Children and Youth Services Review 99 (2019) 451–460

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth

Parents as Teachers (PAT) home-visiting intervention: A path to improved academic outcomes, school behavior, and parenting skills

Michel Lahti[a,b,*], Caroline B.R. Evans[b], Greg Goodman[a], Michele Cranwell Schmidt[a], Craig W. LeCroy[c]

[a] LeCroy & Milligan Associates, Inc., 2002 Forbes Blvd #108, Tucson, AZ 85745, United States
[b] Rhode Island College, 600 Mount Pleasant Ave, Providence, RI 02908, United States
[c] Arizona State University, School of Social Work, Tucson Campus, Tucson, AZ 85745, United States

ARTICLE INFO

Keywords:
Home-visiting
Parents as Teachers
Parenting
Academic achievement
Prevention

ABSTRACT

The time from conception to young childhood is crucial in terms of brain-, social-, emotional-, and cognitive-development. Given the impact that parenting has on child developmental outcomes, home-visiting programs may be a viable means of improving parenting and thus increasing positive child developmental outcomes. Parents as Teachers (PAT) is a home visitation program that promotes parenting skills and abilities to improve parenting skills and child development outcomes. The current study used a matched comparison group quasi-experimental design and assessed three years of academic and school disciplinary data from a school district in Arizona to determine the impact of PAT on student reading, math, and English Language skills as well as absence rates and school suspension rates relative to a non-PAT group of students. Further, the study examined whether the program influenced parenting behavior. Findings indicated that compared to the non-PAT control group, the PAT student group performed better in terms of reading and math achievement and had a significantly lower rate of absenteeism, in-school-suspensions, and out-of-school-suspension. Further, PAT parents showed increased scores on parenting measures at post-test relative to pre-test. Taken together, findings indicate that participation in the PAT program is a viable means for improving child academic outcomes and school behavior and improving parenting behavior. Implications for future research are discussed.

1. Introduction

Home-visiting programs are incredibly diverse with a variety of program models and written curriculums. While goals of home-visiting programs vary, the central goal is to positively impact parenting practices in order to improve long term child development (Haskin, Paxons, & Brooks-Gunn, 2009). More specifically, home-visiting programs aim to enhance parenting knowledge and skills in interacting with and supporting their child, to help parents connect with formal social supports, and to increase parents' ability to cope with and adapt to their new child (Guterman, 2001). These goals are achieved by having a parent educator visit the home to work directly with parents and their children.

In general, various meta-analyses and systematic reviews suggest that home-visiting programs are modestly successful in achieving the aforementioned goals (e.g., Bilukha et al., 2005; Casillas, Fauchier, Derkash, & Garrido, 2016; Filene, Kaminski, Valle, & Cachat, 2013; Guterman, 1999; Sweet & Appelbaum, 2004). For example, one meta-analysis of 51 studies (Filene et al., 2013) found a small, but significant overall mean effect size of 0.20, indicating that the comparison group (i.e., group not enrolled in a home-visiting program) was 1.5 times more likely to have poorer outcomes relative to the group receiving home-visitation services. More specifically, effect sizes had a meaningful impact for 3 of the 6 measured outcomes including maternal life course outcomes (e.g., indicators of maternal health, economic self-sufficiency, educational attainment), child cognitive outcomes (i.e., cognitive and language development), and parent behaviors and skills (e.g., positive parenting behaviors). However, birth outcomes (e.g., prematurity, low birth weight), child physical health (i.e., positive health outcome such as the absence of child injury and illness), and child maltreatment (e.g., self-report of abusive parenting practices), were not significantly impacted by home visitation programs (Filene et al., 2013). These findings highlight that home-visiting programs result in small improvements in a few important areas related to child

* Corresponding author at: LeCroy & Milligan Associates, Inc. 2002 N. Forbes, Suite 108, Tucson, AZ 85745, United States.
E-mail addresses: michel.lahti@lecroymilligan.com (M. Lahti), careyrobertson@gmail.com (C.B.R. Evans), greg@lecroymilligan.com (G. Goodman), michele@lecroymilligan.com (M.C. Schmidt), craig.lecroy@asu.edu (C.W. LeCroy).

https://doi.org/10.1016/j.childyouth.2019.01.022
Received 11 May 2018; Received in revised form 14 January 2019; Accepted 15 January 2019
Available online 17 January 2019
0190-7409/

development.

