Evaluation of the National Family Literacy Program: 2016 Report

Prepared by:

Craig A. Mason, Ph.D.

July 2016

College of Education and Human Development
University of Maine
Orono, Maine
Evaluation of the National Family Literacy Program: 2016 Report

July 2016

Craig A. Mason, Ph.D.
Professor of Education and Applied Quantitative Methods

University of Maine,
5766 Shibles Hall
Orono, Maine 04469-5766
(207) 581-2493

A Member of the University of Maine System
AUTHOR’S BIOGRAPHICAL INFORMATION

Craig A. Mason, Ph.D., is a Professor of Education and Applied Quantitative Methods at the University of Maine with over 25 years of experience providing program evaluation, statistical analysis, and informatics support to national agencies and research teams. Dr. Mason received his PhD in Clinical Child Psychology from the University of Washington, and his research interests are in developmental growth models, informatics, and research methodology. For nearly fifteen years, Dr. Mason has served as a methodological consultant to the U.S. Centers for Disease Control and Prevention, and he has been invited to make presentations on methodology by the Centers for Disease Control and Prevention, the National Association for Public Health Statistics and Information Systems, the U.S. Department of Health and Human Services, the Association of Maternal and Child Health Programs, the Centers for Medicare and Medicaid Services, the Association of Maternal and Child Health Programs, and the Association of University Centers on Disabilities. He has over 80 publications, and has been principal investigator or co-principal investigator on over $10 million in grants from sources including the National Science Foundation, the National Institutes of Health, the Centers for Disease Control and Prevention, and the Department of Education.
EXECUTIVE SUMMARY

The Barbara Bush Foundation believes that families are the key to helping children acquire the fundamental literacy skills they will need to succeed in school. However, with many adults struggling with reading themselves, families need support. The Family Literacy Program therefore seeks to help children and families through a broad-based approach to literacy education. Adult family members can take part in classes that target reading, math, and parenting skills, including English-language skills for non-native speakers and GED preparation. Activities support child literacy education, and promote reading as a family experience, helping caregivers make reading an engaging part of daily life that benefits everyone in the home.

This report summarizes findings from the 2015-2016 evaluation of 20 Barbara Bush Foundation Family Literacy programs across the country. Data included in the evaluation include pre- and posttest surveys of adult participants, and standardized assessments of adults and children.

OVERALL PROGRAM IMPACTS

Programs provided a total of 111,538 service hours to families, including 53,557 hours with 360 caregivers seeking adult education or ESOL skills, and 57,981 hours with 381 children in literacy-focused activities. In exit surveys, adult participants uniformly valued their time in the program, with more than 90% reporting that the Family Literacy Program helped them to become a more confident parent who was better equipped to improve their family’s quality of life, including being better able to evaluate their child’s progress, while also being more skilled at communicating with professionals and advocating for their family.

ADULT EDUCATION IMPACTS

Of the 360 total adults served, pre- and posttest assessments were available for 249. These adult participants were primarily female (93.5%), with over half Hispanic (57.7%) with Spanish the language in the home (53.1%). One-third (33.7%) had an 8th grade
education or less and 72.4% of families had annual income of $20,000 or less. Pre- and posttest assessments addressed both basic/general adult education skills, as well as ESOL skills. Adult impacts were assessed using one of four different instruments: The Comprehensive Adult Student Assessment System – Adult Basic Education test, the Comprehensive Adult Student Assessment System – Adult ESOL Literacy test, the Test of Adult Basic Education, and the BEST Plus. The specific test used varied based on the background, needs, and services provided to participants. Given the use of different tools for assessing adult education and literacy/language skills, results for each of the four instruments were also converted to grade-level equivalent scores providing a common scale for examining general program impacts across all adult participants.

For those in the Adult Basic Education group, average grade-level equivalent scores increased from 6.68 to 7.80, or a 1.11 grade-level increase (with rounding). For those in the Adult ESOL Education group, grade-level equivalent scores increased from 3.21 to 4.45, or a 1.24 grade-level increase. For both groups, a larger increase was seen for those participants who began the year with scores in the bottom two-thirds of the range for their program area (i.e., Adult Basic Education or Adult ESOL Education). In contrast, those whose initial grade-level equivalent scores placed them in the top third of adults in their program area showed relatively smaller increases over time. This suggests that the program may particularly benefit those more in need of help.

When individual scales were examined, results reflected moderate to large effects for Adult Basic Education, ranging from \(d=.48\) for those assessed on the TABE to \(d=.77\) for those assessed using the CASAS Adult Basic Education test. Similar patterns were seen for the individual scales assessing Adult ESOL Education. Based on Cohen’s classification of effect sizes, program effects ranged from moderate/large (\(d=.69\) for those assessed using the CASAS Adult ESOL Literacy test) to large (\(d=.86\) for those assessed using the BEST Plus). This reflected skills for a typical ESOL participant growing from the “Beginning ESL Literacy” or “Low Intermediate ESL” range to the “High Intermediate ESL” range based on National Reporting System benchmarks.
**CHILD IMPACTS**

Pre- and posttest assessments were available for 273 of the 381 total children served. Based upon their age, children either completed (a) the Peabody Picture Vocabulary Test and the Expressive Vocabulary Test, or (b) the Infant / Toddler Checklist.

The *Peabody Picture Vocabulary Test*, or PPVT, is a nationally normed assessment of receptive language vocabulary for children as young as 2 ½ years of age. Children demonstrated significant improvement in PPVT standard scores following participation in the Family Literacy Program. For the typical child, this corresponded to moving from the 18th to the 31st percentile nationally, resulting in an effect size of $d = .43$ or close to a moderate-sized effect. Focusing on lower-performing children, this translated to a meaningful reduction in the number of children scoring in the bottom quartile on the PPVT. At the start of the program, 53.7% of these children scored in the lowest 25 percent nationally. However, by the time of their posttest, only 36.5% were in the lowest quartile: **a 32% reduction**.

The Expressive Vocabulary Test (EVT) is a similar nationally normed assessment of *expressive* language vocabulary for children as young as 2 ½ years of age. Children showed a similar statistically significant increase in EVT standard scores, corresponding to a moderate sized effect ($d = .54$). On average, this translated to the typical child moving from the 14th percentile to the 30th percentile in expressive language skills. Alternatively, this reflected a **23% reduction** in the number of children scoring in the bottom quartile on the EVT: from 57.1% at pretest to only 43.8% at posttest.

