

An Evaluation of the 2009-2010
Step By Step Learning® Implementation of Services
in the Hazleton Area School District

Edward S. Shapiro, Ph.D.

and

Virginia R. Hampton, Ph.D.

Center for Promoting Research to Practice

Lehigh University

Technical Report # 2010-5



December 5, 2010

Executive Summary

The purpose of this evaluation was to determine whether the Step By Step Learning® comprehensive approach to literacy was effective in improving the literacy skills of kindergarten and first grade children in the Hazleton Area School District (HASD). This evaluation examined the outcomes of students in the HASD as compared to students in a school district that did not receive the same level of services from Step By Step Learning® (SBSL) and sought to answer three questions:

1. Were the trainings implemented with fidelity?
2. Did the groups have differences in DIBELS scores in kindergarten and first grade?
3. Did the groups have differences in instructional recommendations in kindergarten and first grade?

SBSL provided multiple services to the HASD during the 2009-2010 school year, which was the initial year of a multi-year effort to improve the literacy skills of students in the district. The services included training in the administration of assessment measures, professional development to help teachers learn how to teach literacy skills, and assistance with implementing effective classroom instruction. This approach employs coaching and mentoring throughout the year, and SBSL associates work closely with teachers both individually and in small groups.

The outcomes of students in the HASD in 2009-2010 were compared to student outcomes in selected schools from the Allentown School District during the 2008-2009 school year. The criteria for selecting a comparison group included: (1) receiving substantially less services during the first year of working with SBSL; (2) having similar third grade PSSA reading results prior to beginning with SBSL; and (3) having similar amounts of students receiving free and reduced lunch. Based on these criteria, seven schools were selected from the ASD.

The only service that SBSL provided to the ASD during 2008-2009 was DIBELS training; teachers did not receive any form of instructional support from SBSL. Among the third grade students in the selected ASD schools during the year before SBSL began, the percentages of students receiving scores of *Advanced/Proficient* and *Basic/Below Basic* on the PSSA reading assessment were within 11% of the HASD results. In the ASD, 67% of students had scores within the *Advanced/Proficient* range and 33% of students had scores of *Basic/Below Basic*, compared to 78% and 22%, respectively, for the HASD. The percentage of students receiving free and reduced lunch was within 10% of the HASD, with 73% in Allentown School District and 63% in the Hazleton School District.

The literacy skills of both groups were examined on the DIBELS (6th edition) measures. In kindergarten, the DIBELS measures included Initial Sound Fluency (ISF), Letter Naming Fluency (LNF), Phoneme Segmentation Fluency (PSF) and Nonsense

Word Fluency (NWF). ISF was administered in the fall and winter, and LNF was administered in the fall, winter and spring. Both PSF and NWF were given in the winter and spring. The first grade measures included LNF, PSF, NWF, and Oral Reading Fluency (ORF). LNF was only administered in the fall, and ORF was administered in the winter and spring. Both PSF and NWF were administered in the fall, winter and spring. The instructional recommendations of *Intensive*, *Strategic* and *Benchmark* were also examined to assess students' levels of risk.

Observations of randomly selected training sessions found that overall, the observed sessions were conducted with fidelity. This indicates that in general, the sessions were implemented as intended. Note that information provided in the trainings is presented in multiple ways across the year, such as through the initial training, refresher trainings, and individual coaching and mentoring sessions. Therefore, school district staff had multiple opportunities to learn any missed information when a particular training session did not include all of the intended material.

The first set of analyses examined students' DIBELS scores in kindergarten and first grade. Among the kindergarten students, the intervention group demonstrated significantly higher skills in initial sound fluency, phoneme segmentation and letter naming fluency. Although the intervention group had significantly higher scores in nonsense word fluency at the winter assessment, there were no differences between the groups by the end of the year. In first grade, analyses indicated that students in the intervention group had significantly higher skills in phonological awareness and letter-sound correspondence. There were no significant differences in oral reading fluency.

Another set of analyses examined the instructional recommendations of the DIBELS, which indicate students' levels of risk. With both kindergarten and first grade, the students from the intervention group had less risk in their literacy skills at the end of the year than did students in the comparison group. Specifically, the intervention group had more students at *Benchmark* and fewer students at *Intensive* than found in the comparison group. Although the amount of change in the percentage of first grade students at *Intensive* was nearly the same for both the intervention and comparison groups, the intervention group had a higher amount of change in students at *Benchmark*.

The results of this evaluation provide support for the effectiveness of the SBSL comprehensive approach to literacy in enhancing the literacy skills of children in kindergarten and first grade. Students in both grades demonstrated significantly higher scores in selected literacy areas, compared to students from a school district that did not receive comprehensive services from SBSL. In addition, students from the HASD ended the year with less risk in their literacy skills. These findings indicate that one year of this intervention can result in improved outcomes for students.

Introduction

Step By Step Learning® (SBSL) implemented a comprehensive approach to literacy in the Hazleton Area School District during the 2009-2010 school year as part of a multi-year effort to improve the literacy skills of students in the district. Through this intervention, SBSL provided services to teachers, specialists and administrators to enhance the literacy skills of children in kindergarten through second grade across seven elementary schools. The services included training in the administration of assessment measures, professional development to help teachers learn how to teach literacy skills, and assistance with implementing effective classroom instruction. This approach employs coaching and mentoring throughout the year, and SBSL associates work closely with teachers both individually and in small groups.

