

**Report on the Effectiveness of Mathnasium System
on Student Performance on Standards-based Math Skills Tests**

Report on 2009 Student Testing Data

prepared by

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



A division of Columbus Research, Inc.
San Diego, California.
619-299-2255

March, 2009 (Rev. C.)

Introduction

Mathnasium is a learning center which students may attend after school to boost their math skills. The center is highly specialized, teaching only math. The program is for students in grades 2 through 8 and high school. Students usually attend the center twice per week, for about an hour each visit. The goal is to significantly increase a student's math skills, understanding of math concepts, and overall school performance, while building confidence and forging a positive attitude toward the subject.

The company has sought to determine the effectiveness of its program, and has commissioned several qualitative and quantitative studies since 2004.

This study was conducted in 2009, using the results of paired testing data for a cross-section of Mathnasium Center students. Using a single group non-experimental 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 4 months.

The Study

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 skills. The tests have been validated for content by an expert panel convened by Mathnasium, and their evaluation confirms that the tests are at grade-level and that the content is consistent with the Mathnasium curriculum framework, and the California State standards.

Between the two tests, each student attended the Learning Center a few times per week for mathematics tutoring. The treatment period averages 3 months.

The design of this statistical study is a 'Single Group Pretest-Posttest Design' (Figure 1). This design compares the same group of participants before and after the program. The

purpose of the single group pretest-posttest design is to determine if participants improved after receiving the program. As is common with most any statistical work, there are limitations and threats to this design which are noted in the Recommendations section of this report.

Figure 1. Single Group Research Design based on Kerlinger (1973)

Students at the Mathnasium Learning Center form a single group. The group receives the treatment for a minimal period of three months. O represents the pretest and posttest.

O X O

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 was 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. One of the grade levels was deemed unsuitable due to low sample size (grade 11, 1 student). Thus, data for grade 11 was not used in final calculations. The *t-test* analysis was performed on the data collected for test levels 1 through 12, except for level 11. The calculations were run using the Data Analysis add-in within Excel.

The Pearson *r* correlations were between .70 and .97, indicators that the paired samples correlated highly.

A statistically significant difference in the testing scores between pre- and posttest is shown at the 95% confidence level (Table 1). The null hypothesis was rejected, demonstrating that there was a statistically significant treatment effect between the pre and posttests.