The Infant / Toddler Checklist (IT) is used to screen young children between 6 and 24 months of age for possible communication impairment. It includes cutoff scores, with performance below the cutoff indicating possible concern that may warrant further attention or continued monitoring. Children demonstrated significant improvement in IT-Total scores following participation in the Family Literacy Program (N=70). As a screening tool, this effect was most clearly seen in the number of children falling below the cutoff scores. At pretest, 25.8% of children scored below the cutoff, indicating a level of concern that may warrant further monitoring or evaluation. By posttest, this had been
reduced to 15.2% of children, or a **41% reduction in the number of children flagged for possible concern.**

The IT *Communication Composite Score* is a subscale of the IT that addresses communications skills, such as emotion and eye contact/gazing, communication, and gestures. Children also showed significant increases specifically in Communication scores with 20.6% of children below the cutoff at pretest, but only 4.4% of children below the cutoff at posttest – a **79% reduction in the number of children flagged for possible concern.** The IT *Expressive Speech Composite Score* is an IT subscale that addresses the use of sound and words. Children showed similar increases in Expressive Speech scores, reflecting a **46% reduction in the number of children flagged for possible concern.** Finally, the *Symbolic Composite Score* addresses the use of objects and the understanding of words. Students showed increases on this scale as well, corresponding to a 13% reduction in the number flagged for possible concern.

**Deaf Family Literacy Programs**

Finally, two Family Literacy programs specifically targeted families of children with hearing loss. Given their unique population and services, these two programs were examined separately from the other 18 programs. As expected, analyses found that adult participants experienced large increases in both their expressive and receptive American Sign Language (ASL) skills following participation in the program. Children in these Deaf Family Literacy Programs also exhibited increases in various literacy-relevant skills, such as cognition, alphabet knowledge, and print awareness.

**CONCLUSION**

In conclusion, this report found that the Barbara Bush Foundation Family Literacy Programs had significant, meaningful impacts on families. For adult and caregivers, this was seen in improvements in basic education skills valuable for daily life and work settings, for children, this reflected increased receptive and expressive language skills and decreased communication and expressive language concerns.
# TABLE OF CONTENTS

Author's Biographical Information ................................................................................... iii

Executive Summary ........................................................................................................iv

Introduction ..................................................................................................................... 1

Family Literacy Programs................................................................................................. 2

Adult Participant Impressions of Program Impacts.......................................................... 4

Analysis Plan for Assessment Data................................................................................. 5

Assessments of Adult Participants .................................................................................. 6

Adult Participants: Change in Grade-Level Equivalent Scores ....................................... 8

  Adult Basic Education ................................................................................................... 8

  Adult ESOL Education ................................................................................................. 10

Results for Individual Instruments................................................................................ 11

  CASAS: Adult Basic Education Test .......................................................................... 11

  CASAS: Adult ESOL Literacy ..................................................................................... 13

  Test of Adult Basic Education (TABE) ................................................................... 14

  BEST Plus ................................................................................................................... 15

Results for Child Participants ......................................................................................... 17

Peabody Picture Vocabulary Test Results....................................................................... 18

Expressive Vocabulary Test .......................................................................................... 21

Infant / Toddler Checklist ............................................................................................. 23

  IT Total Score Results ............................................................................................... 24

  IT Communication Composite Score ...................................................................... 25

  Expressive Speech Composite Score .................................................................... 26

  Symbolic Composite ............................................................................................... 27
Deaf Family Literacy Programs .......................................................................................... 28
Adult Outcomes .............................................................................................................. 28
Child Outcomes ............................................................................................................ 28
Conclusion and Recommendations ................................................................................. 31
References ...................................................................................................................... 32
INTRODUCTION

Literacy is a core component to healthy child development and school-readiness. Children who are poorly prepared to read in kindergarten or early elementary grades are more likely to continue to struggle throughout school. The Barbara Bush Foundation believes that families are the key to helping children acquire literacy skills. However, with many adults struggling with reading themselves, families need support. The Barbara Bush Foundation Family Literacy Program therefore seeks to help children and families grow and succeed through a broad-based approach to literacy education. Adult family members can take part in classes that target reading, math, and parenting skills, including English-language skills for non-native speakers and GED preparation. Activities support child literacy education, and promote reading as a family experience, helping caregivers make reading an engaging part of daily life that benefits everyone in the home.

To help inform these efforts, this report summarizes findings from the 2015-2016 evaluation of the Barbara Bush Foundation Family Literacy Program. Data in the report include pretest and posttest surveys of adult participants, and standardized assessments of adult and child participants. The report first presents a general description of programs, followed by participant impressions of program impacts reported in entry and exit surveys. The report then focuses on the key outcomes of interest: pretest and posttest standardized assessment data from adults and children. These assessments targeted both adult education skills as well as literacy and language skills in children. The report concludes with a review of findings for two programs that specifically targeted families of children with hearing loss. Reflecting the unique families being served by these programs and the need for different assessment instruments, these two programs were examined separately from the others.
FAMILY LITERACY PROGRAMS

This report examines 2015-2016 data from 20 Family Literacy programs operated across 8 states. As summarized in Table 1, five of these programs were based in Florida, four each in Maine and Texas, three in Michigan, and the remaining programs were located in Maryland, Ohio, Oklahoma, and Tennessee. Most were community-based, with three operated through a school or school district (SAIL Into Literacy, SPICE Family Literacy, and Corpus Christi Family Literacy), and one based in a university (Madonna University Deaf Family Literacy). Two programs, one in Michigan and one in Texas, specifically targeted families with a child who had hearing loss.

