Evaluation Matters

Volume 5, Number 2 October 2015 Aleksandr Shneyderman, Ed.D.

Evaluation of the Imagine Learning Program Used with English Language Learners

Introduction

Imagine Learning is a computer-based instructional program used with English Language Learners (ELLs) in the District. It is designed to provide instruction specifically to ELLs. "Imagine Learning provides first-language support in 15 languages to facilitate and enhance ELL learning. Imagine Learning is an engaging language and literacy software program that accelerates English learning. Focused on oral language, academic vocabulary, instruction in the five essential components of reading, and strategic first-language support—it spells success for students everywhere," according to the company's website.

Evaluation Design

This section describes sampling procedures used to select students for the Program and Comparison Samples. In addition, it addresses the outcome measures used in the evaluation and describes the data analyses performed.

Sampling

During the 2014-2015 school year, the Imagine Learning (IL) program was implemented across the District, mostly for students at the initial level of participation in the English for Speakers of Other Languages (ESOL) program. However, not all schools started participating in it at the beginning of the school year. This fact and the almost universal implementation of the program presented a challenge in selecting acceptable samples of students participating and not participating in the program. Students who started participating in the program no later than November of 2014 and accumulated at least 20 hours of program participation were selected for potential inclusion in the Program Sample. The criterion of a minimum of 20 hours of program participation was recommended by the Imagine Learning staff as an indicator of the fidelity of implementation.

In the second step, the District records of students in grades K-12 were used to determine which students did not participate in the IL program during the 2014-2015 academic year. These students were selected for potential inclusion in the Comparison Sample.

Office of Assessment, Research, and Data Analysis
1450 NE 2nd Avenue, Suite 208, Miami, Florida 33132
(305) 995-2943 Fax (305)995-7521

In the third step, the students selected in the previous two steps were split into two groups. The Continuing ELL group consisted of students who participated in the Comprehensive English Language Learning Assessment (CELLA) in the spring of 2014. The New ELL group included students who started ESOL participation in the fall of 2014; accordingly, they had the results on the online CELLA used in the District for ESOL placement.

Finally, for each of the two groups of ELL students defined above, a multivariate matching algorithm was used to find comparison students who would match the selected program students in grades K-12 exactly on their grade levels while minimizing the multivariate distance between program and comparison students on the 2014 CELLA scale scores in Listening/Speaking, Reading, and Writing. Only the records of students with complete sets of 2014 CELLA scores were used in this process. The students selected in this step comprised the Program and Comparison Samples, each consisting of 354 students. These samples were split based on the ELL group and grade level cluster corresponding to the level of the CELLA.

The demographic and academic achievement characteristics of the two samples as well as the mean number of hours of program participation during the 2015-2015 academic year are shown in Table 1 for each of the two ELL groups and for each of the four grade-level clusters.

Table 1 shows that the student samples were reasonably well matched in terms of the percentages of students eligible for the federal free/reduced price lunch (FRL) program. In addition, students in the two samples were well matched on their 2014 CELLA scores in all three modalities. On the other hand, percentages of students participating in the Special Education (SPED) programs were somewhat different in some cases. These differences were taken into account statistically as explained in the *Data Analysis* section.

Table 1

Demographic and English Language Acquisition Characteristics of the two Samples

	Continuing ELL		New ELL	
	Program	Comparison	Program	Comparison
Grades K-2	n = 44	n = 44	n = 9[a]	n = 9
Mean Hours in the Program	35.8	--		
Percentage of Students who are				
Eligible for the FRL program	100	93		
SPED	9	16		
Mean Scale Scores on the 2014 CELLA (SD)[a]				
Listening/Speaking	629 (29)	629 (28)		
Reading	482 (97)	482 (97)		
Writing	584 (44)	584 (44)		
Grades 3-5	n = 61	n = 61	n = 55	n = 55
Mean Hours in the Program	35.2	--	37.2	--
Percentage of Students who are				
Eligible for the FRL program	97	89	93	78
SPED	0	16	2	2
Mean Scale Scores on the 2014 CELLA (SD)				
Listening/Speaking	642 (52)	643 (51)	535 (30)	535 (33)
Reading	628 (61)	628 (62)	570 (28)	570 (29)
Writing	646 (36)	647 (36)	552 (21)	552 (23)
Grades 6-8	n = 32	n = 32	n = 35	n = 35
Mean Hours in the Program	31.2	--	36.2	--
Percentage of Students who are				
Eligible for the FRL program	81	100	86	80
SPED	13	22	0	0
Mean Scale Scores on the 2014 CELLA (SD)				
Listening/Speaking	619 (42)	618 (43)	532 (64)	530 (61)
Reading	658 (39)	657 (39)	642 (34)	643 (35)
Writing	652 (34)	653 (34)	563 (41)	562 (40)
Grades 9-12	n = 84	n = 84	n = 34	n = 34
Mean Hours in the Program	34.0	--	38.7	--
Percentage of Students who are				
Eligible for the FRL program	93	82	90	88
SPED	0	4	1	1
Mean Scale Scores on the 2014 CELLA (SD)				
Listening/Speaking	640 (34)	642 (34)	543 (47)	543 (46)
Reading	692 (45)	693 (45)	596 (76)	596 (76)
Writing	673 (23)	672 (22)	563 (34)	563 (33)