A more recent meta-analysis of 156 studies examining the effectiveness of home-visiting programs found the largest program effects were in increasing positive parenting ($d = 0.26$) and decreasing the likelihood of child maltreatment ($d = 0.22$; Casillas et al., 2016). These small, but significant effect sizes mirror Filene et al. (2013) findings, indicating that home-visiting programs do have the potential to positively impact parenting and thus influence child developmental outcomes. This meta-analysis also distilled a number of implementation factors (e.g., inclusion of role plays, supervision for home-visiting practitioners, fidelity monitoring) that impacted program effectiveness and could help explain the diverse findings with regard to home-visiting effectiveness across studies.

In addition to meta-analyses and systematic reviews, researchers have focused on evaluating the impact of different home-visiting programs. Programs such as Nurse Family Partnership, Healthy Families America, Parents as Teachers, and Home Instruction for Parents of Preschool Youth have grown due to increased federal funding and increasing attention has been focused on evaluating the evidence for these programs. As noted in the systematic reviews, most outcome studies of these programs have found some benefits, but most studies report modest impacts. This has led researchers to further examine factors that may influence outcomes such as degree of engagement with families, program fidelity, worker-family alliance, home visitor characteristics and training, quality of supervision, type of outcome measure, and the addition of mental health experts. Ongoing efforts using various evaluation methods and different outcome indicators are needed. A recent review of home visitation concluded: “There is a growing consensus among researchers and policy makers that a carefully coordinated, comprehensive service system holds the greatest promise for improving the life chances of our nation's most vulnerable young children” (Azzi-Lessing, 2013, p. 385). The current study examines the effect of one home visitation model, Parents as Teachers, a comprehensive home visitation model that provides families parent education services and focuses on linking families to early care and education, parent support groups, and community services.

1.1. Parents as Teachers (PAT) home-visiting intervention

Parents as Teachers (PAT) is an innovative, home-based intervention that provides childhood family support and parent education beginning with pregnancy and extending until entry into kindergarten. Certified parent educators are at the heart of the PAT model; these educators visit parents and children at home for one hour every week, every other week, or monthly, depending upon the specific needs of the family. The Parents as Teachers National Center (PATNC) has developed the comprehensive research- and evidence-based Foundational Curriculum PATNC (2017), covering prenatal to three years, and Foundational 2 Curriculum PATNC (2014), covering three years through Kindergarten. This curriculum is implemented as part of the PAT model during each visit with the goals of: (1) increasing parent knowledge of early childhood development (e.g., parents are taught to identify and encourage the development of age appropriate milestones in language, cognitive, social-emotional, and motor skills); (2) improving parenting practices (e.g., parents practice positive parent-child interactions that are nurturing, responsive, and supportive of the child's learning and development); (3) providing early detection of developmental delays and/or health issues (e.g., PAT children are given a health screening that evaluates health status, safety, vision, and hearing, in addition, a developmental screening is completed that assesses language, cognitive, social-emotional, and motor development. Child development is then monitored during each subsequent visit (PAT, 2018), 4) preventing child abuse and neglect (e.g., strengthening families by developing protective factors, such as parental resilience, social connectedness, and concrete support); and 5) increasing children's school readiness and school success (e.g., teaching parents to promote healthy child brain development through daily play and hands-on experiences).

The curriculum focuses on three central developmental topics: parent-child interactions (i.e., increasing positive parenting behaviors and promoting child development through parent-child activities), development centered parenting (i.e., the connection between child development and parenting including a focus on attachment, discipline, health, nutrition, safety, sleep, transitions/routines, healthy births), and family well-being (i.e., family strengths, capabilities, skills, fostering protective factors; Home Visiting Evidence of Effectiveness [HomVEE], 2017). For the PAT program of this current study, the Parent Educators are bi-lingual and conduct home visits in the family's preferred language. Additionally, curriculum and materials are available in both English and Spanish languages. The *PAT Foundational Training Guide* (PATNC, 2015) includes several modules on cultural competency, encouraging parent educators to understand the cultures of the families served in order to provide culturally appropriate parent education.