Programs provided a total of 111,538 service hours to families, closely divided between literacy services to children (57,981 hours across 381 children) and adults in the family (53,557 hours across 360 adults). This translates to an average of 152.2 service hours per child and 148.8 service hours per adult. As illustrated in Figure 1, the number of both child and adult service hours varied by program (F(18,362)=35.114, p<.001; and F(18,359)=23.026,

<table>
<thead>
<tr>
<th>State</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>Family Literacy Academy of Lake Wales</td>
</tr>
<tr>
<td>Florida</td>
<td>Family Literacy Academy of Tampa Bay</td>
</tr>
<tr>
<td>Florida</td>
<td>Family Literacy Yes! (FLY!)</td>
</tr>
<tr>
<td>Florida</td>
<td>Food Brings Hope</td>
</tr>
<tr>
<td>Florida</td>
<td>Grace Place Naples</td>
</tr>
<tr>
<td>Maine</td>
<td>Passages-Wayfinder School</td>
</tr>
<tr>
<td>Maine</td>
<td>Riverton Family Literacy Initiative-Learning Works</td>
</tr>
<tr>
<td>Maine</td>
<td>SAIL Into Literacy</td>
</tr>
<tr>
<td>Maine</td>
<td>SPICE Family Literacy</td>
</tr>
<tr>
<td>Maryland</td>
<td>Family Discovery Center</td>
</tr>
<tr>
<td>Michigan</td>
<td>Madonna University Deaf Family Literacy</td>
</tr>
<tr>
<td>Michigan</td>
<td>Wayne Metro-Hamtramck</td>
</tr>
<tr>
<td>Michigan</td>
<td>Wayne Metro-Highland Park</td>
</tr>
<tr>
<td>Ohio</td>
<td>Zanesville Family Literacy</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>Ardmore Barbara Bush Literacy Corps</td>
</tr>
<tr>
<td>Tennessee</td>
<td>Deaf Family Literacy Midsouth</td>
</tr>
<tr>
<td>Texas</td>
<td>Aberg Center for Literacy</td>
</tr>
<tr>
<td>Texas</td>
<td>Corpus Christi Family Literacy</td>
</tr>
<tr>
<td>Texas</td>
<td>Northside Inter-Community Agency</td>
</tr>
<tr>
<td>Texas</td>
<td>Rio Grand Valley Literacy Center</td>
</tr>
</tbody>
</table>
Family Literacy Program 2016

$p<.001$; respectively). Note that programs names have been removed and replaced by letters that do not correspond to the order that appears in Table 1, nor to an alphabetical ordering of names.

Programs were loosely organized based on fall and spring terms, with assessments of both adult family members and children conducted at the start and end of each term (4 assessments per year). However, many full-year participants only had a start-of-year and end-of-year assessment. Participation generally reflected three different types of enrollment: Those who took part only in a fall term, those participating only in a spring term, and those participating for the full year. Reflecting the nature of the assessments that adults completed, participants could be viewed as seeking either broadly-defined basic adult education skills or ESOL skills.
ADULT PARTICIPANT IMPRESSIONS OF PROGRAM IMPACTS

At the time of their entry into the program, adults identified various personal reasons for enrolling in the Family Literacy Program. Chief among these was help earning a GED (56.1% of participants), a desire to improve English communication skills (52.7%), learning how to better help their children with homework (43.0%), and helping their children develop skills needed for success in school (35.0%).

By the time of their exit survey, participants indicated that the program had helped them in all of these areas. For example, 61.3% specifically noted that the program had helped them improve their English speaking and writing skills, while 53.5% reported that the program had helped them learn ways to better assist their child with homework, and 41.5% reported learning ways to help their child develop skills needed to succeed in school. Among basic adult education participants who entered the program interested in obtaining a GED, 28.3% had in fact obtained their GED by the time of their exit survey.

Participants also reported receiving help and benefits in areas beyond those they were initially seeking. For example, when entering the program only 17.4% of adults indicated they were looking for help with parenting skills; however by the end of the year, 50.1% volunteered that the program had helped improve their parenting. In fact, by the end of the program nearly all of the adult participants (98.6%) reported that they felt more confident as a parent, better able to evaluate how much progress their child is making (96.0%) and more aware of what is expected in their child’s school (89.9%).

More broadly, participants reported that they felt better able to communicate with people working with them and their child (94.9%), and advocating for themselves and their family (94.6%). All of this made them better equipped to improve their family’s quality of life (94.6%).

In terms of child-literacy related skills and resources, at the end of the program, participants reported more child-focused books in the home (t(271)=4.106, p<.001). At the start of the program, 41.3% of adults reported having 10 or fewer child books at
home; but by the end of the program that had been reduced to only 24.5% of families. They also reported an increase in literacy-related resources in their home, such as ABC books, magnetic letters, and reading games ($t(280)=3.735, p<.001$), and more frequent use of engaging reading behaviors, such as asking their child what they think will happen next when reading, or asking them questions about the material they are reading ($t(226)=2.011, p=.046$).

Participants found the experience sufficiently rewarding that 75.9% indicated they planned to return next year.

**Analysis Plan for Assessment Data**

For analyses examining change in adult and child assessment data, participants were classified as either “full-year” or “term-only” through the following procedure: If someone completed the first assessment (i.e., at the start of the fall term) and fourth assessment (i.e., at the end of the spring term) they were considered “full-year”. For the remaining participants, if they completed both the first and second assessment in the fall or both the third and fourth assessment in the spring, they were considered “term-only” participants. Given the small number of fall-term-only and spring-term-only participants, data for both of these groups were combined for analysis.

For succinctness, the report will refer to the first assessment based on these groups as a participant’s “pretest”. This would reflect the 1st assessment in the fall for full-year participants and term-only participants enrolled in the fall, and the 1st assessment in the spring for term-only participants enrolled in the spring. The report will refer to the second assessment based on these groups as a participant’s “posttest”. This would be the 2nd assessment in the spring for full-year participants and term-only participants enrolled in the spring, and the 2nd assessment in the fall for term-only participants enrolled in the fall.

Given that the two programs targeting families of children with hearing loss (understandably) used fundamentally different assessment instruments, these two programs were examined separately from the others and are discussed at the end of
this report. Therefore, the assessment data discussed in the following sections on adult
and assessments are based on the 18 other family literacy programs.

**ASSESSMENTS OF ADULT PARTICIPANTS**

As noted previously, programs served a total of 360 adults. The majority of these adults
were female (93.6%) and Hispanic (58.6%), with Spanish as the primary language in
the home (51.0%). For most of these families, home-life reflected an adult couple
(62.7%) or extended family (10.5%) with income of less than $20,000 (70.9%). Nearly a
third (32.1%) of adult participants had the equivalent of an 8th grade education or less.
Thirty-one percent had been involved with the Family Literacy Program in the previous
year.