An evaluation of this intervention was conducted to examine the effectiveness of the SBSL services in the Hazleton Area School District (HASD) after one year of implementation. To help determine whether the intervention was related to student outcomes, rather than what would have occurred due to time, instruction and normal activities across the school year, selected schools within the Allentown School District (ASD) comprised the comparison group. ASD was selected as a comparison based on several criteria: (1) minimal services from SBSL; (2) comparable literacy skills among the students as indicated by third grade PSSA results; and (3) similar percentages of students receiving free and reduced lunch. Student outcomes as measured by the DIBELS (Dynamic Indicators of Basic Early Literacy Skills, 6th edition) were compared between the two school districts for kindergarten and first grade students; no second grade DIBELS scores were collected within the ASD.

Measures

The measures used to assess student outcomes were the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (6th edition). The DIBELS measures included: Initial Sound Fluency, Letter Naming Fluency, Phoneme Segmentation Fluency, Nonsense Word Fluency and Oral Reading Fluency. Initial Sound Fluency (ISF) measures phonological awareness through examining children's skill in recognizing the beginning sound in words. Letter Naming Fluency (LNF) examines students' ability to identify uppercase and lowercase letters. Phoneme Segmentation Fluency (PSF) measures phonological awareness, and Nonsense Word Fluency (NWF) measures letter-sound correspondence. Oral Reading Fluency (ORF) examines the reading of connected text in grade level passages and serves as a measure of general proficiency in reading. The DIBELS were administered in the fall, winter and spring. In kindergarten, the fall assessments included ISF and LNF and the winter assessments included ISF, LNF, PSF, and NWF. Assessments in the spring included LNF, PSF and NWF. The DIBELS also provide an indication of students' risk levels, based on their scores across all the measures at each assessment point. These recommendations include Benchmark (Low Risk status), Strategic (Some Risk status) and Intensive (High Risk status).

The Comprehensive Approach to Literacy in the HASD

The comprehensive approach to literacy employed by Step By Step Learning® included multiple services provided to teachers, specialists and administrators across the year. These services were provided through large and small group training sessions, “refresher” training sessions later in the year, and modeling and coaching sessions with teachers in their classrooms. A primary aspect of the SBSL approach is to provide an “I do, we do, you do” model, in which school staff receive extensive training, modeling and coaching to eventually implement this approach to intervention on their own. Another aspect of this approach is to involve the administrators in the trainings, so that they understand and can help facilitate this approach to intervention. Throughout the year, the SBSL associates worked closely in the schools with the teachers, specialists and administrators to provide the services described below as well as to problem-solve as needed.

School district staff received training in the administration of DIBELS, allowing teachers to conduct assessments of students’ literacy skills and provide progress monitoring across the year. Teachers also were trained in using the DIBELS data and informal diagnostic assessments to inform the development of appropriate instructional plans and intervention groups. In addition, teachers learned how to monitor student’s progress and to modify instructional plans and intervention groups when needed.

Another aspect of the approach included professional development to help teachers learn how to teach children literacy skills. SBSL provided training in the Language Essentials for Teachers of Reading and Spelling (LETRS) program, developed by Louisa C. Moats (2005, Sopris West Educational Services). This professional development program consists of multiple modules that focus on different components of reading instruction. During the first year in Hazleton, school staff received training in the following modules: (1) The Challenge of Learning to Read; (2) The Speech Sounds of English: Phonetics, Phonology, and Phoneme Awareness; and (3) Spellography for Teachers: How English Spelling Works. Separate training sessions also provided assistance with connecting the information learned from each LETRS module to use in the classroom.

SBSL associates also worked with teachers in their classrooms, to help teachers apply the information learned from the trainings directly to their students. The associates modeled appropriate instructional intervention activities with small groups of students and later provided coaching to the teachers when they demonstrated their small group instruction to the associates. This modeling and coaching occurred across the year.

School District Comparisons

Several criteria informed the selection of the comparison group: (1) differences in the amount of SBSL services; (2) comparable literacy skills before SBSL began working with the school district; and (3) similarity in the percentage of students receiving free and reduced lunch. Allentown School District was selected based on these criteria. ASD is a

much larger school district than the HASD, so seven schools within the district were selected to help meet these criteria without having a disproportionate number of students between the two groups. Among these schools, four serve grades kindergarten through fifth, two serve grades one through five, and one serves prekindergarten and kindergarten.

The first criterion was that the comparison school district receives minimal SBSL services. By comparing a school district that receives comprehensive services from SBSL (the Hazleton Area School District) to a school district that receives few if any services, results can be interpreted with more certainty that any differences in outcomes reflect the intervention. During the first year that SBSL worked with ASD (2008-2009), SBSL provided DIBELS training to kindergarten and first grade teachers. Teacher learned how to administer the DIBELS with the use of handheld personal digital devices (PDA). Students in the ASD had DIBELS assessments, but teachers did not have any other professional development from SBSL.

The second criterion for selection of a comparison group was to have comparable literacy skills to the intervention group, as indicated by PSSA reading results. The PSSA scores were compared for third grade students in the year before SBSL began providing services. The goal was to have school district scores for the combined *Advanced/Proficient* and for the combined *Basic/Below Basic* categories that were within 10% of one another. Table 1 provides a summary of the PSSA results for each school district. Note that the Allentown School District results are for six of the seven schools selected as best meeting the criteria for PSSA results and free/reduced lunch because one school does not have third grade students.

Table 1. PSSA Results for Third Grade Students Prior to SBSL Services to District

School District	PSSA Year	Percentage of Students at <i>Advanced/Proficient</i>	Percentage of Students at <i>Basic/Below Basic</i>
Hazleton Area School District	2008-2009	78%	22%
Allentown School District (selected schools)	2007-2008	67%	33%

As the table indicates, the *Advanced/Proficient* scores for HASD and the selected schools in ASD were within 11 percentage points of one another. The difference between the two groups was also 11% for the *Basic/Below Basic*. This difference was considered adequate, given the importance of having a comparison group that also had few services and had similar demographic data, and the difficulty of finding school districts that met all three criteria.

