was included in the analysis. A total of 8,031 students completed this form of the reading test. Of these students, 6,375 students completed all three sections of the assessment during the same week and 6,185 completed the sections in the correct order.

Materials

The materials for this study consisted of four equated computerized forms of the state's mathematics assessment and one form of the computerized version of the state's reading assessment for grade seven. The mathematics and reading assessments were the state assessments used to measure annual yearly progress under the 2001 No Child Left Behind

legislation. Consequently, the assessment was a direct measure of the mathematics and reading standards set forth by the state for seventh grade students.

The mathematics and reading assessments each consisted of three sections. The item totals were 28, 29, and 27 for the three sections of the seventh grade mathematics assessment regardless of form, respectively, and 25, 25, and 34 items for the three sections on the reading assessment. There was no time limit for the completing each section of the assessment, but it was recommended that students spend approximately 45 minutes per section and that at most two sections be administered in one day.

Procedure

Datafile.

For this study, the focus was on the date and time that students completed each section of the assessment. Since there were three sections of each assessment and only the total scores were equated for operational purposes, the section scores were equated for this study by estimating each examinee's ability for each section tested. These estimates were computed using a 3-parameter logistic model and BILOG-MG. Theoretically, a student should have the same ability regardless of which section of the test he or she is completing. Therefore, if an increase in ability level was detected between a student's performance on Section 1 and Section 2, for example, then this might indicate that the student performed better on Section 2 than on Section 1.

In summary, the reading and mathematics data files used consisted of seventh grade binary data, section subtotals, ability levels per section, and the exact date and time each student completed each section of the assessment. Only those students that completed the sections of the test in the intended order and completed all sections of the reading assessment within one week and all sections of the mathematics assessment within one week were included.

Methods of Analysis.

This study sought to answer the question of whether or not the time of day or the day of the week a statewide assessment is administered significantly affects the score on the assessment. Before answering this question, a pairwise comparison was completed to determine if there was a significant difference in ability between performance on Section 1 and Section 2 of the reading and mathematics tests, respectively.

To answer the day of week question, a series of ANOVAs were conducted for both the mathematics and reading assessments separately. The dependent variable in the ANOVAs was the difference in ability as estimated separately by the responses to items from Section 1 and Section 2. Using the difference in ability as the dependent variable allowed for the ability on Section 1 to be "controlled" and the increase or decrease of ability on Section 2 to be analyzed. A two-way ANOVA was not completed due to the 10 empty cells that would result from no students completing Section 1 of the assessment after Section 2 (see Table 1 or 2). Therefore, four separate ANOVAs and one t-test were conducted to determine the following: 1) If a student completed Section 1 on Monday, was there a significant difference in the student's ability if the second section was completed on Monday, Tuesday, Wednesday, Thursday, or Friday; 2) If a student completed Section 1 on Tuesday, was there a significant difference in a student's ability if the second section was completed on Tuesday, Wednesday, Thursday or Friday; 3) If a student completed Section 1 on Wednesday, was there a significant difference in ability if the second section was completed on Wednesday, Thursday, or Friday; 4) If a student completed Section 1 on Thursday, was there a significant difference in ability if the second section was completed on Thursday or Friday; and 5) If a student completed Section 1 on Friday, was there a significant difference in ability if the second section was also completed on

Friday. Bonferroni adjustments were made to each ANOVA and pairwise comparisons were planned as follow up tests if significant results were found.

To answer the time of day question, a 4 x 4 ANOVA was conducted for the mathematics and reading assessments separately. The first factor was the time in which each student completed Section 1 of the test and the second factor was the time in which each student completed the Section 2 of the test. The four levels of each factor represented four time intervals: before 10:00 a.m.; between 10:00 a.m. and 11:59 a.m.; between 12:00 p.m. and 1:59 p.m.; and *at 2:00 p.m. or later* One way ANOVAs and pairwise comparisons with a Bonferroni adjustment were planned as follow up tests if significant results were found.