Of the 360 total adults served, 21 were participants in the two Deaf Family Literacy
programs and are discussed separately at the end of this report. Records indicated that
of the remaining 339 adults, 199 had successfully completed the program. Of these,
184 had also completed both a pretest and posttest assessment. An additional 47
participants were missing completion-status information, but had also finished both a
pre- and posttest. Finally, there were 18 adults whose completion status indicated that
they had withdrawn or dropped out of the program, and yet they nevertheless had
finished both a pretest and posttest. For example, this would include participants who
completed the fall term, but then withdrew midway through the spring term.

Therefore, the following outcome analyses focus on the 249 adults who (a) were
documented to have successfully completed the program and had a pre and posttest
assessment, (b) had no documented completion status, but had completed a pretest
and posttest, or (c) had completion status indicating they dropped out, but nevertheless
had finished both a pretest and posttest. These adult participants were very similar to
the total served: 93.5% were female, 57.7% were Hispanic with Spanish the language in
the home (53.1%). Black/African American adults comprised 13.8% of the sample. The
most common home settings reflected a couple with children (65.6%) or an extended
family (10.4%). One-third (33.7%) had an 8th grade education or less and 72.4% of families had annual income of $20,000 or less (see Figure 2).

Four out of five of these adults (80.1%) participated for the entire year, with the balance participating for one term – either fall or spring. The 249 participants averaged 182.0 hours of literacy sessions over the year.

![Figure 2. Participant family income](image)

Adult impacts were assessed using one of four different instruments: The Comprehensive Adult Student Assessment System – Adult Basic Education test, the Comprehensive Adult Student Assessment System – Adult ESOL Literacy test, the Test of Adult Basic Education, and the BEST Plus. The specific test used varied based on the background, needs, and services provided to participants. The following sections first summarize results pooled across all four instruments, and then examine findings based on each individual tool.
ADULT PARTICIPANTS: CHANGE IN GRADE-LEVEL EQUIVALENT SCORES

Given the use of different tools for assessing adult education and literacy/language skills, scores for each of the four instruments were converted to grade-level equivalent scores. These estimated grade-level equivalent scores provided a common scale for examining general program impacts across all adult participants. Using these scores, analyses found that mean grade-level equivalent scores for the 249 adults with pretest and posttest data increased from 4.70 to 5.89, or 1.19 grade-levels ($t(248)=11.924$, $p<.001$).

Follow-up analyses examined whether adults who entered the program with lower levels of skill demonstrated more (or less) improvement. To do this, participants were divided into three approximately equal groups: those who scored in the lowest third at pretest (grade-level equivalent scores less than or equal to 2.60), those who scored in the middle third at pretest (grade-level equivalent scores greater than 2.60, but less than or equal to 5.60), and those who scored in the top third at pretest (grade-level equivalent scores greater than 5.6). Those who entered the program with the lowest initial grade-level equivalent scores showed the greatest increase ($t(78)=7.277$, $p<.001$). Their scores improved by approximately 1.55 grade levels (from 1.57 to 3.13, with rounding). Those in the middle third increased by approximately 1.26 grade levels ($t(85)=9.484$, $p<.001$), increasing from 4.06 to 5.33. Those in the top third increased approximately 0.76 grade levels ($t(83)=4.820$, $p<.001$) from 8.30 to 9.06 grade levels.

Adult Basic Education

While the previous analyses found significant overall effects, large differences between those adults in the ESOL and Adult Basic Education groups warrant examining each separately. For those in the Adult Basic Education group, grade-level equivalent scores increased on average from 6.68 to 7.80, or a 1.11 grade-level increase (with rounding, $t(106)=6.963$, $p<.001$).
As with the previous analyses, Adult Basic Education participants were similarly divided into 3 groups of approximately the same size: those who scored in the lowest third at the pretest (grade-level equivalent scores less than or equal to 4.60), those in the middle third at the pretest (grade-level equivalent scores greater than 4.60, but less than or equal to 8.20), and those in the top third at the pretest (grade-level equivalent scores greater than 8.20). As illustrated in Figure 3, those who entered the program with the lowest scores continued to show the largest increase ($t(34)=3.999, p<.001$), with scores improving by 1.34 grade levels (from 3.41 to 4.75). Scores for those in the middle third increased by approximately 1.29 grade levels (with rounding, $t(36)=4.890, p<.001$), improving from 6.72 to 8.02. Scores for those in the top third increased by 0.69 grade levels ($t(34)=3.271, p=.002$), improving from 9.92 to 10.61 grade levels. In other words, when Adult Basic Education participants are analyzed separately, those who entered the program with pretest scores in the upper third of participants showed relatively smaller gains. There was relatively minimal difference in the gains seen among those who entered with pretest scores the lowest or middle third of the range.

![Figure 3. Grade-level equivalent scores for Adult Basic Education.](image-url)
Adult ESOL Education

For those in the Adult ESOL Education group, grade-level equivalent scores increased from 3.21 to 4.45, or a 1.24 grade-level increase ($t(141)=9.817$, $p<.001$). As with the Adult Basic Education analyses, Adult ESOL Education participants were similarly divided into 3 groups of approximately equal size: those who scored in the lowest third at the pretest (grade-level equivalent scores less than or equal to 1.70), those in the middle third at the pretest (grade-level equivalent scores greater than 1.70, but less than or equal to 3.69), and those in the top third at the pretest (grade-level equivalent scores greater than 3.69).

As illustrated in Figure 4, those who entered the program with the lowest scores showed an increase of 1.44 grade levels – from 1.26 to 2.70 ($t(48)=6.037$, $p<.001$). An almost identical increase was seen among Adult ESOL Education participants who entered the program in the middle third of the range. On average, those participants saw their scores increase by 1.47 grade levels ($t(44)=7.694$, $p<.001$), improving from a grade-
level equivalent of 2.68 to 4.15. Scores for those in the top third at baseline increased approximately 0.83 grade levels (with rounding, \(t(47)=3.889, p<.001\)) from 5.70 to 6.52 grade levels. In essence, when Adult ESOL participants are analyzed separately, those who entered the program with pretest scores in the upper third of the range continued to show relatively smaller gains, while larger and essentially equivalent gains were seen among those who entered with pretest scores in the lowest third or middle third of the range.

**RESULTS FOR INDIVIDUAL INSTRUMENTS**

Additional follow-up analyses examined pretest to posttest gains based on individual assessment instruments used by the various Family Literacy participants.