In addition to these structured home visits, PAT provides families with monthly (or more frequent) group gatherings including family activities, presentations, community events, parent cafes, and group meetings. These venues provide families with social support and the opportunity to obtain information and share experiences with their peers. Further, PAT provides annual health, hearing, vision, and developmental screenings, starting within 90 days of enrollment (HomVEE, 2017).

Past research on the effectiveness of PAT is limited with mixed results, but in line with research on other home-visiting programs generally suggests that the program successfully supports families and improves parent and child outcomes. Improved parent outcomes included increased health- and self-care literacy, decreased child maltreatment, and increased knowledge about child development and child rearing (Carroll, Smith, & Thomson, 2015; Chaiyachati, Gaither, Hughes, Foley-Schain, & Leventhal, 2018; Guastaferro et al., 2018; Pfannenstiel & Seltzer, 1989). Further, parents engaged in PAT demonstrated increased knowledge about the importance of physical stimulation, appropriate discipline, and child development compared to control parents (Pfannenstiel & Seltzer, 1989).

In terms of child outcomes, children engaged in the PAT program demonstrated higher mental processing and language abilities compared to the control group and were more likely to have positive adult relationships, coping capabilities, and engage in social play (Pfannenstiel & Seltzer, 1989). PAT participation was also significantly associated with school readiness; parents engaged in the PAT program read to their children more frequently and were more likely to enroll them in preschools, which resulted in increased school readiness. Further, preschool children from impoverished homes who participated in PAT, began preschool with school readiness scores equal to those of children from more affluent homes. School readiness was the most important predictor of achievement in third grade. These findings suggest that the PAT program resulted in improved parenting practices that likely promoted school readiness and subsequent academic achievement for children (Pfannenstiel, Seitz, & Zigler, 2002; Ziegler, Pfannenstiel, & Seitz, 2008). Taken together, the past research on PAT suggests that targeting parents' knowledge about child development and increasing positive parenting behaviors is an effective means of positively impacting child outcomes, especially as related to school readiness and academic achievement.

2. Material and methods for current study

The aim of current study was to estimate the program effect of PAT on students and their parents who participated in the program (PAT group), relative to a group of students who did not participate in the PAT program (non-PAT group). Using a matched comparison group quasi-experimental design, the current study assessed three years of academic and school disciplinary data. Informed consent was obtained

for all participants.

2.1. Research questions

There is limited research establishing if and how PAT participation effects specific forms of academic achievement such as reading and math abilities and no past research on the effect that PAT has on school absenteeism or suspensions; the current study fills these gaps by answering the following research questions:

1. To what extent does the PAT intervention result in statistically significant increases in student reading and math assessment scores compared to a comparison group of non-PAT students?
2. To what extent do English Language Learner (ELL) students who receive the PAT intervention achieve English language proficiency compared to a comparison group of ELL non-PAT participants?
3. Is there a difference between PAT and non-PAT students on child school indicators of absentee rates and number of suspensions (in-school and out-of-school) in the last three school years?
4. To what extent does the PAT intervention improve parenting skills from baseline to posttest assessment (within group change) on two parenting measures?

Based on past research, it was hypothesized that the PAT group would outperform the non-PAT group on each of the aforementioned outcomes and that parenting post-test scores would be significantly higher than pre-test scores.

2.2. Participant sampling

The initial data for this study were contained in four samples: (one) a sample of 1443 students that were former PAT participants, with birthdates ranging from September 1999 through October of 2010; (two) a sample of 16,766 students with state standardized achievement data (Arizona Measurement of Educational Readiness to Inform Teaching [AzMERIT]) in Language Arts/Reading and Mathematics from Spring 2015, 2016, and 2017; (three) a sample of 642 children pre-Kindergarten through 3rd grade with birthdates ranging from September 2008 to June 2017, who are currently receiving or had participated in PAT services during early childhood, and; (four) a parent/guardian sample with 675 PAT participants with enrollment dates ranging from November 2008 through November 2017.

Samples one and two were merged to represent former PAT participants as students in grades 3–12 and having state standardized test data. When the two samples were merged, 625 of the 1443 PAT participant records (43%) from sample one were linked to their district administrative records from sample two. (See Table 1 for demographic information).