Effect size measures were computed with the r^2 and partial-eta squared statistics. Values less than 0.01 were considered not practically significant. Values between 0.01, inclusive, and 0.06 were considered small effect sizes. Values between 0.06, inclusive, and 0.14 were considered medium effect sizes and values greater than or equal to 0.14 were considered large effect sizes (Cohen, 1988).

Results

Is there a significant difference in test scores between the first and second sections of the assessment?

For mathematics, student performance on Section 1 (M = -0.015, SD = 0.931) was less than that on Section 2 (M = 0.008, SD = 0.931). Based on a dependent sample t-test, this ability difference was statistically significant, $t(17029) = -4.75$, $p < 0.001$, $r^2 = 0.002$. However, the small effect size indicates that this difference was not practically significant.

For reading, student performance on Section 1 (M = -0.002, SD = 0.873) was less than that on Section 2 (M = 0.003, SD = 0.895). However, using a dependent sample t-test these results were not significant, $t(6185) = -0.645$, $p = 0.519$, $r^2 < 0.001$.

Is there a significant difference in test scores between the weekday students completed the first and second sections of the assessment?

Four separate ANOVAs and a t-test were conducted to evaluate differences between a student's ability difference on Section 1 and Section 2 of the test and the weekday in which they completed each section. Table 1 displays the mean differences, standard deviations, and sample sizes for the mean ability difference between the two sections of the mathematics assessment. Table 2 displays a similar table for the reading assessment.

For the mathematics assessment, 13 comparisons were done so the alpha level was set at 0.05/13 = 0.004. None of the ANOVAs or the t-test revealed significant results at this level. The first ANOVA was conducted to determine if there was a significant difference in ability for those students who completed Section 1 of the assessment on Monday and Section 2 on Monday, Tuesday, Wednesday, or Thursday. Friday was not included in the analysis since Table 1 indicates that only 2 people completed Section 1 on Monday and Section 2 on Friday. For those who completed Section 1 on Monday, the results indicated no significant results based on the day of the week Section 2 was completed, $F(3, 6087) = 1.151$, $p = 0.327$, partial $\eta^2 = 0.001$.

A second ANOVA was conducted to determine if there was a significant difference in ability for those students who completed Section 1 of the assessment on Tuesday and Section 2 of the assessment on Tuesday, Wednesday, or Thursday. Friday was not included in the analysis since only 17 people completed Section 1 of the assessment on Tuesday and Section 2 on Friday. For those who completed Section 1 on Tuesday, the results indicated no significant difference

based on the day of the week Section 2 was completed, $F(2, 9308) = 0.518$, $p = 0.277$, partial $\eta^2 < 0.001$.

A third ANOVA was conducted to determine if there was a significant difference in ability between those students who completed Section 1 of the assessment on Wednesday and Section 2 on Wednesday, Thursday, or Friday. The differences in ability were not significant, $F(2, 1347) = 0.518$, $p = 0.596$, partial $\eta^2 = 0.001$. A fourth ANOVA was conducted to determine if there was a significant difference in ability between those students who completed Section 1 of the assessment on Thursday and Section 2 on either Thursday or Friday. No significant difference in scores were detected, $F(1, 242) = 4.284$, $p = 0.040$, partial $\eta^2 = 0.017$. Finally, a comparison between the differences in ability between those who completed both Section 1 and Section 2 of the mathematics assessment on Friday did not indicate a significant difference in performance, $t(14) = 0.272$, $p = 0.790$, $r^2 = 0.005$.

For the reading assessment, 11 comparisons were made so the alpha level was set at 0.05/11 = 0.005. None of the ANOVAs or the t-tests revealed significant results. The first ANOVA was conducted to determine if there was a significant difference in ability for those students who completed Section 1 of the assessment on Monday and Section 2 on Monday, Tuesday, Wednesday, or Thursday. Friday was not included in the analysis since Table 2 indicates that only 1 person completed Section 1 on Monday and Section 2 on Friday. For those who completed Section 1 on Monday, the results indicated no significant results between the day of the week Section 2 was completed, $F(3, 2283) = 0.490$, $p = 0.689$, partial $\eta^2 = 0.001$.