**CASAS: Adult Basic Education Test**

The *Comprehensive Adult Student Assessment System*, or CASAS is a competency-based instrument designed to “measures the basic skills and the English language and literacy skills needed to function effectively at work and in life” (CASAS, 2016a). It is widely used with adult learners and has been found reliable and valid. A total of 143 adult participants had pretest and posttest results on the CASAS, with 44 of these completing the *Adult Basic Education* test and 99 completing the *Adult ESOL Literacy* test. Results for each specific test are presented separately below.

The 44 participants completing the CASAS Adult Basic Education test showed a statistically significant increase in scores following participation in the Family Literacy Program, on average increasing from a score of 233.3 at the pretest (SD=9.47) to a score of 239.6 at the posttest \(t(43)=5.856, p<.001\). In terms of an effect size, this resulted in a \(d=0.67\), midway between a moderate and large effect\(^1\).

\(^1\) A small effect is typically defined as approximately \(d=0.20\), a moderate effect for \(d=0.50\), and a large effect for \(d=0.80\)
Analyses found a marginally significant difference in the size of the effect based on whether adults participated for a full year or a single term ($F(1,42)=3.359, p=.074$), with the ability to detect a difference limited by the small number of participants in the “term-only” group (n=8). Therefore, further interpretation focused on the 36 full-year participants.

Figure 5 shows the pretest and posttest difference in CASAS scores for full-year participants. Scores for these adults increased from 232.6 at the start of the year to 239.9 at the end of the year. The blue horizontal bars in Figure 5 indicate the mean score for the corresponding assessment, while the red vertical bars show one standard deviation above and below the mean – reflecting the variability in the scores across the 44 participants. In terms of grade-levels, this translates to an increase from middle eighth grade to upper ninth grade in scores (CASAS, 2016b).

An effect size was subsequently calculated specifically for the full-year participants. This resulted in an effect size of $d=.77$, or an increase of 0.77 standard deviations following participation in the Family Literacy program. Based on widely used criteria for evaluating effect sizes, this would be seen as a large effect.

Follow-up analyses found that the magnitude of the treatment effects varied significantly from program to program ($F(4,39)=3,735, p=.011$). This reflected very large effects ($d=1.23$ and $d=.97$) in two of the five programs reporting CASAS Adult Basic Education data. Further analysis found that the size of the effect did not vary based on family income ($F(3,39)=0.897, p=.451$), Hispanic ethnicity ($F(1,42)=0.024, p=.877$), identification as being Black/African American ($F(1,42)=0.057, p=.812$), or the number hours in which the adult took part in the program ($F(1,42)=2.456, p=.125$).
CASAS: Adult ESOL Literacy

The 99 participants completing the CASAS Adult ESOL Literacy test showed a statistically significant increase in scores following participation in the Family Literacy Program, on average increasing from 207.7 at the pretest (SD=12.94) to 214.6 at the posttest ($t(98)=8.336$, $p<.001$). In terms of an effect size, this resulted in a $d=.54$, a moderate sized effect.

Not surprisingly, follow-up analyses indicated that the size of the effect adults experienced varied based on whether they participated for a full year or a single term ($F(1,97)=7.992$, $p=.006$), with larger effects seen with a full year’s participation. While it may be expected that additional time in the program would lead to greater change over time, in this case it also reflected most of the single-term participants already performing much higher at the pretest than full-year participants. The exact reason for this is unclear, but may reflect program-to-program differences in recruitment, enrollment patterns, or services.

Therefore, Figure 6 shows the pretest and posttest difference in CASAS scores on the Adult ESOL Literacy test for full-year participants. Scores for these adults increased from 203.8 at the start of the year to 212.3 at the end of the year. Based on National Reporting System benchmarks (National Reporting System, 2016), this reflects improvement from the “Low Intermediate ESL” range into the “High Intermediate ESL” range of performance. This corresponds to an effect size of $d=.69$ – approaching a large effect.
Further analysis found that the size of the effect did not vary based on family income ($F(3,87)=1.923$, $p=.132$), but did vary significantly from program to program ($F(8,90)=5.133$, $p<.001$). For individual programs, effects sizes ranged from small ($d=.09$) to extremely large ($d=1.39$).

**Test of Adult Basic Education (TABE)**

The second assessment used with adult participants was the *Test of Adult Basic Education*, or TABE (CTB, 2016). The TABE assesses reading, language, language mechanics, and vocabulary, in adult learners. A total of 63 adult participants had pretest and posttest data on the TABE. Results for these participants showed a statistically significant increase in scores following participation in the Family Literacy Program – on average increasing from a score of 489.0 (SD=67.5) at the pretest to 516.9 at the posttest ($t(62)=3.532$, $p=.001$). Surprisingly, the degree to which a person’s score increased did not vary based on whether they took part for a single term or the full year ($F(1,61)=1.374$, $p=.246$); however, this likely reflected the small number of participants in the “term-only” group (n=11).

Nevertheless, for consistency in evaluating effect sizes, Figure 7 illustrates the pretest to posttest change observed among the full-year participants. Scores for these adults increased from 481.8 at the start of the year to 513.9 at the end of the year, reflecting a moderate effect size ($d=.48$). In terms of grade-level equivalent scores, this translates to an increase from a grade-level of 4.8 to 5.8.
Unlike CASAS scores, the increase in TABE scores did not vary by program ($F(7,55)=0.858, p=.546$); however, this also likely reflects the small number of participants completing the TABE at most sites. Similarly, this increase appeared consistent and did not vary based on family income ($F(4,55)=0.207, p=.934$), Hispanic ethnicity ($F(1,61)=0.138, p=.711$), identification as Black/African American ($F(1,61)=0.035, p=.852$), or the number of literacy hours in which the adult participant took part ($F(1,61)=0.279, p=.599$).

**BEST Plus**

The final assessment measure used with adult participants was the *BEST Plus* (Center for Applied Linguistics, 2015). The BEST Plus is designed to assess English language skills in English language adult learners. It targets interpersonal communication relevant for daily life and work. A total of 43 adult participants (all Hispanic) had pretest and posttest data on the BEST Plus. These participants showed a statistically significant increase in scores following involvement in the Family Literacy Program, with scores increasing on average from 383.1 (SD=100.0) at the pretest to 469.5 at the posttest ($t(42)=8.744, p<.001$). This translates to an effect size of $d=.86$, which corresponds to a “large” effect – nearly a standard deviation increase in skill and knowledge. Alternatively, based on National Reporting System for Adult Education benchmarks, it reflects a typical participant moving from the “Beginning ESL Literacy” level to being only 3.5 points below the “High Intermediate ESL” level.