The last criterion for selection of the comparison group was to have similar percentages of students receiving free and reduced lunch. Again, the goal was to have school districts within 10% of one another. In HASD, 63% of the students received free

and reduced lunch, whereas the figure was 73% among the seven selected schools within the Allentown School District. Therefore, the selected schools within the ASD meet this demographic criterion.

Evaluation Questions

The purpose of the evaluation was to determine the effectiveness of one year of the comprehensive approach to literacy implemented in the Hazleton Area School District. This included an evaluation of several aspects of the intervention. First, the evaluation assessed whether the trainings were implemented in the manner in which they were intended (that is, whether they were conducted with fidelity). If the intervention was not administered with fidelity, interpretations of its results would be limited. Second, the evaluation examined student literacy skills in kindergarten and first grade, to compare the outcomes of students. Finally, the evaluation examined the risk levels of students in the two school districts, to determine whether students in the intervention group had less risk in their literacy skills by the end of the year.

Specifically, three questions were examined:

1. Were the trainings implemented with fidelity?
2. Did the groups have differences in DIBELS scores in kindergarten and first grade?
3. Did the groups have differences in instructional recommendations in kindergarten and first grade?

Evaluation Question 1: Were trainings implemented with fidelity?

To determine whether trainings were implemented in the manner in which they were intended (i.e., with fidelity), observations were conducted of randomly selected training sessions. The checklists used for the observations were developed to reflect the information that the presenter planned to include in the session. For the large group sessions, the items on the checklist corresponded to the Powerpoint slides, to indicate whether all of the topics, corresponding information, and activities were presented, as well as whether the materials and handouts were provided. The small group sessions did not utilize slides, so the items on the checklists corresponded to the specific topics, corresponding information and activities that were meant to be presented, including the necessary materials and handouts.

As seen in Table 2, the fidelity for many of the observed sessions ranged from 93% to 100%. This indicates that the sessions were typically implemented as planned. Among the sessions with fidelity below 90%, the presenters did not have enough time to cover all the material and certain sections of the training were not presented, which resulted in lower ratings of fidelity. However, when only examining the sections that were presented, the fidelity was higher, indicating that the presenters were following the intended format for the trainings until they ran out of time. Note that information provided in the trainings is presented in multiple ways across the year, such as through the initial training, refresher trainings, and individual coaching and mentoring sessions. Therefore, school district staff had multiple opportunities to learn any missed information.

Table 2. Fidelity For Observed Sessions

Type of Session	Fidelity
DIBELS	100%
DIBELS	94%
DIBELS	93%
LETRS	69%*
LETRS Connect to Classroom	76%*
Literacy Leadership	72%
Data analysis/Informal diagnostics	100%
Instructional Planning	100%

* Certain sections of the training were not presented due to time constraints.

Overall, the observed sessions were conducted with fidelity. The mean percentage across the observed sessions is 88%, which indicates that on average, the sessions were conducted as planned.

Evaluation Question 2: Did the groups have differences in DIBELS scores in kindergarten and first grade?

To determine the effectiveness of the intervention provided by Step By Step Learning®, students' DIBELS outcomes were examined across the year. Comparisons were made between students who participated in the intervention (the intervention group) and students who did not participate in the intervention (the comparison group).

Kindergarten. The kindergarten DIBELS scores were examined across the year. For kindergarten, the DIBELS measures include ISF, LNF, PSF, and NWF. Table 3 presents the means and standard deviations for each measure across the year. Students from both groups began the year with similar skills in Initial Sound Fluency, and the intervention group students had slightly higher skills in Letter Naming Fluency. By the end of the year, students in the intervention group had higher scores on LNF and PSF. However, both groups had the same score for NWF at the end of the year. Figures 1-4 provide graphic displays of the results across the year.

Table 3. Kindergarten DIBELS Scores By Group

Measure	Intervention			Comparison		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Initial Sound Fluency						
<i>Fall ISF</i>	732	7.98	7.270	623	7.40	6.930
<i>Winter ISF</i>	763	28.05	17.645	652	17.71	11.162
Letter Naming Fluency						
<i>Fall LNF</i>	732	13.38	14.081	623	10.67	13.323
<i>Winter LNF</i>	763	35.24	17.962	652	26.12	17.336
<i>Spring LNF</i>	762	47.71	16.918	651	42.97	17.797
Phoneme Segmentation Fluency						
<i>Winter PSF</i>	762	18.35	13.629	652	10.88	10.895
<i>Spring PSF</i>	762	42.20	14.791	651	36.85	18.243
Nonsense Word Fluency						
<i>Winter NWF</i>	762	18.96	15.265	649	10.50	11.082
<i>Spring NWF</i>	762	33.61	16.948	651	33.82	19.697