A second ANOVA was conducted to determine if there was a significant difference in ability for those students who completed Section 1 of the assessment on Tuesday and Section 2 of the assessment on Tuesday, Wednesday, or Thursday. Friday was not included in the analysis

since only 5 people completed Section 1 of the assessment on Tuesday and Section 2 on Friday. For those who completed Section 1 on Tuesday, the results indicated no significant difference in performance based on the day of the week Section 2 was completed, $F(2, 3073) = 0.876$, $p = 0.416$, partial $\eta^2 < 0.001$.

A third ANOVA was conducted to determine if there was a significant difference in ability between those students who completed Section 1 of the assessment on Wednesday and Section 2 on either Wednesday or Thursday. Again, Friday was not included in the analysis since only 8 people completed Section 1 on Wednesday and Section 2 on Friday. The differences in ability were not significant, $F(1, 737) = 1.633$, $p = 0.202$, partial $\eta^2 = 0.002$.

Two t-tests were conducted to determine if there was a significant difference in ability between those students who completed both Section 1 and Section 2 on Thursday and those who completed both sections on Friday. An ANOVA was not conducted for those who completed Section 1 on Thursday and Section 2 on either Thursday or Friday because only 9 people who completed Section 1 on Thursday completed Section 2 on Friday. No significant difference was found between those who completed both sections of the test on Thursday, $t(49) = -1.187$, $p = 0.241$, $r^2 = 0.0279$. Similarly, for the 10 people that completed both sections of the reading assessment on Friday, no significant difference in performance was detected, $t(9) = -3.359$, $p = 0.008$, $r^2 = 0.556$.

Tables 1 and 2 provide some additional information about the day of the week students are tested. In general, the tables show a pattern of students performing worse on Section 2 than on Section 1 of the test if both sections of the test were administered on the same day. With the exception of the testing on Monday followed by Tuesday, there also appears to be a pattern of students performing better on Section 2 than on Section 1 when the test was given on sequential

days. Finally, by examining the least and greatest abilities in the total row and total column, a pattern in Table 2 suggests that Wednesday is the best day to test reading ability and Friday is the worst day, regardless of which section of the test is administered. There is not a clear pattern suggesting the best or worst day for the mathematics test.

Table 1

Mean Differences in Ability Between the First and Section Sections of the Mathematics Assessment Based on Day of Week Assessment Completed: Ability on Section 2 – Ability on Section 1 (standard deviation of the difference scores are indicated in parentheses)

		Section 1					
		Monday	Tuesday	Wednesday	Thursday	Friday	**TOTAL**
Section 2	Monday	-0.055 (0.622) N = 126	X	X	X	X	**-0.055** **(0.622)** **N = 126**
	Tuesday	0.029 (0.648) N = 5294	-0.165 (0.562) N = 207	X	X	X	**0.027** **(0.645)** **N = 5501**
	Wednesday	0.009 (0.652) N = 576	0.030 (0.635) N = 8686	-0.046 (0.621) N = 208	X	X	**0.027** **(0.635)** **N = 9470**
	Thursday	-0.043 (0.625) N = 95	0.068 (0.681) N = 418	0.004 (0.662) N = 1096	-0.184 (0.673) N = 159	X	**0.000** **(0.668)** **N = 1768**
	Friday	-0.352 (0.190) N = 2	0.095 (0.667) N = 17	0.004 (0.510) N = 301	0.006 (0.703) N = 85	0.041 (0.577) N = 15	**0.013** **(0.632)** **N = 165**
	TOTAL	**0.024** **(0.647)** **N=6093**	**0.031** **(0.635)** **N = 9328**	**-0.003** **(0.651)** **N = 1350**	**-0.118** **(0.688)** **N = 244**	**0.041** **(0.577)** **N = 15**	**0.023** **(0.642)** **N = 17030**