This increase did not vary statistically based on whether a person participated for one term or the full year ($F(1,41)=0.052, p=.821$); however, only 10 participants in this group were in the “term-only” group. Nevertheless, one alternative explanation may be that program effects occur in the first few months of participation. A test of program-to-program differences in the degree to which BEST scores increased approached statistical significance ($F(2,40)=3.110, p=.056$), and so pretest to posttest mean scores
for the three programs reporting BEST data are presented in Figure 8\(^2\). When an effect size was calculated for each program, all reflected large to extremely large effect sizes (see Figure 9).

Additional analyses found that this increase did vary based on family income \((F(4,30)=3.902, p=.011)\). Given the modest sample size and limited income range, the exact nature of this variation is not definitive, but may reflect a larger positive impact for the adults in lower-income families.

\(^2\) Note that data for full-year and term-only participants are both included in this figure.
RESULTS FOR CHILD PARTICIPANTS

Child assessments consisted of the (a) Peabody Picture Vocabulary Test (PPVT) and the Expressive Vocabulary Test (EVT), or (b) the Infant / Toddler Checklist (IT). Of the 381 total children provided services, 22 were participants in the two Deaf Family Literacy programs and are discussed separately at the end of this report. Records indicated that of the remaining 359 children, 199 successfully completed their program. Of those, 182 also completed both a pretest and posttest assessment—139 using the PPVT/EVT and 43 using the IT. Seventeen children completed their program but did not have assessment data.

In addition, records for 76 children indicated that they did not complete or “dropped out” for various reasons. However, 33 of these children nevertheless had a pretest and posttest (22 using the PPVT/EVT, 11 using the IT). For example, this might be a result of a child completing the fall term, but not completing the spring term. Finally, while 84
children were missing completion-status information, 58 nevertheless had pretest and posttest data (42 using the PPVT, 16 using the IT).

Therefore, for outcome analyses the following sections focus on the 273 child participants who (a) were documented to have successfully completed the program and had a pretest and posttest, or (b) had no documented completion status, but had completed a pretest and posttest, or (c) had a closure status indicating they dropped out, but nevertheless had a completed pretest and posttest. This reflects 76.0% of child participants.

The review of child outcomes begins with the findings from the Peabody Picture Vocabulary Test, before reviewing findings from the Expressive Vocabulary Test, and then concluding with findings from the Infant / Toddler Checklist.

**Peabody Picture Vocabulary Test Results**

The *Peabody Picture Vocabulary Test* (PPVT) is a nationally normed assessment of receptive language vocabulary for children as young as 2 ½ years of age (Pearson Assessments, 2007). It has been extensively evaluated for both reliability and validity. PPVT analyses focused on PPVT *standard scores* (national mean = 100, SD = 15) for the 203 children with both pretest and posttest PPVT data. The majority of these children were in preschool or not yet in school (63.1%) and Hispanic (67.5%). The sample was evenly divided by gender (male: 50.7%) and 15.3% identified as Black/African American. Nearly all (85.7%) participated for a full year.

Children enrolled in the Family Literacy Program showed a statistically significant increase in their PPVT standard scores following participation in the Family Literacy Program, with scores increasing from 86.2 to 92.7 on average ($t(202)=5.618$, $p<.001$). While less increase was observed among those participating for a single term, the difference was not statistically significant ($F(1,201)=0.633$, $p=.427$), in part reflecting the small number in that group (n=29).
Based on national norms, this translates to the typical child moving from the 18\textsuperscript{th} percentile to the 31\textsuperscript{st} percentile nationally, corresponding to an effect size of $d=.43$ or a moderate-sized effect. Focusing on lower-performing children, this reflected a meaningful reduction in the number of children scoring in the bottom quartile on the PPVT. At the start of the program, 53.7\% of these children scored in the lowest 25 percent nationally. However, by the time of their posttest, only 36.5\% were in the lowest quartile: a 32\% reduction.

The size of this increase did vary significantly by program ($F(14,187)=4.914$, $p<.001$). In particular, data from two of the programs reflected moderate ($d=-.49$) to large ($d=-.98$) decreases in child PPVT standard scores between pretest and posttest. These would reflect a one-half to a nearly full standard deviation decrease in performance. However, one of these programs reported PPVT scores for only three children, and the pretest mean for these three children was more than one standard deviation above the mean. The other program reported on a larger number of children with an average pretest performance below the mean. Regardless, such large declines in standard scores raise questions about these specific assessments that may merit further review. For example, were there differences in the protocol or procedures that may have either artificially enhanced scores at the pretest or lowered them at the posttest? Alternatively, are there possible data entry or data management issues that may be impacting the reported results? Excluding these two programs, the overall mean effect size for the PPVT increased to $d=.54$. This would reflect a child moving from the 16\textsuperscript{th} percentile to the 33\textsuperscript{rd} percentile in receptive language skills. Figure 10 illustrates the pretest to posttest change in PPVT scores by program, excluding the two programs just noted. Figure 11 presents effect sizes separately for each program, with three programs showing large to extremely large increases.
Figure 10. Change in PPVT standard scores by program.

Figure 11. PPVT standard scores effect sizes by program.

Note: Effect sizes for E = -.98, I = -.49, K = -.02, and N = -.04
The degree of improvement in PPVT scores was consistent for both males and females ($F(1,201)=0.678, p=.411$), while a marginally significant difference in the size of the increase was observed based on children’s age ($F(1,201)=3.416, p=.066$). This suggested that the effect may potentially be larger for younger children. There was stronger evidence that the effect may be larger for Hispanic children ($F(1,201)=5.758, p=.017$), with mean standard scores increasing from 83.4 to 91.8, or an effect size of $d=0.56$. Unfortunately, on average standard scores did not increase for African American children in the sample ($F(1,201)=11.344, p<.001$). However, it should be emphasized that as these standard scores are age-adjusted and so reflect children maintaining the same relative level, not a decline. Finally, the size of improvement in PPVT scores also increased with higher levels of literacy hours ($F(1,201)=7.845, p=.006$).

**Expressive Vocabulary Test**

The Expressive Vocabulary Test (EVT) is a similar nationally normed assessment of expressive language vocabulary for children as young as 2½ years of age (Williams, 2014). It has similarly been extensively evaluated for both reliability and validity. EVT analyses focused on EVT standard scores (national mean = 100, SD = 15) for the same group of children as seen with the PPVT.