Samples three and four were merged to represent the younger pre-school participants, linked with their parents' or guardians' data. Birthdates in sample three starting in 2008 captured data on current elementary school students up to third grade. The overlap in birthdates from sample one, with participants born through October 2010, captured additional PAT participants in the group variable. This merged sample was then used to analyze reading assessment scores (The Reading Analysis and Prescription System [RAPS 360] and The Arizona English Language Learner Assessment [AZELLA]) from School Year (SY) 2014 through SY2017.

Thus, two study samples were generated: **Sample one**- Former PAT participants (n = 625) and well-matched comparisons (n = 3125) with state student achievement data in English language arts and mathematics. **Sample two**-Younger pre-school PAT participants plus elementary school participants (n = 983) and well-matched comparisons (n = 4915) that include reading assessment scores with parent/guardian data and accompanying program dosage information. It is important to note that Sample one and Sample two represent two separate data sets for analysis and are not compared with each other, due to a limited overlap of outcome data and intervention data.

Table 1
Sample demographic comparisons of PAT and Non-PAT Groups after propensity score matching.

	Sample 1 (n = 3750)		Sample 2 (n = 5898)	
Demographic Characteristic	PAT (n = 625)	Non-PAT (n = 3125)	PAT (n = 983)	Non-PAT (n = 4915)
Gender				
Female	51.7%	51.7%	48.6%	48.6%
Male	48.3%	48.3%	51.4%	51.4%
Race/Ethnicity				
Hispanic	97.8%	97.8%	96.3%	96.3%
White	0.6%	0.6%	1.3%	1.3%
African American	0.2%	0.2%	0.2%	0.2%
Native American	1.0%	1.0%	2.0%	2.0%
Asian/Pacific Islander	0.4%	0.4%	0.1%	0.1%
Other Demographics				
Special Education Student	12.5%	12.5%	8.2%	8.2%
English Language Learner	13.6%	13.6%	11.8%	11.8%
Free/Reduced Lunch	86.1%	86.1%	75.3%	75.3%

2.3. Measures

2.3.1. Student outcomes

2.3.1.1. AzMERIT. The Arizona Measurement of Educational Readiness to Inform Teaching (AzMERIT; Arizona Department of Education [ADE] & American Institutes for Research, 2017) is an annual statewide student achievement test administered every Spring beginning in SY2015. This test assesses English Language Arts (ELA)/reading achievement for students in grades 3–11 and math achievement. For the current study, the SY2015 AzMERIT ELA and math assessment scores were used as the baseline measures for comparing PAT and non-PAT students in the analytic sample. The grade level of students in the sample ranged from 3rd to 11th grade. AzMERIT data collected in SY2017 were used as the comparison data. The AzMERIT is an online assessment taken in two or three discrete test sessions. The pool of items includes a variety of selected responses, machine scored constructed responses (i.e., graphic response, natural language, equation response, hot text, table input items), and essay responses. The Cronbach's alpha for the ELA and math AzMERIT is uniformly in the 0.90 range, which is consistent with most other similar length achievement tests (ADE & American Institutes for Research, 2017).

2.3.1.2. RAPS 360. The Reading Analysis and Prescription System (RAPS 360; Mindplay, & Methods and Solutions, Inc., 2015) is a computerized reading assessment program administered to students to identify strengths and weaknesses in various areas of reading. The RAPS 360 assesses eight domains of reading including: Comprehension, Phoneme Segmentation, Listening Vocabulary, Phonics-Decoding/Encoding, Visual Scanning Efficiency, Natural Fluency, Expected Fluency, and the Pause-Assisted Fluency and takes between 20 and 45 min to complete (Mindplay, & Methods and Solutions, Inc., 2015). The current study evaluated RAPS 360 results in the three specific areas of Comprehension, Phonics, and Fluency. The Fall SY2014 benchmark assessments in these three areas were used as baseline measures for comparing PAT and non-PAT students in the analytic sample.