Table 2

Mean Differences in Ability Between the First and Section Sections of the Reading Assessment Based on Day of Week Assessment Completed: Ability on Section 2 – Ability on Section 1 (standard deviations indicated in parentheses)

		Section 1					
		Monday	Tuesday	Wednesday	Thursday	Friday	**TOTAL**
Section 2	Monday	-0.054 (0.642) N = 25	X	X	X	X	**-0.054** **(0.642)** **N = 25**
	Tuesday	-0.015 (0.604) N = 1924	-0.056 (0.558) N = 76	X	X	X	**-0.016** **(0.602)** **N = 2000**
	Wednesday	-0.026 (0.601) N = 310	0.030 (0.614) N = 2868	-0.080 (0.582) N = 76	X	X	**0.022** **(0.612)** **N = 3254**
	Thursday	-0.146 (0.630) N = 28	-0.002 (0.583) N = 132	0.009 (0.579) N = 663	-0.082 (0.486) N = 50	X	**-0.003** **(0.576)** **N = 873**
	Friday	0.2166 (N/A) N = 1	0.082 (0.577) N = 5	-0.342 (0.726) N = 8	0.081 (0.447) N = 9	-0.451 (0.424) N = 10	**-0.178** **(0.567)** **N = 33**
	TOTAL	**-0.018** **(0.604)** **2288**	**0.027** **(0.611)** **N = 3081**	**-0.004** **(0.582)** **N = 747**	**-0.057** **(0.480)** **N = 59**	**-0.451** **(0.424)** **N = 10**	**0.005** **(0.604)** **N = 6185**

Is there a significant difference in test scores between the time of day students completed the first and second sections of the assessment?

A 4 x 4 ANOVA was conducted to evaluate any differences between a student's ability difference on Section 1 and Section 2 of the test and the time of day in which students completed each of these sections. The mean ability differences, standard deviations, and sample sizes for the mathematics and reading assessments are displayed in Tables 3 and 4, respectively. For the mathematics assessment, the results of the ANOVA indicated no significant findings. The interaction between the time of day in which students completed Section 1 and Section 2 of the assessment and the difference in ability between the two sections was not significant, $F(9, 16978) = 1.033$, $p = 0.410$, partial $\eta^2 = 0.001$. In addition, the main effect of the time of day Section 1 was completed was not significant, $F(3, 16978) = 1.717$, $p = 0.161$, partial $\eta^2 < 0.001$.

The main effect of the time of day Section 2 was completed was also not significant, $F(3, 16978) = 0.345$, $p = 0.792$, partial $\eta^2 < 0.001$.

For the reading assessment, the interaction between the time of day Section 1 and Section 2 was administered was not statistically significant, $F(9, 6152) = 1.398$, $p = 0.183$, partial $\eta^2 = 0.002$. The main effect of time the first section was given was also not statistically significant, $F(3, 6152) = 0.955$, $p = 0.413$, partial $\eta^2 < 0.001$. The main effect of time the second section was given, however, was statistically significant, $F(3, 6152) = 4.240$, $p = 0.005$, partial $\eta^2 = 0.002$. This small effect size indicates that the time Section 2 was administered only accounts for 0.2% of the variance in the ability difference between the two sections and is not practically significant.

Table 3

Mean Differences in Ability Between the First and Section Sections of the Mathematics Assessment Based on Time of Day Section Completed: Ability on Section 2 – Ability on Section 1 (standard deviations indicated in parentheses)