Figure 1. Kindergarten ISF Scores By Group

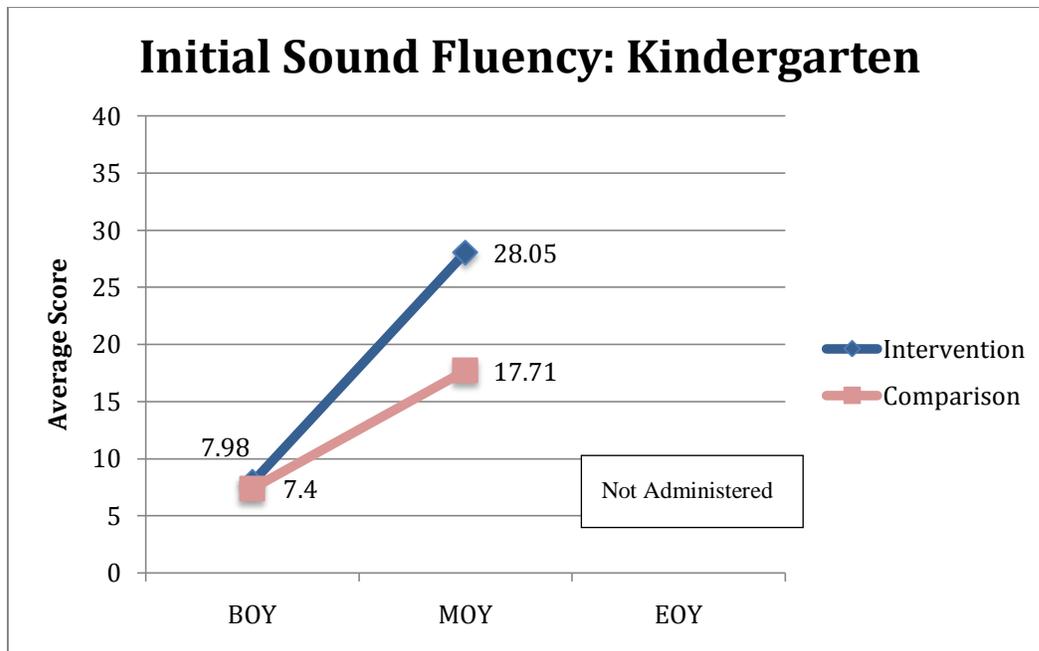


Figure 2. Kindergarten LNF Scores By Group

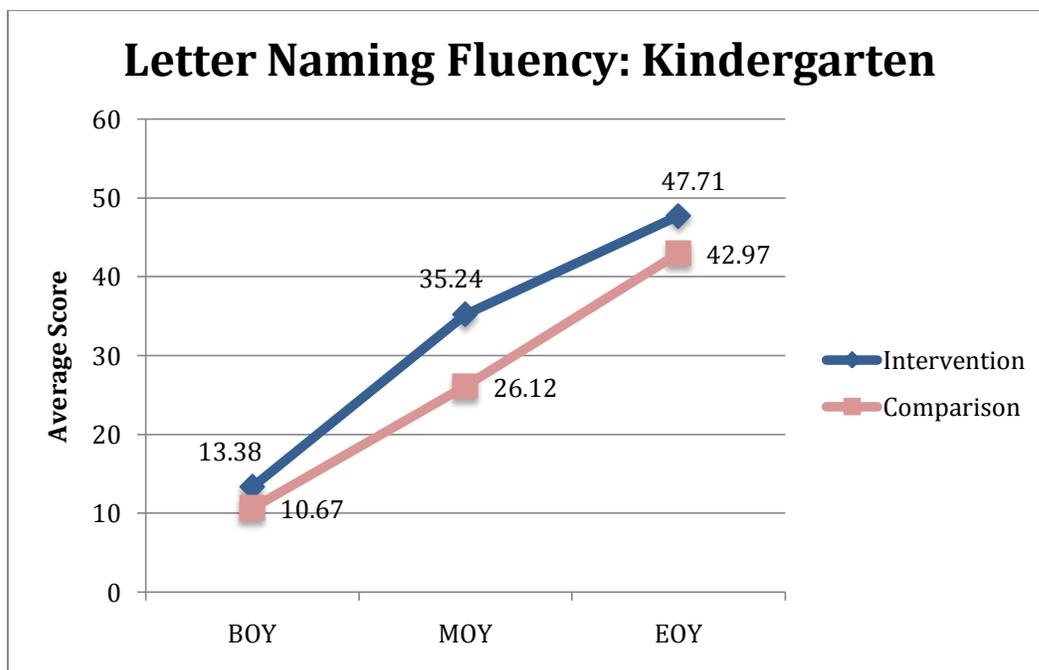


Figure 3. Kindergarten PSF Scores By Group

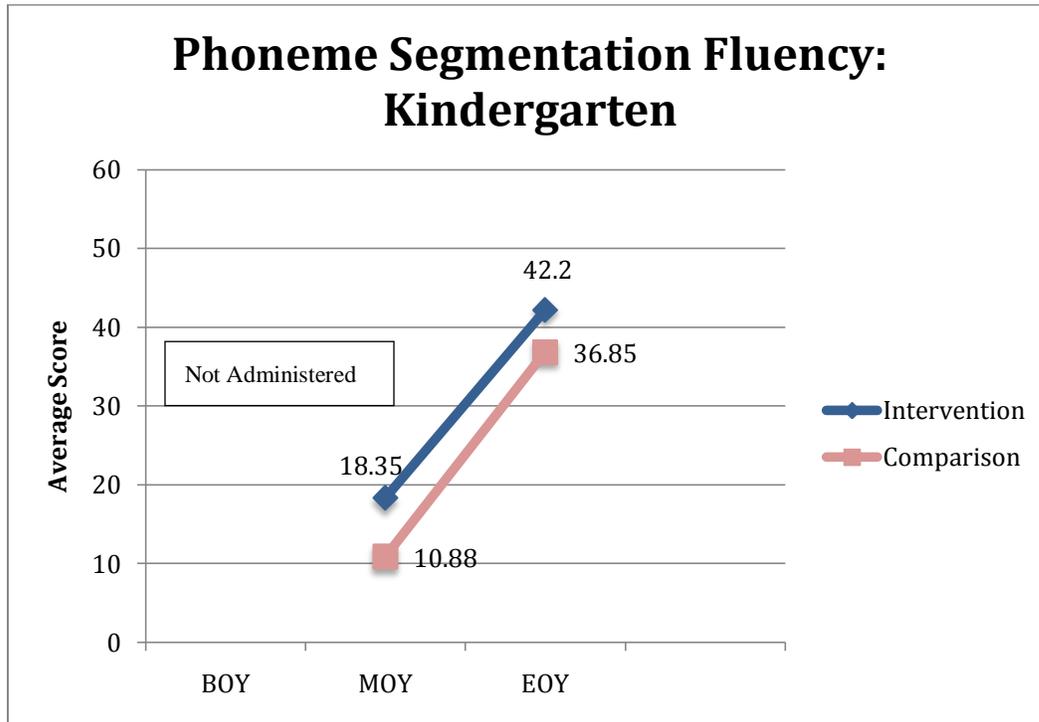
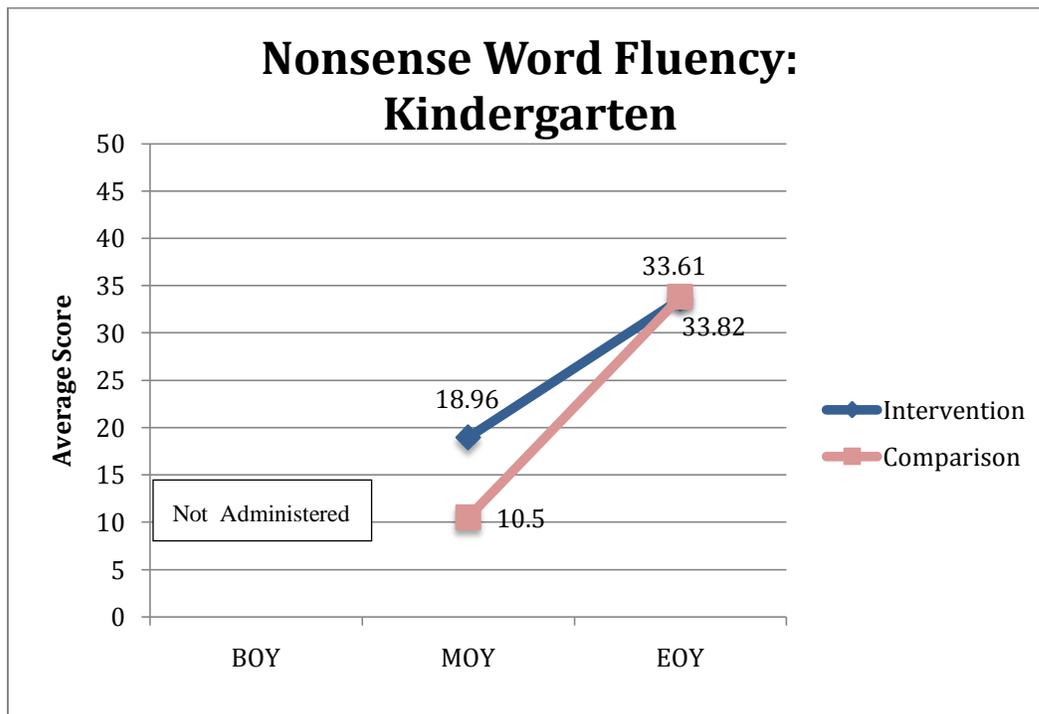


Figure 4. Kindergarten NWF Scores By Group



Repeated measures Analysis of Variance (ANOVA) analyses were conducted for ISF, LNF, PSF, and NWF to determine whether there were any significant differences between the groups in their scores over time. Significant differences were found for all four measures. These results were explored further by conducting t-tests at each DIBELS assessment point (fall, winter, and spring). For Initial Sound Fluency, significant differences were found for the winter score, but not for the fall scores. For Phoneme Segmentation Fluency, the differences were significant at both winter and spring. Nonsense Word Fluency scores were significant at the winter but not at the spring assessment points. The Letter Naming Fluency scores were significantly different at all three assessment points, so additional analyses further explored these results. Separate Analysis of Covariance (ANCOVA) analyses were conducted for the winter and spring LNF scores, using the fall LNF scores as the covariate to take into account the differences in students' letter naming fluency prior to the intervention. The scores for the intervention group were significantly higher at both the winter and spring assessments than the scores for the comparison group.

These results indicate that students in both groups began kindergarten with comparable skills in Initial Sound Fluency but the intervention group students demonstrated significantly higher skills by the winter assessment. In addition, student performance for Phoneme Segmentation Fluency was significantly higher among the intervention group. Performance for Letter Naming Fluency was also significantly higher at the middle and end of year for students in the intervention group when accounting for differences in skills at the beginning of the year. However, the groups did not demonstrate significant differences in their Nonsense Word Fluency scores by the end of the year.

First Grade. The DIBELS scores were examined for first grade students across the year. Table 4 presents the means and standard deviations for each measure. Students began the year with comparable scores for LNF and NWF, with the comparison group having slightly higher scores for PSF. At the winter and spring assessment points, the intervention group demonstrated higher scores for all measures. Figures 5-7 display the scores across the year for PSF, NWF and ORF.

