

**Report on the Effectiveness of Mathnasium Learning Center Teaching
on Student Performance on Standards-based Mathematics Tests**

2008 Report Addendum

Analysis of Student Testing Data

prepared by

John B. Watson, Ph.D. and the staff of

Watson Education
San Diego, California.
619-299-2255

October, 2008 (Rev. A.)

This report presents additional data analysis for Mathnasium Learning Centers, to complement previous data analysis prepared in Summer 2008. Mathnasium Learning Centers has sought to determine the effectiveness of its program, and has commissioned several qualitative and quantitative studies since 2004.

Study

This study was conducted in 2008, using the results of paired testing data for a cross-section of Mathnasium Center students. Using a single group pre-posttest design, this study was to determine whether there exists a positive treatment effect on mathematics testing performance of elementary and middle school children as a result of their attending the Mathnasium teaching center for a period of approximately 3 months. To see whether students' performance on tests of math skills are improving as a result of Mathnasium tutoring, two math tests were given to students, one at the beginning of the study period (pretest), and one at the end (posttest). The students who participated in this study are Mathnasium customers who also attend local elementary schools.

Students are placed at a specific learning level within the Mathnasium curriculum based on an initial interview at the Mathnasium Center, a review (if possible) of recent testing results from school, and whether or not they are able to pass a pre-test at the level one lower than their current grade. These data are used to place the student at the level where they begin the Mathnasium learning process. This is the level where the subject's pre-test and post-test are given.

The instruments used in this study are the Mathnasium 'Check-up Tests'. These tests have been aligned to math standards from all States in which Mathnasium operates, including the State where Mathnasium is headquartered, California. The pre- and posttests are equivalent, containing the same level and number of questions and testing the same exact skills. The tests have been independently validated by an experienced credentialed mathematics teacher, showing that they test at grade-level and that the content is consistent with the California State standards.

The null hypothesis of this study is that attending the Learning Center will have no positive causal effect on posttest performance. A *t-test* comparing matched pairs of pre- and posttest results was used to statistically determine if there is a significant difference between the two test scores across the study population. A one-tailed test was used because the effect of the Mathnasium tutoring treatment was assumed to be positive.

Analysis

Once the pre and posttest data was collected, the data were entered into Microsoft Excel, and manipulated to combine and match student pairs of data. This information was then sorted by test level. The data was reviewed to determine which grades could serve as the subject of further statistical analysis. Several of the grade level statistics were deemed unreliable due to low sample size. Data for testing levels 1, 5, 6, 7, 8, and 9 were

provided for this study but not included in analysis due to low sample sizes (< 7 students).

Table 1. Statistical results

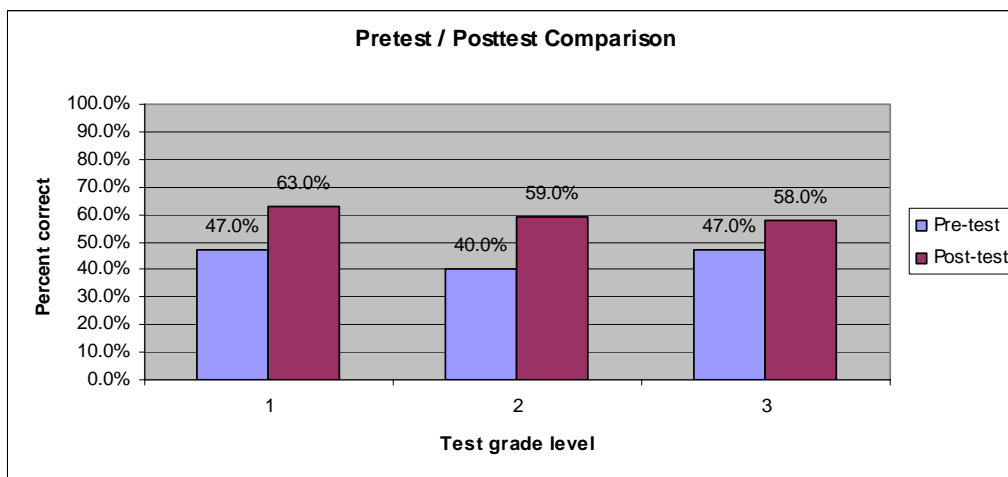
		Mean	Variance	n	Pearson <i>r</i>	df	t Stat	P(T<=t) one-tail	t Critical one-tail (at .05)
Grade 2	Pretest	47%	0.0104	7	0.77	6	-6.1	0.000442	1.94
	Posttest	63%	0.0121	7					
Grade 3	Pretest	40%	0.0179	18	0.50	17	-4.9	6.92E-05	1.74
	Posttest	59%	0.0352	18					
Grade 4	Pretest	47%	0.0275	11	0.86	10	-2.7	0.011989	1.81
	Posttest	58%	0.0663	11					

The *t-test* analysis was performed on the data collected for test levels 2 through 4. The calculations were run using the Data Analysis Add-in statistical functionality within Microsoft Excel. The paired samples were found to correlate highly. A statistically significant difference in the testing scores between pre- and posttest is shown at the 95% confidence level (Table 1).

Conclusion

The statistical results show a positive treatment effect across all grades analyzed. The improvement in mean scores between pre and post-tests across all grades analyzed ranges from 23 to 30%. The students performed significantly better on a math posttest after receiving instruction through the learning center.

Figure 2. A graphical comparison of mean pre- and posttest results.



Appendix A: References

American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME). (1999). *The Standards for Educational and Psychological Testing*.

Kerlinger, F. M. (1973). *Foundations of behavioral research*. New York: Holt Rinehart & Winston.

Mathnasium, LLC. (2004). Results of Parent Satisfaction Survey. (Web Site) URL: www.mathnasium.com.

Trochim, W. (2000). *The Research Methods Knowledge Base*, 2nd Edition. Atomic Dog Publishing, Cincinnati, OH.

Trochim, W. and Land, D. (1982). Designing Designs for Research. The Researcher, 1, 1, 1-6.

Appendix B: Statistical Results

Grade 2

t-Test: Paired Two Sample for Means

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	0.465714	0.632857
Variance	0.010395	0.012124
Observations	7	7
Pearson Correlation	0.768815	
Hypothesized Mean Difference	0	
df	6	
t Stat	-6.099047	
P(T<=t) one-tail	0.000442	
t Critical one-tail	1.94318	
P(T<=t) two-tail	0.000885	
t Critical two-tail	2.446912	

Grade 3

t-Test: Paired Two Sample for Means

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	0.396667	0.59
Variance	0.017929	0.035176
Observations	18	18
Pearson Correlation	0.496801	
Hypothesized Mean Difference	0	
df	17	
t Stat	-4.88856	
P(T<=t) one-tail	6.92E-05	
t Critical one-tail	1.739607	
P(T<=t) two-tail	0.000138	
t Critical two-tail	2.109816	

Grade 4

t-Test: Paired Two Sample for Means

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	0.465455	0.58
Variance	0.027507	0.06632
Observations	11	11
Pearson Correlation	0.859251	
Hypothesized Mean Difference	0	
df	10	
t Stat	-2.65809	
P(T<=t) one-tail	0.011989	
t Critical one-tail	1.812461	
P(T<=t) two-tail	0.023978	
t Critical two-tail	2.228139	