		Section 1				
		Before 10:00 a.m.	10:00 a.m. – 11:59 a.m.	12:00 p.m. – 2:00 p.m.	At 2:00 p.m. or later	**TOTAL**
Section 2	Before 10:00 a.m.	0.008 (0.634) N = 4311	0.026 (0.640) N = 954	0.034 (0.625) N = 244	0.036 (0.711) N = 160	**0.013** **(0.637)** **N = 5669**
	10:00 a.m. – 11:59 a.m.	-0.035 (0.618) N = 469	0.030 (0.640) N = 4143	0.010 (0.606) N = 700	0.095 (0.592) N = 64	**0.022** **(0.633)** **N = 5376**
	12:00 p.m. – 2:00 p.m.	-0.039 (0.606) N = 126	0.019 (0.750) N = 215	0.035 (0.648) N = 3483	0.047 (0.635) N = 533	**0.034** **(0.650)** **N = 4357**
	At 2:00 p.m. or later	0.074 (0.719) N = 53	-0.150 (0.563) N = 68	-0.015 (0.643) N = 113	0.050 (0.667) N = 1358	**0.037** **(0.664)** **N = 1592**
	TOTAL	**0.003** **(0.633)** **4959**	**0.026** **(0.644)** **N = 5380**	**0.030** **(0.640)** **N = 4540**	**0.049** **(0.660)** **N = 2115**	**0.023** **(0.642)** **N=16994**

Table 4

Mean Differences in Ability Between the First and Section Sections of the Reading Assessment Based on Time of Day Section Completed: Ability on Section 2 – Ability on Section 1 (standard deviations indicated in parentheses)

		Section 1				
		Before 10:00 a.m.	10:00 a.m. – 11:59 a.m.	12:00 p.m. – 2:00 p.m.	At 2:00 p.m. or later	**TOTAL**
Section 2	Before 10:00 a.m.	0.012 (0.613) N = 1576	0.005 (0.628) N = 226	0.112 (0.599) N = 92	-0.107 (0.627) N = 55	**0.012** **(0.615)** **N=1949**
	10:00 a.m. – 11:59 a.m.	-0.063 (0.636) N = 169	0.012 (0.600) N = 1608	-0.059 (0.601) N = 114	-0.157 (0.660) N = 31	**-0.002** **(0.605)** **N=1922**
	12:00 p.m. – 2:00 p.m.	-0.079 (0.597) N = 63	0.036 (0.693) N = 96	-0.005 (0.581) N = 1308	-0.081 (0.554) N = 76	**-0.009** **(0.588)** **N=1543**
	At 2:00 p.m. or later	-0.266 (0.749) N = 21	0.088 (0.564) N = 67	0.063 (0.668) N = 79	0.038 (0.598) N = 587	**0.036** **(0.609)** **N=754**
	TOTAL	**-0.002** **(0.617)** **N = 1829**	**0.015** **(0.607)** **N = 1997**	**0.001** **(0.589)** **N = 1593**	**0.007** **(0.601)** **N = 749**	**0.006** **(0.604)** **N=6168**

Discussion

This study investigated when seventh grade students perform best on statewide mathematics and reading assessments. Based on the sample sizes displayed in Tables 1 and 2, both mathematics and reading teachers tend to start testing on a Monday or Tuesday, and then administer the second section of the test on either a Tuesday or Wednesday. Relatively few students take two sections of a test in one day. It appears that teachers are trying to test at times they believe will facilitate optimal student performance, reinforcing the importance of this research.

The first question investigated in this study compared student performance on the first two sections of the mathematics and the first two sections of the reading assessment. It was hypothesized that no significant difference in scores would exist. The score difference for the

reading test was not statistically significant, and although the difference for mathematics was statistically significant, it was not practically significant.

The second question investigated the best day of the week to assess students. For both the reading and mathematics assessments, the weekday in which the assessment was administered did not significantly affect a student's performance on the test. Looking for general patterns in Tables 1 and 2, the ability levels of students tended to decrease from the Section 1 estimate to the Section 2 estimate when students were administered both sections of the assessment on the same day (though this was not subject to a statistical test). This decrease in estimated ability level was most noticeable (-0.18) for those students who completed both sections of the mathematics test on a Thursday. This pattern was found for both the reading and mathematics assessments. Also, in general, students tended to show the greatest increase in score when Section 2 was administered the day after Section 1 was completed.