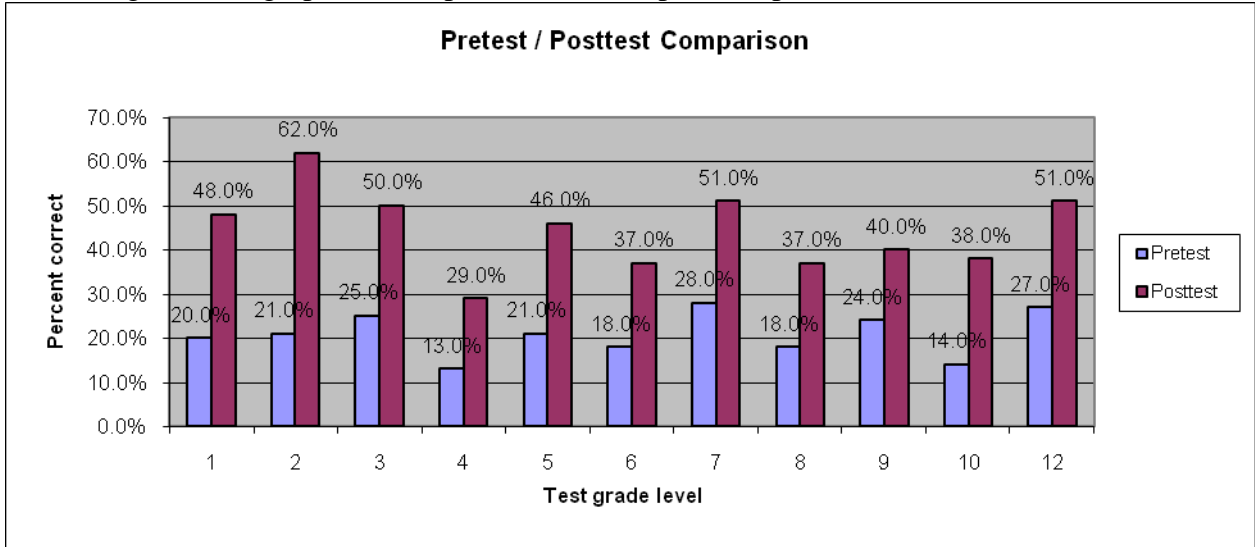
Table 1. Statistical results

		Mean	Variance	n	Pearson <i>r</i>	df	t Stat	P(T<=t) two-tail	t Critical two-tail (at .05)
Grade 1	<i>Pretest</i>	20%	0.10	5	0.74	4	-2.52	0.065	2.78
	<i>Posttest</i>	48%	0.13	5					
Grade 2	<i>Pretest</i>	21%	0.03	6	0.74	5	-7.22	0.001	2.57
	<i>Posttest</i>	62%	0.04	6					
Grade 3	<i>Pretest</i>	25%	0.04	8	0.85	7	-4.22	0.004	2.36
	<i>Posttest</i>	50%	0.09	8					
Grade 4	<i>Pretest</i>	13%	0.00	5	0.97	4	-9.56	0.001	2.78
	<i>Posttest</i>	29%	0.01	5					
Grade 5	<i>Pretest</i>	21%	0.04	9	0.70	8	-4.64	0.002	2.31
	<i>Posttest</i>	46%	0.05	9					
Grade 6	<i>Pretest</i>	18%	0.02	17	0.77	16	-5.94	0.000	2.12
	<i>Posttest</i>	37%	0.04	17					
Grade 7	<i>Pretest</i>	28%	0.07	18	0.83	17	-5.93	0.000	2.11
	<i>Posttest</i>	51%	0.08	18					
Grade 8	<i>Pretest</i>	18%	0.04	16	0.88	15	-4.86	0.000	2.13
	<i>Posttest</i>	37%	0.09	16					
Grade 9	<i>Pretest</i>	24%	0.06	4	0.96	3	-2.75	0.071	3.18
	<i>Posttest</i>	40%	0.11	4					
Grade 10	<i>Pretest</i>	14%	0.04	5	0.92	4	-4.88	0.008	2.78
	<i>Posttest</i>	38%	0.07	5					
Grade 12	<i>Pretest</i>	27%	0.11	4	0.85	3	-2.68	0.075	3.18
	<i>Posttest</i>	51%	0.09	4					

Conclusion

The statistical results show a positive treatment effect. The difference in mean percent correct score between pretest and posttest ranges from 20% to 40% (see Figure 1). The students performed significantly better on a math posttest after receiving instruction through the learning center.

Figure 1. A graphical comparison of mean pre- and posttest results.



Recommendations

While these results show a positive treatment effect, there are a number of threats to the statistical results.

- The statistical design of this study has inherent limitations, namely participants may improve over time without intervention of any kind, and these changes can be mistakenly attributed to the program under evaluation. This design could not indicate whether the program solely caused improvement in participants; as there is no way to distinguish between changes over time due to other factors and effects specific to the program. If the study had a true experimental design, featuring randomly assigned subjects to control and experimental groups, the analysis could more accurately isolate the treatment effect of the learning center.
- The sample size was very small in this study. It is recommended that the center conduct additional studies using larger numbers of students. To produce reliable statistics, the minimum size of the groups ought to be a minimum of 20 subjects per group; of course, the larger the group, the better.
- The length of the student treatment is another threat to the study, because 4-6 months treatment duration likely overlaps time when students are attending regular math classes in school. It may be possible to reduce this effect by running studies for students whose pretest, treatment and posttest take place over a time frame like a summer break.

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. (2008). 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.