Table 4. First Grade DIBELS Scores By Group

Measure	Intervention			Comparison		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Letter Naming Fluency						
<i>Fall LNF</i>	743	38.64	16.366	552	38.85	16.558
Phoneme Segmentation Fluency						
<i>Fall PSF</i>	743	26.94	14.422	551	28.21	13.951
<i>Winter PSF</i>	769	46.64	14.638	562	42.42	14.499
<i>Spring PSF</i>	765	51.89	11.671	555	44.75	12.349

Nonsense Word Fluency						
<i>Fall NWF</i>	744	25.33	20.541	552	26.30	19.305
<i>Winter NWF</i>	769	53.60	26.356	561	44.06	21.990
<i>Spring NWF</i>	765	71.42	30.305	555	59.78	28.251
Oral Reading Fluency						
<i>Winter ORF</i>	769	32.01	30.246	561	25.31	27.213
<i>Spring ORF</i>	765	54.54	33.934	554	49.98	31.261

Figure 5. First Grade PSF Scores By Group

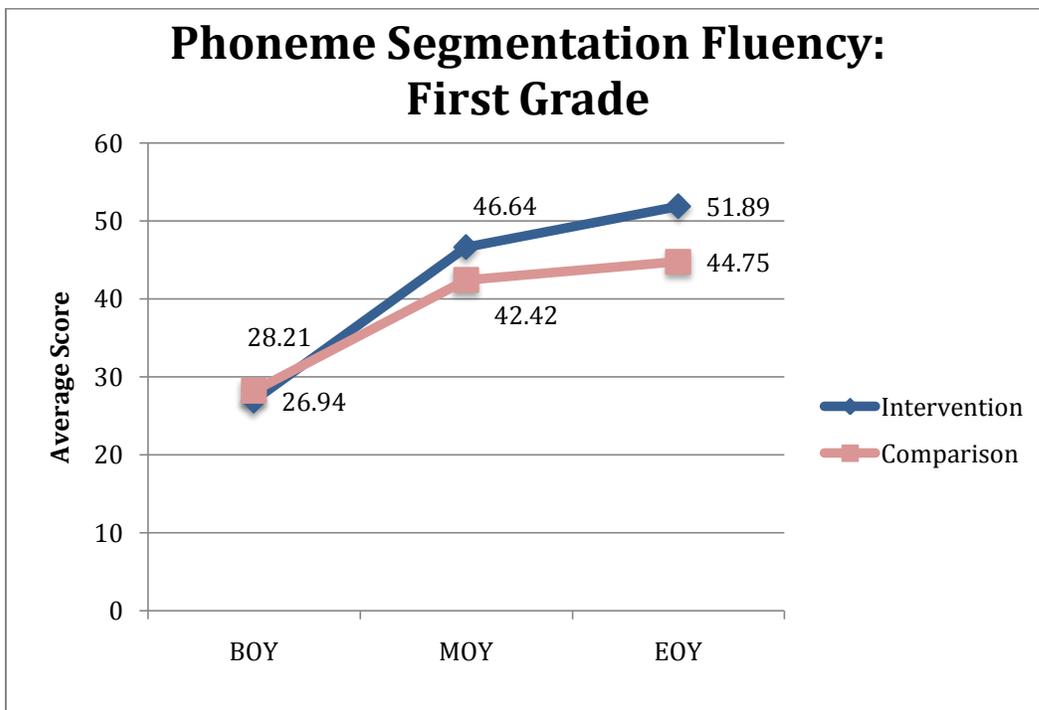


Figure 6. First Grade NWF Scores By Group

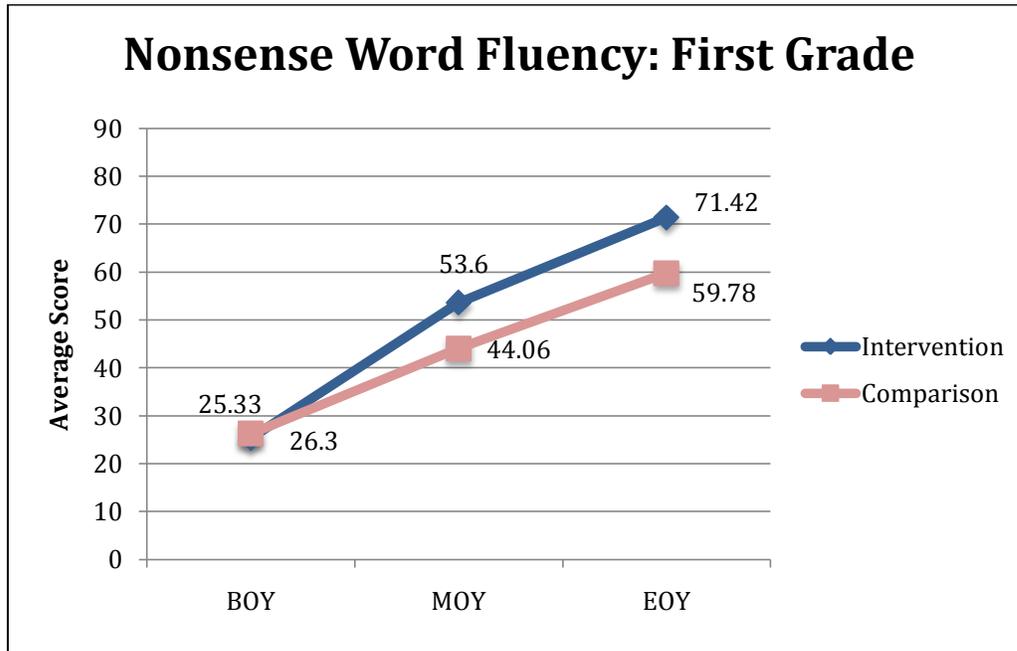
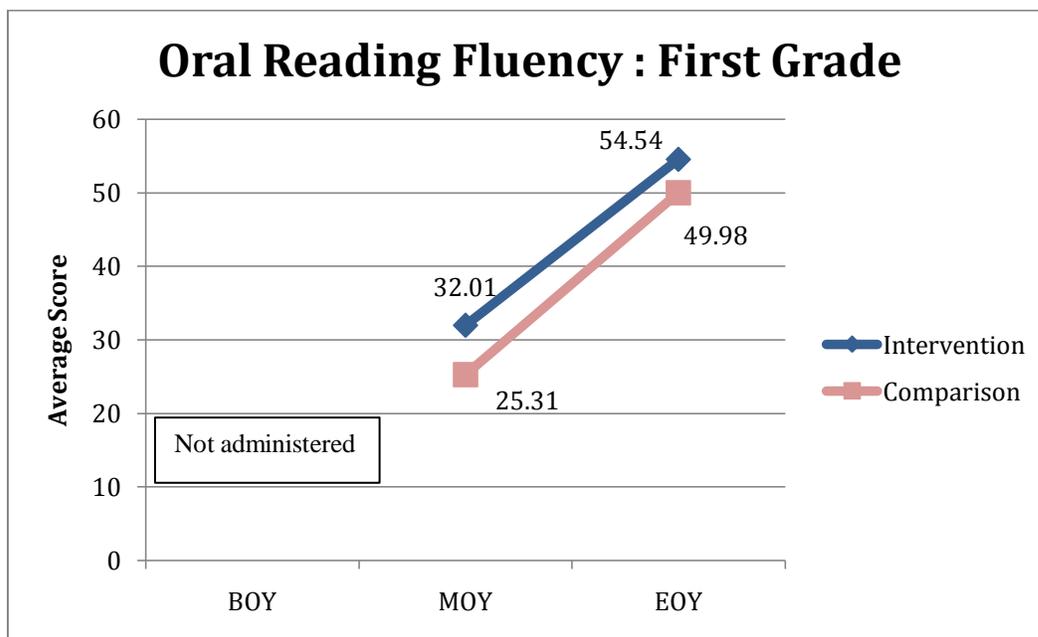


