

2016-17 Evaluation of the Home Instruction for Parents of Preschool Youngsters (HIPPY) Program

Heather Turner, Ph.D. & Sarah Hinkley, Ph.D.

EA17-178-4

At-a-Glance¹

Home Instruction for Parents of Preschool Youngsters (HIPPY) is a home visiting program for parents of three-, four-, and five-year-olds targeting school readiness and parent involvement. HIPPY Dallas Independent School District (ISD) had a total budget of \$2,572,106 for program year 2016-17, including \$1,000,000 from Dallas ISD general operating funds, \$1,158,446 from Title I funding, and \$413,660 from external sources. The HIPPY program in Dallas began in 1988 and is the largest of 11 HIPPY programs in the state of Texas, serving 1,226 children and 1,113 parents in 2016-17, with 1,087 children and 989 parents successfully completing the program year.

Purpose of the Evaluation

This purpose of this evaluation was to evaluate and report outcomes of the HIPPY program. Specifically, this evaluation reports demographic data for participants; pre- and post-test results from a school readiness assessment; kindergarten readiness levels; long-term reading and mathematics outcomes from former HIPPY participants; and results from a parental involvement survey for first year HIPPY parents.

Participants

Methodology

HIPPY Dallas ISD program implementation information primarily came from an interview with the HIPPY Dallas ISD program director and data files from the Texas HIPPY office. The evaluator conducted frequency analyses to summarize participant characteristics.

Results

Of the 1,226 children served by HIPPY Dallas ISD in 2016-17, 1,087 completed the program (989 parents). Enrollment was highest in HIPPY Year 1² (51%, $n = 526$). A majority of the children and parents who completed the program were Hispanic (both 86%). Most children participated with a biological parent (93%, $n = 1,010$). A majority of the parents were married (70%, $n = 698$), did not work outside the home (69%, $n = 680$), and completed the program using the Spanish curriculum (76%, $n = 766$). A majority of the children were female (52%, $n = 568$). Participation in HIPPY

Dallas ISD decreased this year to 1,226 children compared to 1,237 in 2015-16. There was an increase, however, in the number of completers ($n = 1,087$ in 2016-17 compared to $n = 985$ in 2015-16).

School Readiness

Methodology

The evaluator used paired-sample t tests (statistical significance) and Cohen's d effect sizes (practical significance)³ to compare pre- and post-test scores on the English and Spanish versions of the *Bracken School Readiness Assessment (BSRA-3)*.

Results

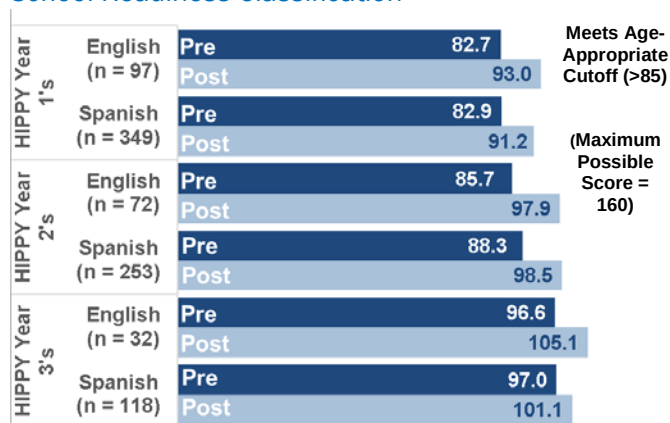
There were statistically significant pre/post improvements on composite test raw scores at all curriculum levels in both the English and Spanish versions ($p < .001$). All pre-test to post-test differences were practically significant with Cohen's d effect size values ranging from .27 (small effect size, HIPPY Year 3 Spanish) to .66 (medium effect size, HIPPY Year 2 English). Both English and Spanish average standard scores for all curriculum years improved from pre-test to post-test (see Figure 1). Average standard scores in both languages improved across all curriculum years and exceeded the school-ready threshold by the end of the program year. A majority of HIPPY Year 1's (71% English; 62% Spanish), Year 2's (72% English; 79% Spanish), and Year 3's (78% English; 86% Spanish) were classified as meeting age-appropriate standards by the end of the year.

¹ Portions of this evaluation were adapted from McEnturff (2014) and Palladino (2015, 2016).

² There are three years of HIPPY curriculum, Year 1 (first year) through Year 3 (last year).