2.3.1.3. AZELLA. The Arizona English Language Learner Assessment (AZELLA; ADE, & Harcourt Assessment, Inc., 2007) measures students' English language proficiency; students who score *Proficient* on the test are deemed to have sufficient English proficiency to be placed in mainstream, English speaking classrooms (ADE, 2014). The AZELLA assesses listening, speaking, reading, and writing ability; the current study evaluated the reading and writing subscales along with AZELLA

total scores, with SY2014 as the baseline measure for comparing PAT and non-PAT students in the analytic sample. Past research indicates that the Cronbach's alpha for the total AZELLA scores was 0.87 and the average machine Cronbach's alpha across grades was 0.83 (ADE, 2016).

2.3.1.4. School attendance and suspensions. School data on number of absences, in-school suspensions, and out-of-school suspensions were collected and analyzed to compare school attendance and behavior for PAT participants versus non-PAT participants for SY2015, 2016, and 2017. This is administrative data the school district routinely collects on its students.

2.3.2. Parent outcomes

2.3.2.1. KIPS. The Keys to Interactive Parenting Scales (KIPS) is a validated, structured observational assessment that examines caregiver-child interactions during play (Comfort et al., 2010; Comfort & Gordon, 2006; Comfort, Gordon, & Naples, 2011; Comfort, Gordon, & Unger, 2006). The KIPS was completed annually by PAT staff and was used to determine the extent to which the PAT intervention improved skills in parent-child relationships, learning, and confidence as parents. The data were analyzed as pretest/posttest paired scores. The KIPS is an observational measure that assesses a caregiver's interaction with a child over a 20-min time period. Scores were obtained by PAT staff who were trained in the use of the KIPS, however, there was no assessment of inter rater reliability of this measure for this part of the study. KIPS developers report results of high interrater reliability ($r = 0.88$), and good coefficient alphas of 0.89. The KIPS has significant positive correlations with Nursing Child Assessment Teaching Scale (NCATS) and the Home Observation for Measurement of the Environment (HOME) subscale scores (Comfort & Gordon, 2006).

2.3.2.2. PFS. The Protective Factors Survey (PFS; Counts, Buffington, Chang-Rios, Rasmussen, & Preacher, 2010) is a self-administered pre-post evaluation tool designed for use with caregivers receiving child abuse prevention services and to provide feedback to programs for continuous improvement and evaluation purposes. Results are designed to identify changes in protective factors and measure individual family protective outcomes. The PFS has 65 items that assess the domains of family functioning, social support (i.e., from family and friends), concrete support (i.e., access to tangible goods and services), nurturing and attachment (i.e., positive interactions between parent and child), and knowledge of parenting and child development. Studies have reported subscale reliability ranging from 0.76 to 0.89, internal consistency = 0.93, and test–retest reliability ranging from 0.63 to 0.88 (Counts et al., 2010).

2.4. Data analysis

Analysis on student outcomes was guided by What Works Clearinghouse (WWC) recommendations (U.S. Department of Education, WWC, 2017) and applied as part of a quasi-experimental, between groups design whereby treatment (PAT) and comparison (non-PAT) groups on student reading and math achievement as well as school attendance and behavior were evaluated. The WWC recommends establishing a Baseline Equivalence estimate, and suggests guidelines for implementing statistical adjustments (See Ho, Imai, King, & Stuart, 2007).

The WWC (2017) also recommends assessment of the practical importance of an intervention's effect by translating effect sizes into “improvement index” values. The improvement index represents the difference between the percentile rank consistent with the mean value of the outcome for the intervention group and the percentile rank consistent with the mean value of the outcome for the comparison group distribution. The improvement index represents the expected change in percentile rank for an average control group student if the student had received the intervention.

Finally, The WWC (2017) has adopted the use of effect size estimates known as Hedges' g or Cohen's d. It is defined as the difference between the outcome (posttest) means for the intervention group and comparison group, divided by the unadjusted pooled standard deviation of the outcome measure. An effect size of 0.2 is considered small, 0.5 moderate, and 0.8 large (Cohen, 1988).