Figure 7. First Grade ORF Scores By Group



Repeated measures ANOVA were conducted for PSF, NWF, and ORF. Significant differences were found for both PSF and NWF, but the scores for ORF did not differ significantly between the groups. To further explore these results, a series of t-tests were conducted for PSF and NWF. The fall scores for PSF and NWF did not differ significantly between the groups. However, the winter and spring scores for both measures were significantly higher for the students in the intervention group.

These results indicate that the two groups began with comparable skills in phonological awareness (as measured by PSF) and letter-sound correspondence (as measured by NWF). Over time, students in the intervention group demonstrated significantly higher skills in both areas. However, the groups did not have significant differences in scores for ORF.

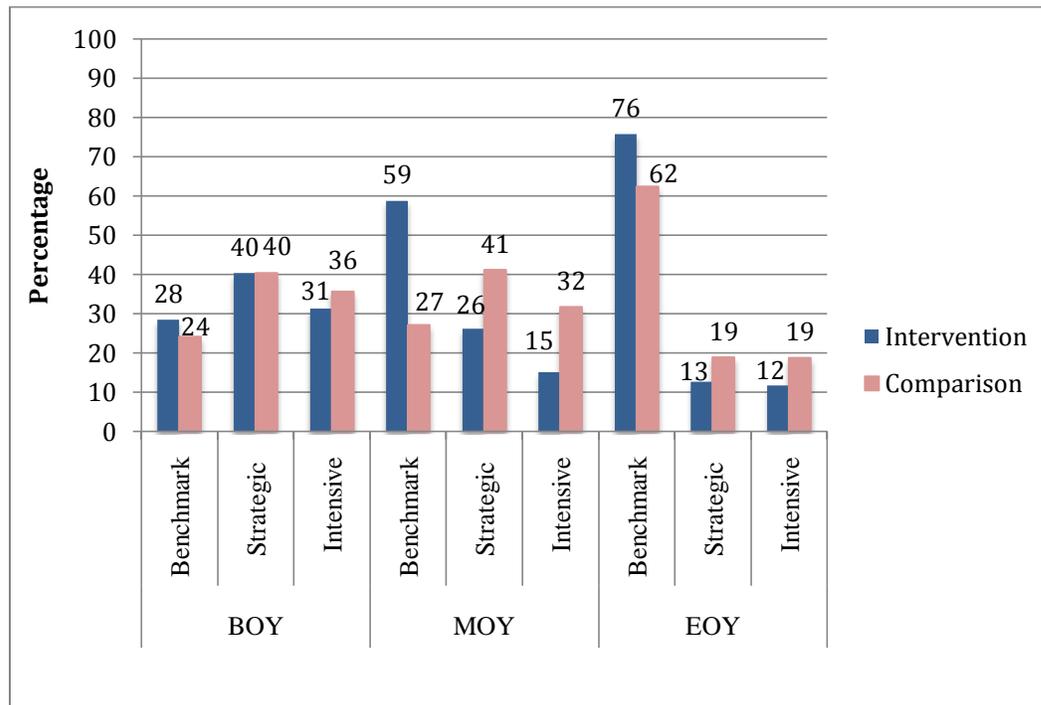
In summary, the intervention group demonstrated an advantage over the comparison group in kindergarten and first grade. Over time, kindergarten students in the intervention group developed significantly higher skills in ISF, LNF and PSF. In first grade, students in the intervention group also displayed significantly higher skills in PSF, as well as with NWF, at the winter and spring assessments. These findings suggest that students who participate in the intervention gain skills in selected literacy areas across the year.

Evaluation Question 3: Did the groups have differences in instructional recommendations in kindergarten and first grade?

Students' overall instructional recommendations based on the risk levels of each DIBELS measure were examined across the year for kindergarten and for first grade. The instructional recommendations include *Benchmark*, *Strategic*, and *Intensive*.