Not surprisingly, in many ways the EVT findings for these children followed a similar pattern to their PPVT results. Children showed a statistically significant increase in EVT standard scores following participation in the Family Literacy Program, on average increasing from a score of 83.8 at pretest to a score of 92.0 at the posttest ($t(202)=8.024, p<.001$). Based on national norms, this reflected a slightly larger effect for expressive language ($d=.54$). On average, this translates to the typical participating child moving from the 14th percentile to the 30th percentile in expressive language skills. Focusing on lower-performing children, this was associated with a 23% reduction in the number of children scoring in the bottom quartile on the EVT: from 57.1% scoring in the lowest 25 quartile at pretest, to only 43.8% scoring in the lowest quartile at posttest.
As with PPVT scores, this effect did not vary based on whether children were participating for one term or a full year \((F(1,201)=1.018, p=.314)\) – again, likely due to the small number of children participating in the “term-only” group. The effect did vary by program \((F(15,187)=4.248, p<.001)\). Changes in EVT standard scores for individual programs are presented in Figure 12. Given these are age-adjusted standard scores, a horizontal line would indicate children are growing at the expected rate for their age. In contrast, nearly all programs are showing increases in standard scores, reflecting an acceleration in child expressive language skills.

Figure 13 shows the corresponding effect sizes for individual programs. One program shows an extremely large increase \((d>1.50)\), while several are approaching the classical definition of a “large” effect size \((i.e., d=.80)\). One program showed a small-to-moderate decrease in EVT standard scores \((d=-.31)\), which may also warrant closer review.

Improvement in EVT scores did not vary by child gender \((F(1,201)=0.020, p=.887)\); however, analyses suggested that the increase may be greater for younger children \((F(1,201)=5.222, p=.023)\), with an effect size of \(d=.67\) for children in preschool – midway between a moderate and large effect.
Scores for Hispanic children again reflected large increases following participation in the Family Literacy Program \((F(1,201)=10.768, p=.001)\), increasing from 79.2 to 89.6 \((d=.70)\). Increases were also greater with higher levels of literacy sessions \((F(1,201)=7.845, p=.006)\).

**INFANT / TODDLER CHECKLIST**

The Infant / Toddler Checklist is a component of the *Communication and Symbolic Behavior Scales–Developmental Profile* (Wetherby and Prizant, 2001). It is used to screen young children between 6 and 24 months of age for possible communication impairment. It includes age-adjusted cutoff scores based on a nationally normed sample, with scores below the cutoff suggesting possible concern that may warrant further attention or continued monitoring.

A total of 70 children had pretest and posttest Infant / Toddler Checklist (IT) data. Of these, records indicated that 43 had formally completed the program. Sixteen children had no information on program completion, but had both pretest and posttest data and were included in these analyses. Finally, 11 children whose records indicated did not complete the program but nevertheless completed a pretest and posttest were also included in these analyses.

Reflecting the IT age range, all of these children were in preschool or not yet in school. The sample was fairly balanced between males (47.1%) and females. Slightly less than
half of the children were of Hispanic ethnicity (45.7%), and 15.7% were Black/African American. Three-out-of five participated for a full year (60.0%).

Given the IT includes multiple subscales, the following material first focuses on Total Score results, with subsequent follow-up analyses examining the Communication Composite Score, the Expressive Speech Composite Score, and the Symbolic Composite Score.

**IT Total Score Results**

Analyses found that children showed significant increases in IT Total scores following participation in the Family Literacy Program ($t(65)=7.224$, $p<.001$). Scores increased from 36.7 (SD=12.80) at the pretest to 45.4 at the posttest (see Figure 14). Further examination based on the cutoff scores can help to interpret this effect. At pretest, 25.8% of children scored below the cutoff, indicating a level of concern that may warrant further monitoring or evaluation. By posttest, this had been reduced to 15.2% of children, or a **41% reduction in the number of children flagged for possible concern**.

Follow-up analyses found that the size of this effect was greater following one year in the program, versus a single term ($F(1,64)=15.599$, $p<.001$), with scores increasing from 33.5 to 45.6 for full-year participants. Among full-year participants, the number of children falling below the desired cutoff decreased by half (50.0%).

Analyses also suggested that size of the effect varied by program ($F(10,55)=4.680$, $p<.001$), although the number of children with IT scores in many programs was too
small for any interpretation (e.g., nine programs had six or fewer cases). Results did not vary based on child gender \((F(1,64)=1.942, p=.168)\). While analyses suggested that Hispanic children may have shown less of an increase \((F(1,64)=9.151, p=.004)\), these same children showed higher initial scores and so were performing equivalent to non-Hispanic children at the posttest.

**IT Communication Composite Score**

The IT *Communication Composite Score* addresses communications skills, such as emotion and eye contact/gazing, communication, and gestures. Analyses found that children showed significant increases in Communication scores following participation in the Family Literacy Program \((t(67)=5.355, p<.001)\). Scores increased from 19.21 (SD=6.41) at the pretest to 22.72 at the posttest (see Figure 15). This increase was associated with a decline in the number of children falling below cutoff. At pretest, 20.6% of children scored below the cutoff, by posttest this had been reduced to only 4.4% of children, or a *79% reduction in the number of children flagged for possible concern*.

As with the IT Total Score findings, analyses suggested the effect was greater following one year in the program, versus a single term \((F(1,66)=8.959, p=.004)\) and that it was consistent for both males and females \((F(1,66)=2.381, p=.128)\), with the same pattern continuing for Hispanic children \((F(1,66)=6.107, p=.016)\).
Expressive Speech Composite Score

The IT Expressive Speech Composite Score addresses the use of sound and words. Analyses found that children showed significant increases in Expressive Speech scores following participation in the Family Literacy Program ($t(67)=7.298$, $p<.001$). Scores increased from 8.28 (SD=3.30) at the pretest to 10.16 at the posttest (see Figure 16). This was associated with a drop from 19.1% of children falling below cutoffs at pretest to 10.3% of children below cutoff, at posttest, or a **46% reduction in the number of children flagged for possible concern**.