³ Cohen's d values of 0.2, 0.5, and 0.8 are considered small, medium, and large effect sizes, respectively (Cohen, 1988). Practical significance was only examined when statistical significance (where $p < .05$) was observed.

Figure 1: 2016-17 HIPPY Dallas ISD Average BSRA-3 Standard Scores at Pre-Test and Post-Test with School Readiness Classification



Source: Texas HIPPY program data files as of July 11, 2017.

Note: Standard score categorizations were established by the publisher (Bracken, 2002), and were: Very Delayed (≤ 70), Delayed (71 to 85), Average (86 to 114), Advanced (115 to 129), and Very Advanced (≥ 130 to 160). There was no normative sample of Spanish version test-takers. Refer to Palladino (2015, 2016) regarding limitations of the analysis and normative sample.

Kindergarten Readiness

Methodology

The evaluator used the beginning-of-year *Istation's Indicators of Progress (ISIP)* exam data file dated October 18, 2016 to compare differences in the percentage of HIPPY and Non-HIPPY kindergarten students that met Tier 1 levels at the beginning of the year. *ISIP* is a computer-adaptive assessment that provides continuous monitoring of progress in reading skills.⁴ *ISIP* scores are assigned to tiers based on student score percentiles. An assignment of Tier 1 indicates the student's score fell into the 40th percentile or above. A score assigned to Tier 1 is considered to be at or above grade level. The evaluator merged *ISIP* data with demographic data dated October 31, 2016. For the analysis, the English and Spanish administrations were combined into one file to allow for examination of overall student performance.⁵ *ISIP* data were analyzed using a *Chi-square* analysis to compare the proportions of HIPPY and Non-HIPPY kindergarten students who performed at or above grade level.

Results

Out of 10,180 kindergarten students with beginning-of-year *ISIP* data in Dallas ISD in 2016-17, 270 were flagged as former HIPPY participants. Of

these former HIPPY participants, 51 percent ($n = 138$) were reading at or above grade level compared to 55 percent ($n = 5,443$) of Non-HIPPY kindergarten students. While the Non-HIPPY students slightly outperformed former HIPPY students the proportional difference between groups was not statistically significant, $\chi^2 (2, n = 10,180) = 1.64, p = .44$.

Long-Term Reading and Mathematics Outcomes

Methodology

The evaluator used *Terra Nova/SUPERA* (kindergarten through grade two) and *State of Texas Assessments of Academic Readiness (STAAR; grades three through five)* to compare passing rates and mean standardized scores between former HIPPY students and matched Non-HIPPY control groups.⁶ The matched control groups were created using propensity score matching to identify students who took the same exam but did not participate in the HIPPY program.⁷ The evaluator used *Chi-square* analyses to test for group differences in passing rates, and *t* tests (statistical significance) and Cohen's *d* effect sizes (practical significance) to compare average scale scores. District passing rates were provided for reference only.

Results

Terra Nova/SUPERA

The evaluator examined *Terra Nova* reading and mathematics and *SUPERA* reading passing rates and average scale scores for kindergarten through grade two. HIPPY students outperformed Non-HIPPY students on eight of nine *Terra Nova/SUPERA* reading and mathematics exams (see Figure 2). The grade one *Terra Nova* reading passing rate difference between HIPPY and Non-HIPPY students was statistically significant, $\chi^2 (1, n = 100) = 5.19, p = .02$. Although the HIPPY students were not statistically compared to the district, the HIPPY students also outperformed the district on all exams except for *Terra Nova* reading for kindergarten students.

For scale scores, grade one and grade two *Terra Nova* reading exam average scale score differences were statistically significant (grade one: $p = .04$, grade two: $p = .045$) and average scale scores were higher for HIPPY students compared to Non-HIPPY students. Scale score differences in kindergarten *SUPERA*

⁴ *ISIP* Tier 1 level outcomes have been analyzed in several early childhood program evaluations (Kim, 2017; Palladino, 2017).