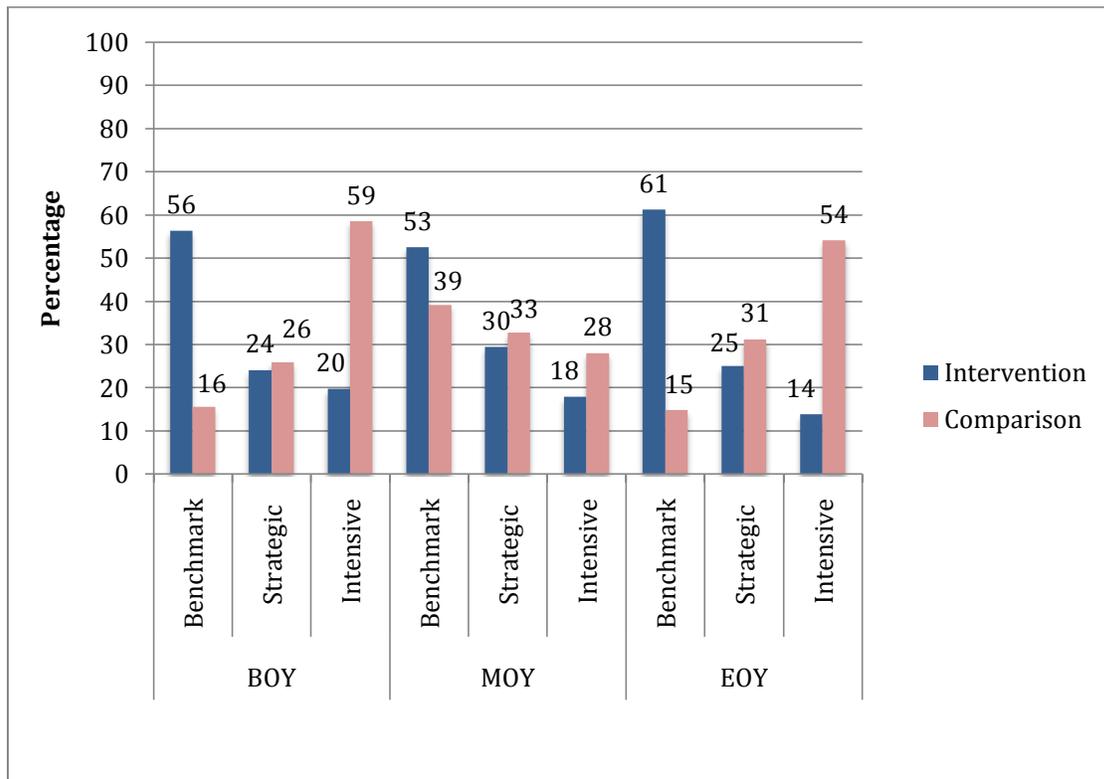
Kindergarten. Figure 8 displays the instructional recommendations for each group across the year. The groups began kindergarten with similar percentages of students at *Benchmark*, *Strategic*, and *Intensive*. For both groups, 40% of their students began kindergarten at the *Strategic* level. Each group began kindergarten with more students at *Strategic* and *Intensive* than at *Benchmark*. At the middle of the year, students in the intervention group displayed a large increase in the percentage of students at *Benchmark*, with a substantially higher percentage of students at *Benchmark* than the comparison group. By the end of the year, both groups had large increases in the percentage of students at *Benchmark* and had decreases in the percentage of students at *Intensive*. However, the intervention group had a higher percentage of students at *Benchmark* and a smaller percentage of students at *Intensive* at the end of the year. Further, the increase in the percentage of students at *Benchmark* was larger among the intervention group.

Figure 8. Kindergarten Instructional Recommendations



First Grade. In first grade, the intervention group began the year with a substantially higher percentage of students at *Benchmark* and a substantially lower percentage of students at *Intensive* than found among the comparison group (see Figure 9). By the end of the year, the intervention group had increased the percentage of students at *Benchmark* and decreased the percentage of students at *Intensive*. Among the comparison group, the percentage of students at *Benchmark* was essentially the same at the end of the year as in the beginning of the year (with a decrease of one percentage point). The percentage of students at *Intensive* decreased by approximately the same amount of points for both the intervention and comparison groups.

Figure 9. First Grade Instructional Recommendations



In summary, these results suggest that students who participated in the intervention group displayed less risk by the end of the year than students in the comparison group for both kindergarten and first grade. In kindergarten, both groups of students began with similar percentages of students at the three levels of instructional recommendations and

Conclusions

The results of this evaluation indicate that students in the intervention group significantly outperformed the students in the comparison group on several literacy skills in kindergarten and first grade. An examination of students' scores across kindergarten found that students in the intervention group demonstrated significantly higher skills in phonological awareness and the identification of letters in the winter and spring, and letter-sound correspondence in the winter. In addition, the intervention group had significantly higher skills in letter naming fluency when accounting for differences in skills at the beginning of the year. Among the first grade students, the intervention group had significantly higher scores for letter-sound correspondence and phonological awareness. In addition, kindergarten and first grade students from the intervention group had lower levels of risk in their literacy skills at the end of the year.

These findings suggest that with one year of intervention, students gained significantly higher skills in selected literacy areas and had lower risk levels at the end of the year. The SBSL comprehensive approach to literacy provided teachers with skills intended to enhance students' literacy. The evaluation of the use of this approach in the HASD indicated that students in kindergarten and first grade developed more skills in selected literacy areas than students from a school district which did not utilize these services. The results from this evaluation support the use of this comprehensive approach for improving literacy skills in kindergarten and first grade.