In Laird's (1925) research, he noted that students' performances on assessments tended to gradually increase to Wednesday, and then sharply decrease. The patterns found in Table 2 for the reading assessment tend to agree with this outcome. From this data, students had their best test performance on Wednesday, regardless of whether they completed Section 1 or Section 2 that day, and the worst performance was given on Fridays. For the mathematics test, this pattern did not exist. For Section 1, students performed best on Mondays and worst on Wednesdays. For Section 2, students performed best on Wednesdays and worst on Mondays. Since the reading and mathematics results were not consistent with each other, the only overall conclusion that can be drawn from the research presented in this paper is that the day of the week an assessment is administered does not affect a student's performance on that assessment.

The third question addressed in this study was the best time of day to assess students. The day was divided into four intervals: before 10:00 a.m.; 10:00 a.m. – 11:59 a.m.; 12:00 p.m. – 1:59 p.m.; and at 2:00 p.m. or later. It was hypothesized that no difference in student performance on either the reading or mathematics assessment would be found based on the time of day the assessment was completed. The results supported this hypothesis. In light of the nonsignificant findings, Tables 3 and 4 were analyzed for general patterns. For both the reading and mathematics assessment, in general, students tended to perform best on assessments completed at 2:00 p.m. or later and worst on those completed before 10:00 a.m. However, these results are based on general trends and not on significant findings.

The results of the studies by Callan (1999) and Dunn (1985) suggested that students would perform better in the morning than in the afternoon. Ammons, et al. (1995) and Anderson et al. (1991) disagreed and thought that any differences found between performance on assessments and time of day would most likely be related to individual learning styles and perhaps teachers' best instructional time and not on the time of day the assessment was administered. The results of this study fail to support the former research, but tend to support the latter research. The results do not indicate any advantage in students testing in the morning versus the afternoon.

Conclusion

When is the best time for students to take a state assessment? In this study of seventh grade students taking mathematics and reading assessments it does not seem to matter. Based on the results in this study, the time of day or the day of week that seventh grade students complete a state assessment does not affect students' performance on the assessment. Individual preferences may affect individual students' scores; however, when looking at the population of

test takers as a whole, test performance is unrelated to the time of day or day of week the assessment is administered.

Limitations

There are several limitations to this study. First, the sample only included seventh grade students and only included the score difference from the first section to the second section of the assessment. Although similar results would be expected for other grade levels and for comparing the second and third sections of the exam, this was not done in the current research. Second, there are other more powerful methods that could be used to conduct this study. Although repeated runs of ANOVA addressed the research question of the best day of the week to test, this design lacks power. Third, the results from this study only include data from seventh grade students in one Midwestern state and only include one year's worth of data. Similar studies with an expanded testing population would increase the external validity of this study.

Further Research

There are several possible directions to extend this research. First, this study could be repeated with this and subsequent year's data, as well as other grade levels, to determine if the findings generalize. In addition, reasons why no significant differences were found could be investigated. Finally, other research methodologies could be applied to the data that are more powerful in design that that used in the current research.

References

Ammons, T. L., & et al. (1995). *The Effects of Time of Day on Student Attention and Achievement* (ERIC Reproduction Service NO. ED 384 592).

Anderson, M. J., Petros, T. V., Beckwith, B. E., Mitchell, W. W., & Fritz, S. (1991). Individual differences in the effect of time of day on long-term memory access. *American Journal of Psychology, 104*, 241-255.

Callan, R. J. (1999). Effects of matching and mismatching students' time-of-day preferences. *The Journal of Educational Research, 92*(5), 295-299.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Dunn, R. (1985). It's time to handle instructional time correctly. *Early Years K-8, 16*, 47-49.

Laird, D. A. (1925). Relative performance of college students as conditioned by time of day and day of week. *Journal of Experimental Psychology 8*(1), 50-63.

Lievens, F., Buyse, T., Sackett, P.R. (2005). Retest effects in operational selection settings: Development and Test of a Framework. *Personnel Psychology, 58*(4), 981-1007.