[a] When the number of students is small (<30), the statistics are not shown because the statistical analysis for this subgroup was not carried out. [b] SD represents Standard Deviation.

Outcome Measures

Student results on the 2015 CELLA were used to examine the effects of the program on students' English language acquisition. CELLA is a four-skill language proficiency assessment that tests ELL students' listening, speaking, reading, and writing skills. The results are provided as scale scores in the three domains: Listening/Speaking), Reading and Writing.

Data Analysis

The General Linear Model (GLM) was used to compare the 2015 CELLA mean scale scores for students in the Program and Comparison Samples. The numbers of students whose CELLA results were included in the statistical analyses were slightly different from those in Table 1 above because not all students had their 2015 CELLA results. The analyses were carried out separately for each of the two ELL groups, four grade-level clusters, and the three areas in which CELLA scale scores are reported: Listening/Speaking, Reading, and Writing. The 2014 CELLA scores in each modality and students' SPED status were used as covariates. An attempt to use students' free/reduced price lunch (FRL) status as a covariate was made. However, this variable was found not to be a statistically significant predictor of the 2015 CELLA outcomes once the 2014 CELLA scores and SPED status were used in the model as covariates. Consequently, the FRL status was removed from the final model. A program participation dichotomous indicator was used as a fixed factor in the GLM process. All separate analyses were carried out at the .05 level of statistical significance.

Results

The results of the statistical analyses are presented in Table 2. The adjusted mean scale scores whose differences were found to be statistically significant are shown in bold.

The results of the analyses of the IL program shown in Table 2 indicate that in most comparisons, the differences between the 2015 CELLA adjusted mean scale scores for Program and Comparison Samples were not statistically significant. In one case, these differences were found to be statistically significant, but not in favor of the program.

Table 2

Results of the GLM Analysis

Adjusted 2015 CELLA Mean Scale Scores	Continuing ELL		New ELL	
	Program	Comparison	Program	Comparison
Grades K-2				
Listening/Speaking	657	660		
Reading	606	594		
Writing	635	637		
Grades 3-5				
Listening/Speaking	690	689	640	651
Reading	681	684	**655**	**677**
Writing	675	683	654	667
Grades 6-8				
Listening/Speaking	676	678	669	659
Reading	703	696	702	709
Writing	690	682	675	683
Grades 9-12				
Listening/Speaking	691	694	692	684
Reading	722	725	729	721
Writing	696	691	698	693

Discussion

In this evaluation, the results of only those students in the program who completed at least 20 hours of the program's lessons were included in the analyses. This was done in an effort to include in the Program Sample only those students for whom there was a minimum "dosage" of exposure to the program activities. Students in the Program Sample were matched with students in the Comparison Sample based on their 2014 English language acquisition results as well as certain demographic characteristics. The 2015 CELLA results were used to assess the potential program effects.

The results of the statistical analyses indicate that program students did not outperform comparison students (in the sense of statistically significant differences) in any of the comparisons. Of course, it is possible that some of the program's potential positive effects were not measured by CELLA. This can happen if the IL curriculum is not matched to the standards assessed by CELLA.

In addition, it is possible and even likely that students who participated in the IL instructional program would be more familiar with and have more positive attitudes toward computer technology in general and learning through computer technology in particular. These possible positive effects were not assessed by this evaluation.