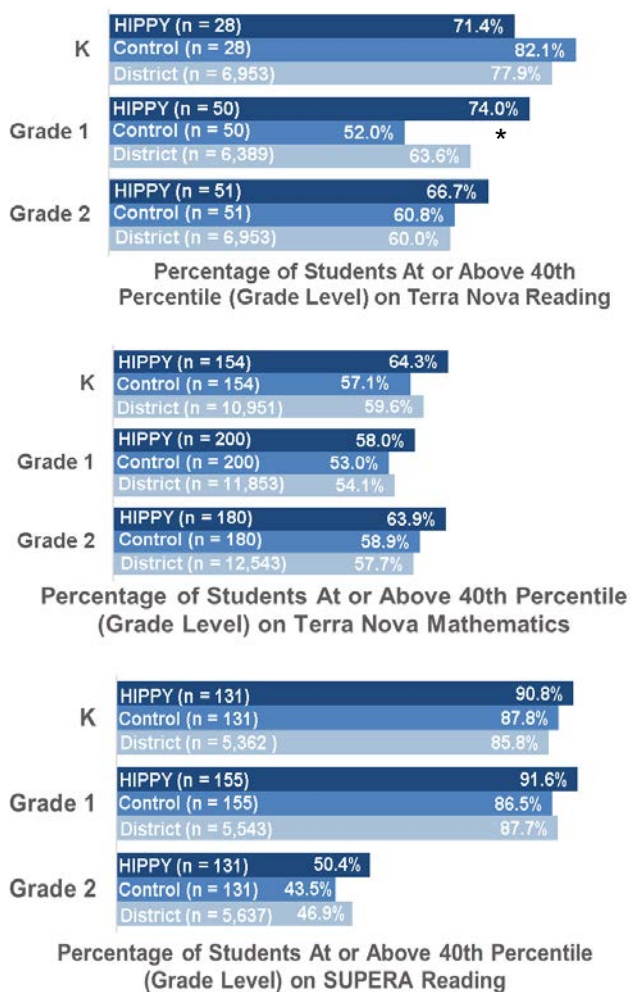
⁵ Comparisons between the English and Spanish versions of *ISIP* are not recommended because the two versions are normed differently. Therefore, the percentage of students at or above Tier 1 (which indicates performing at grade level) was used.

⁶ *SUPERA* file dated 8/10/17; grade one and grade two data were missing for two schools.

⁷ Non-HIPPY matched controls included students who took these exams at each grade who did not participate in HIPPY but had the same probability of HIPPY enrollment. This methodology allows for group comparisons with reduced concern about systematic differences that may confound the effects of the treatment. For a detailed methodology on propensity score matching, readers are referred McEnturff (2014) and Palladino (2015, 2016).

reading were statistically and practically significant ($p = .02$, Cohen's $d = .28$).

Figure 2: 2016-17 Terra Nova/SUPERA Reading and Mathematics Passing Rates for Former HIPPY Participants, Matched Controls, and the District



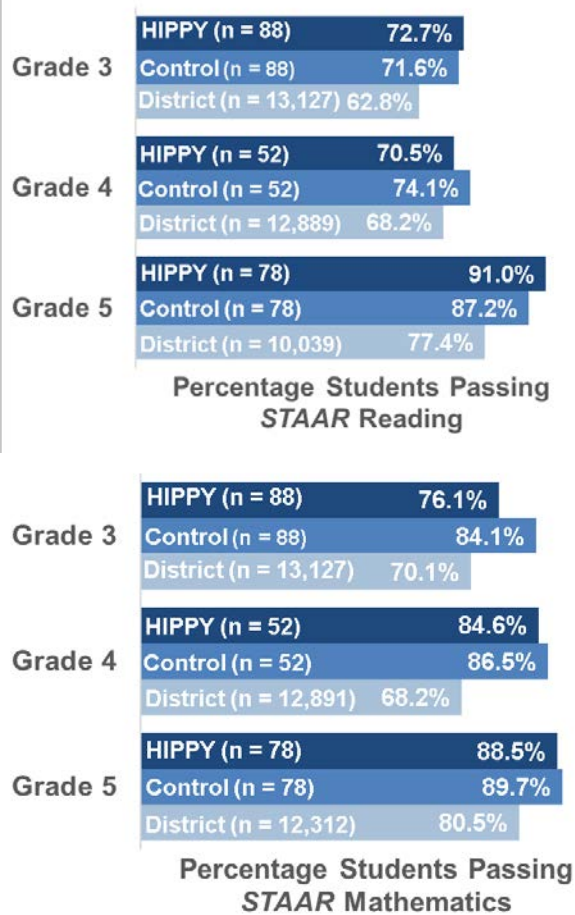
Source: Terra Nova data reported as of July 19, 2017 for students enrolled on March 27, 2017. SUPERA file dated August 10, 2017; grade one and grade two data were missing for two schools.
Note: Students scoring at or above the 40th percentile were considered passing. * indicates statistically significant differences were found between HIPPY and Non-HIPPY groups on the Terra Nova exam passing rates.