2.4.1. Propensity score matching

Because random assignment was not conducted, propensity score analysis was used to control for possible selection bias and ensured that participants in the PAT group were statistically indistinguishable across observed covariates from participants in the non-PAT group. A propensity score is the predicted probability of receiving treatment given observed covariates (Rosenbaum & Rubin, 1983). To demonstrate that the propensity score technique was successful in matching the two groups on key demographic variables of gender, ethnicity, special education, ELL and free/reduced lunch, a non-parametric Mann-Whitney U analysis was performed. This test identifies differences in mean ranks as well as the observed significance level (p -value) between the treatment and comparison groups for Samples one and two. The test indicated that for all key demographic variables, (Sample one) PAT participants and comparisons were well-matched, sharing the same Mean Rank (1875.50), $U = 976{,}562.50$, and $p = 1.00$. The same method was used for Sample two and again, the two groups were well-matched, sharing the same Mean Rank (2949.50), $U = 2{,}415{,}722.50$, and $p = 1.00$. This is important because the more similar the two groups are at baseline, the more observed differences between the two groups could be attributed to the intervention. Following Propensity Score matching, sample one consisted of a well-matched group of 625 PAT participants with 3125 comparison students and sample two consisted of a well-matched group of 983 PAT participants with 4915 comparison students.

2.4.2. Statistical tests

For questions that examined effects of PAT on student ELA achievement and the reading assessment, Analysis of Covariance (ANCOVA) was used with the relevant ELA or reading posttest as the dependent measure (spring SY2017) and the treatment/comparison groups as the independent variable. The analysis controlled for the baseline achievement measure (pretest) in the same domain as the outcome (i.e., the ELA pretest [spring SY2015] or the reading pretest [spring SY2015]). For the question that examined the effect of PAT on math achievement, an independent samples t-test was performed on the gain scores calculated from pre-test (SY2105) to post-test (SY2017). The contrasts were the treatment/comparison conditions evaluated for significance and effect size at the student level.

RAPS 360 data were also analyzed using gain score analysis. An independent samples t-test was performed on the gain scores calculated from pre-test (SY 2014) to post-tes (SY 2017).

To test the hypothesis that there would be one or more mean differences between PAT students and comparisons in AZELLA assessment outcomes in reading, writing, and total combined scores in SY2015, 2016, and 2017, a one-way multivariate analysis of variance (MANOVA) was conducted. Prior to conducting the MANOVA, a series of Pearson correlations were performed between all dependent variables to test the MANOVA assumption that the dependent variables would be correlated with each other in the moderate range (i.e., 0.20–0.60; Meyers, Gamst, & Guarino, 2006). Multivariate tests with exact statistics, degrees of freedom, significance level, and effect size are reported.

Absence rates and school suspensions were analyzed to examine the effect of PAT participation on school indicators of attendance and behavior. A four by two chi square test of independence (four categories of absence rates or suspensions by two groups) was calculated comparing the observed counts and percentages between PAT and non-PAT students. Similar tests were conducted on school indicators among the

subgroups of special education, ELL, and financial status. The contrasts were the treatment/comparison conditions at the student level evaluated for chi square value, degrees of freedom, significance, and Cramer's V effect size.

Prior to conducting the analysis on both KIPS and the PFS, the assumption of normally distributed difference scores was examined. The assumption was considered satisfied for both assessments, as the skew and kurtosis levels were less than the maximum allowable values for a *t*-test (i.e., skew $< |2.0|$ and kurtosis $< |9.0|$; Posten, 1984). Dependent samples *t*-tests were performed.

3. Results

3.1. Propensity score matching

In both the PAT and non-PAT groups within each sample, the demographic distribution of students was identical. Demographic characteristics between samples differed slightly. Females were the slim majority in Sample one (51.7%) and males were the slim majority in Sample two (51.4%). The vast majority were of Hispanic origin in both Sample one (97.8%) and Sample two (96.3%). While the two Samples differed in the proportion of students who were identified as special education students, English language learners (ELLs) and students who were on free/reduced lunch, these subgroups were identical between PAT and comparisons within each Sample. The largest dissimilarity was free/reduced lunch students between Sample one (86.1%) and Sample two (75.3%). Sample one and Sample two represent two separate data sets for analysis and are not compared with each other, due to a limited overlap of outcome data and intervention data.