As with the Total Score findings, analyses suggested the effect was greater following one year in the program, versus a single term ($F(1,66)=13.720$, $p<.001$) and that it was consistent for both males and females ($F(1,66)=3.190$, $p=.079$), with the same pattern continuing for Hispanic children ($F(1,66)=4.751$, $p=.033$).

In contrast to the Communication scores, analyses suggested that Expressive Speech scores may have increased with higher levels of literacy hours ($F(1,66)=5.693$, $p=.020$). While the true impact is difficult to disentangle from other effects (e.g., one-term vs. full-year participation, program-to-program differences), the data suggested that this may be an area where additional time in the program plays a larger role in increasing expressive language skills.
Symbolic Composite

The Infant Toddle Checklist Symbolic Composite Score addresses the use of objects and the understanding of words. Analyses found that children showed significant increases in Symbolic scores following participation in the Family Literacy Program ($t(68)=7.232, p<.001$). Scores increased from 9.36 (SD=4.37) at the pretest to 12.48 at the posttest (see Figure 17). However, in contrast to the previous two areas, this was associated with a smaller drop in the number of children flagged for possible concern – from 34.8% of at pretest to 30.4% at the posttest, or a 13% reduction.

Other results were similar to the prior findings. Analyses found evidence that the effect was greater following one year in the program, versus a single term ($F(1,67)=20.701, p<.001$) and that it was consistent for both males and females ($F(1,67)=0.732, p=.395$). Data for Hispanic children showed less of an increase ($F(1,67)=13.930, p<.001$), but their initial scores were higher than non-Hispanic children leading to equivalent posttest scores.

As with Expressive Speech scores, results also suggested that this may be an area where the number of literacy hours plays a larger role in increasing these skills ($F(1,67)=4.732, p=.033$).
**DEAF FAMILY LITERACY PROGRAMS**

As noted previously, two Family Literacy programs (*Madonna University Deaf Family Literacy* in Michigan and *Deaf Family Literacy Midsouth* in Tennessee) specifically targeted families of children with hearing loss. This included 21 adults and 22 of the total children served. Given the unique population and services, a different set of assessment instruments were used for these two programs.

**ADULT OUTCOMES**

A key goal of these programs is to enhance the American Sign Language (ASL) skills for participating adults. Therefore, adult participants were assessed on both ASL expressive and receptive skills at the start and end of the year, with approximately half of adult participants having data for both the pretest and posttest (n=9 for ASL Expressive, n=8 for ASL Receptive). Analyses found significant increases in both expressive and receptive skills following participation in the Deaf Family Literacy programs (ASL Expressive: $t(8)=3.222, p=.012$; ASL Receptive: $t(7)=3.503, p=.010$). As illustrated in Figure 18 the magnitude of these increases corresponded to very large effect sizes, with $d=1.53$ for ASL Expressive and $d=2.19$ for ASL Receptive, although it must be noted that if one began the program with relatively little ASL knowledge, large effect sizes (in terms of standard deviations) may be easier to initially achieve.

**CHILD OUTCOMES**

Children were assessed on four different instruments: the *Rosetti Infant-Toddler Language Scale* (adapted for deaf or hard of hearing children), the *Central Institute for the Deaf Preschool Developmental Rating Forms* (CID), the *Test of Early Reading Ability-Deaf/Hard of Hearing* (TERA D/HH), and the *Early Reading/Reading Checklist*. Pretest and posttest Rosetti data were only available for one student, and so it was not further considered. Analyses found significant increases across all three other measures.
The CID was designed to addresses several areas of child development specifically for children with hearing loss, including literacy, cognition, social behavior, and fine/gross motor development. It is a criterion-based instrument, which assesses children against various identified skills, rather than norms. Following participation in the Deaf Family Literacy Program, CID scores increased from 109.3 to 197.3 ($t(11)=8.092$, $p<.001$), which translates to a very large effect (d=1.32), although this increase would also include anticipate child growth and development.

The TERA D/HH addresses children’s understanding of various literacy-based constructs, alphabet knowledge, and print awareness. Following participation in the
Deaf Family Literacy Program, TERA D/HH scores also grew significantly, increasing from 13.6 to 20.8 over the course of the academic year ($t(7)=7.959$, $p<.001$), which corresponds to a very large effect ($d=1.48$). This is a norm-based instrument, although standard scores were not available for this report. Raw scores were compared to “Instructional Target Zones” identified by the developer, with performance increasing from approximately a 6 year and 3 month level to a 7 year 3 month level between the fall and spring assessments.

Finally, the Early Reading/Reading Checklist assesses alphabet knowledge, comprehension, sight word vocabulary, reading comprehension strategies, and other skills related to literacy. This is also a criterion-based instrument. Scores increased from 19.0 to 40.3 ($t(5)=3.653$, $p=.015$), which corresponded to a very large effect ($d=1.82$).

While these findings provide evidence that participation in the Deaf Family Literacy Program has positive impacts on children with hearing loss, it must be remembered that the actual number of participants is small, with only 12 children having both pretest and posttest CID scores, only 8 having both tests for the TERA D/HH, and only 6 having both tests for the Early Reading/Reading Checklist.
CONCLUSION AND RECOMMENDATIONS

This report summarizes findings from the 2015-2016 evaluation of the Barbara Bush Foundation Family Literacy program. Data included in the evaluation include pretest and posttest surveys of adult participants, and standardized assessments of adults and children. Analyses found that the program had significant, meaningful impacts on families. For caregivers, this was seen in improvements in basic education skills valuable for daily life and work. For children, this reflected increases in receptive and expressive language and decreased risk for communication and expressive language concerns in young children.

While these findings are positive, the following recommendations are offered as possible ways to further understand and document the complexity and strength of program effects on families and children.

Additional Program and Service Data. Program-to-program differences were consistently observed across outcomes. To some degree, this likely reflected fundamental differences in the ages and populations being targeted by different programs. However, it also likely reflects subtle differences in program design and implementation, as well as the specific activities and levels of engagement by participants. Additional detail regarding both program activities and adult/child participation may help to identify strategies that have the largest impact on the most families.

Parenting Data. While it is important for interventions – and evaluations – to stay focused on the specific goals of a program, feedback from the Family Literacy Program pretest and posttest surveys suggest that adult participants may have also gained substantive parenting skills. While this is not a recommendation to fundamentally change the nature of programs or activities, it may be worth further exploring the role that parenting and parent education (even if not intentionally targeted) may be having in the family-based literacy environment.
REFERENCES