STAAR

The evaluator compared STAAR reading and mathematics passing rates for grades three through five. Despite HIPPY students having a lower percentage of passing rates on four of six STAAR exams compared to Non-HIPPY students (see Figure 3), none of these differences were statistically significant. No comparisons between HIPPY and Non-HIPPY groups on STAAR average scale scores were statistically significant. Although it was not appropriate to statistically compare HIPPY passing

rates to those of the district because of differing sample sizes and possible confounding factors, HIPPY passing rates exceeded the district for all STAAR reading and mathematics exams in grades three through five.

Figure 3: 2016-17 STAAR Reading and Mathematics Passing Rates for Former HIPPY Participants, Matched Controls, and the District



Source: District STAAR files dated June 22, 2017.
Note: Passing rates reflect approaching grade level or higher scores using the year-specific standard for all versions of STAAR excluding Alternate 2. Scores included for students enrolled as of March 27, 2016. No comparisons between HIPPY and control groups on the percentage of students passing or on average scale scores were statistically significant. For more information on the methodology on propensity score matching, readers are referred to Thoemmes, F. (2012), Bush and Kim (2016), McEnturff (2014), and Palladino (2015, 2016).

Parent Involvement

Methodology

First-year HIPPY parents completed the Parent Involvement Interview at the beginning and the end of the program year. Frequency analyses were conducted to compare pre- and post-test outcomes on completed cases.

Results

New HIPPY Dallas ISD participants achieved marked gains in parental involvement, as reported on the Parental Involvement Interview. First, parents reported spending an average of over 3 extra minutes per day reading to their children at post-test (16.9 minutes) compared to pre-test (13.1 minutes). In addition, at least 90 percent of parents reported at post-test that they engaged in four literacy enrichment activities when they read to their children, with two of the activities approaching 100 percent participation (stopped reading and asked the child what is in the picture and talked about the story when the book is done).

Recommendations

- **Continue to seek areas for improvements in effectiveness.** The program has enjoyed pre/post gains in developmental skills and parent involvement every year it has been evaluated. HIPPY leadership should continue to nurture its staff and review processes to maintain these strong results.
- **Include HIPPY in the evaluation of other major Dallas ISD early childhood programs.** Many children in the Dallas ISD attend more than one pre-kindergarten program. Thus, it has become more difficult to evaluate the efficacy of any one early childhood program exclusively, particularly for long-term outcomes. Evaluation and program managers should consider program participation and dosage as a variable in future evaluations of early childhood education. Data structures and systems must be carefully reviewed and enhanced to support this goal.

References

- Bracken, B. A. (2002). *Bracken school readiness assessment*. San Antonio, TX: The Psychological Corporation.
- Bush, J. & Kim, M. (2016). Evaluation of Reasoning Mind: 2015-16. Dallas Independent School District.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd. ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Kim, M. (2017). Updated Evaluation of Kindergarten to Grade 2 Final Report 2015-16. Dallas Independent School District.
- McEnturff. (2014). Evaluation of the 2013-14 Home Instruction for Parents of Preschool Youngsters (HIPPY) program. Dallas Independent School District.

Palladino, D. (2015). Evaluation of the 2014-15 Home Instruction for Parents of Preschool Youngsters (HIPPY) program. Dallas Independent School District.

Palladino, D. (2016). Evaluation of the 2015-16 Home Instruction for Parents of Preschool Youngsters (HIPPY) program. Dallas Independent School District.

Palladino, D. (2017). Updated Evaluation of Accelerating Campus Excellence (ACE) 2015-16. Dallas Independent School District.

Thoemmes, F. (2012). Propensity score matching in SPSS. Retrieved from <http://arxiv.org/abs/1201.6385>. Accessed July 25, 2017.

An electronic copy of this report can be found at www.dallasisd.org/Page/888. For more information, please contact Program Evaluation at evaluation@dallasisd.org.

© 2017 Dallas Independent School District (Dallas ISD)

Mail a written request for permission to use or reproduce any part of the document to:
Dallas ISD
Department of Evaluation and Assessment
H. B. Bell School Support Service Center
2909 N. Buckner Blvd.
Box 10
Dallas, TX 75228