

# Marygrove College Institute for Arts Infused Education Pretest Reading Assessment

Shlomo S. Sawilowsky

September 3, 2009

A reading pretest and posttest was developed from November, 2008 through February, 2009. It is based on the 3rd grade Dolch vocabulary. The instruments were constructed based on the blueprint approach to test construction. Teachers from the five elementary schools participated in developing the test blueprint, where each topic was sequenced and then broken down into sub-competencies. The teachers weighted the importance of each entry in the blueprint. Next, a table of specifications was used to assign the cognitive level to each skill (i.e., recall, application, problem solving). A task analysis was completed to differentiate prerequisite and enabling skills from the sub-competencies, of which only the latter appear on the test.

A 250 word vignette serves as the reading stimulus. Responses were created as test items in multiple choice, fill in the blank, and matching formats. Fifteen of the items assess the cognitive reading skills, and the remaining two items pertain to the affective domain (e.g., students' feelings regarding reading), for both the pretest and the posttest.

The participating schools and sample size are depicted in Table 1.

School	Frequency	Percent
Timbuktu	40	17.2
Golightly	48	20.7
George Crockett	43	18.5
Maybury	52	22.4
Academy of Americas	49	21.1
Total	232	100.0
Total	239	

SPSS 17 and MicroCat 3.5 software were used to input and cleanse the data. Spearman Brown, a measure of internal consistency reliability, for the 15 item reading skills subscale was .82. The reliability estimate for the 2 item affective subscale was .86. Thus, both subscales have excellent reliability evidence.

At the pretest, there were N=232 students. The mean score was 66.2% ( $\sigma = .18$ ). A passing score of 70% was obtained by only 56.4% (131/232) students, of whom 2 (< 1%) scored

100%. At the posttest, there were N = 172 students. The mean score was 82.7% ( $\sigma = .15$ ). A passing score of 70% was obtained by 86% (148/172) students, of whom 33 (19.2%) obtained perfect scores.

A paired samples t test was conducted to confirm the post test scores were statistically significantly improved based on the AIE intervention. The obtained t was 9.46,  $df = 167$ ,  $p < .000$ . **The estimated effect size (a measure of the strength of the intervention) was .88, which Cohen (1988) identifies as a “large” treatment effect. Thus, the AIE intervention has brought about a huge difference in students’ mastery of the 3<sup>rd</sup> grade Dolch vocabulary.** A breakdown of Pretest to Posttest scores for each school in Tables 2 – 6.

1. Timbuktu. The Posttest – Pretest difference was 2.1, which is a statistically significant improvement ( $t=4.792$ ,  $df = 29$ ,  $p = .000$ ).

	Mean	N	Std. Deviation	Std. Error Mean
Posttest	13.2667	30	3.02784	.55281
Posttest	11.1633	30	2.55228	.46598

2. Golightly. The Posttest – Pretest difference was 3.63, which is a statistically significant improvement ( $t=6.14$ ,  $df = 24$ ,  $p = .000$ ).

	Mean	N	Std. Deviation	Std. Error Mean
Posttest	14.2000	25	2.67706	.53541
Posttest	10.5740	25	2.49124	.49825

3. George Crockett. The Posttest – Pretest difference was 1.26, which is not a statistically significant improvement ( $t=1.61$ ,  $df = 24$ ,  $p = .12$ ). Note, however, that the comparatively small sample size is a major factor in depressing the statistical power. The more powerful nonparametric Wilcoxon Rank-Sum test, which was also not statistically significant, had a p value of .09, which is much closer to the  $\alpha = .05$  standard.

Table 4. George Crockett				
	Mean	N	Std. Deviation	Std. Error Mean
Posttest	13.1600	25	2.67208	.53442
Pretest	11.9000	25	2.76522	.55304

4. Maybury. The Posttest – Pretest difference was 2.96, which is a statistically significant improvement ( $t=6.625$ ,  $df = 42$ ,  $p = .000$ ).

Table 5. Maybury				
	Mean	N	Std. Deviation	Std. Error Mean
Posttest	14.9767	43	1.88337	.28721
Pretest	12.0186	43	2.67884	.40852

6. Academy of Americas. The Posttest – Pretest difference was 1.65, which is a statistically significant improvement ( $t=3.47$ ,  $df = 44$ ,  $p = .001$ ).

Table 6. Academy of Americas				
	Mean	N	Std. Deviation	Std. Error Mean
Posttest	14.1556	45	2.20491	.32869
Pretest	12.5044	45	2.71603	.40488

The Pretest to Posttest increase for all schools is depicted in Figure 1.

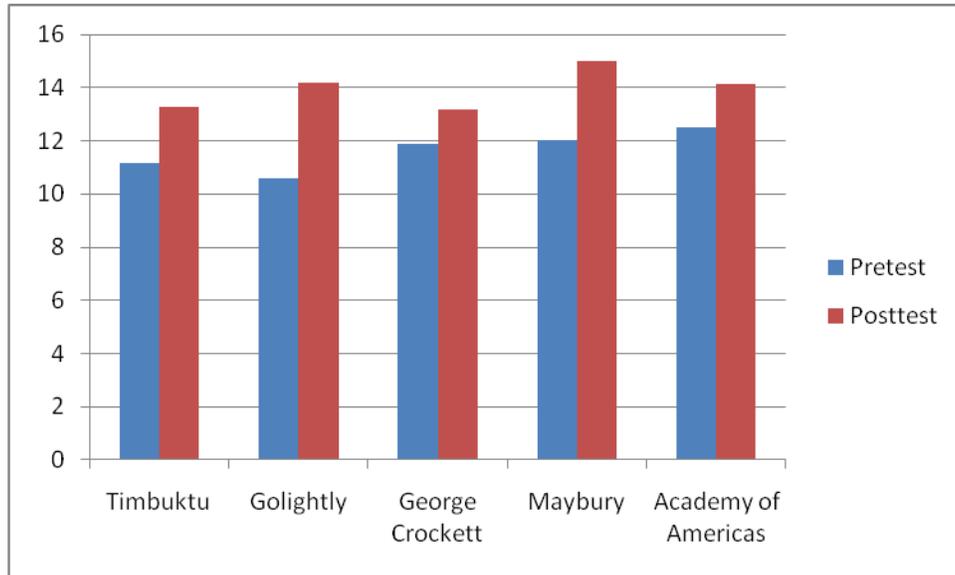


Figure 1