3.2. English Language Arts (ELA)/reading achievement-AzMERIT

The results found a difference in AzMERIT ELA/reading achievement outcome scores, with PAT students performing better than the matched comparison group. The PAT participants had a statistically significantly larger adjusted posttest mean than the comparison students, after controlling for pretest scores. Baseline equivalence on ELA scale scores showed that the difference in pretest means of the two groups was 12% (0.12) of the pooled standard deviation (PSD). ANCOVA was performed because mean differences were < 25% of the PSD between the PAT group ($n = 382$) and the comparison group ($n = 1175$). ANCOVA showed a statistically significant difference in the adjusted posttest means between PAT participants ($M = 2544.64$; $SD = 32.22$) and comparisons ($M = 2541.64$; $SD = 30.44$) after controlling for the pretest covariate. The fixed factor 'group' was significant, $F(1, 1554) = 6.00$, $p = .014$ as was the covariate pretest, $F(1, 1556) = 1883.41$, $p < .001$. The homogeneity of variance assumption held with a Levene's test of 0.591. Hedges' *g* effect size calculation = 0.10, which is a small effect size based on Ellis' (2010) guidelines. The resulting improvement index was +4, corresponding to moving performance for the average student from the 50th to the 54th percentile of the comparison group distribution if the student had received the intervention.

3.3. Math achievement-AzMERIT

Results found a difference in AzMERIT math outcome gain scores, with PAT students performing better than non-PAT comparison students. The PAT participants demonstrated a statistically significantly larger gain in mean score than the comparison students. Baseline equivalence on math scale scores (MSS) between the two groups showed that the difference in pretest means was 32% (0.32) of the PSD. Gain Score Analysis was used as a result of pretest differences being > 25% of the PSD. The PAT participant group ($n = 364$) had a pretest mean of $M = 3567.07$ ($SD = 50.15$) and a posttest mean of $M = 3623.97$ ($SD = 46.89$); by contrast, the comparison group ($n = 1193$) demonstrated a numerically larger pretest mean ($M = 3584.75$; $SD = 57.28$) and posttest mean ($M = 3632.93$; $SD = 50.82$). An independent samples *t*-test was performed on gain scores calculated from pretest to posttest. The assumption of homogeneity of variances was tested and satisfied via Levene's *F* test $F(1555) = 2.21$, $p = .137$. The independent samples *t*-test on gain scores was associated with a statistically significant effect, $t(1555) = 4.67$, $p < .001$. Thus, PAT participation was associated with a statistically significantly larger gain in mean math scale scores from 2015 to 2017 compared to non-PAT participation. Cohen's *d* (effect size) was estimated at 0.28 (a small to moderate effect size [0.2 is small, 0.5 moderate, 0.8 large; Cohen, 1988]), and the improvement index was +11.

3.4. Reading achievement-RAPS 360

PAT participants showed larger mean gains from pretest to posttest compared to the non-PAT participants and PAT participation was associated with a larger adjusted posttest mean than non-PAT participation after controlling for pretest scores; however, the difference observed was not statistically significant. Baseline equivalence on RAPS 360 Fluency scale scores showed that the difference in pretest means of the two groups was 6% (0.06) of the PSD. ANCOVA was conducted between the PAT group ($n = 362$) and the comparison group ($n = 932$) and did not show a statistically significant difference in the adjusted posttest means between the PAT participant's ($M = 131.00$; $SD = 58.52$) and comparison group ($M = 126.55$; $SD = 58.80$), after controlling for the pretest covariate. The fixed factor 'group' was not significant, $F(1, 1291) = 2.03$, $p = .155$, however the covariate pretest was significant, $F(1, 1293) = 456.45$, $p < .001$. The homogeneity of variance assumption held with a Levene's test of 0.886. Hedges' *g* effect size calculation = 0.08 (Small effect size; Cohen, 1988). Although PAT participation was associated with a larger adjusted posttest mean than non-PAT participation after controlling for pretest scores, the difference was not statistically significant.

3.5. English language learner achievement-AZELLA

The results suggest PAT ELL students performed better than non-PAT ELL students on the AZELLA. PAT participation was associated with consistently larger means compared to non-PAT participation in reading, writing, and total scores for all three years and a statistically significant MANOVA effect was obtained, as six of the nine *F*-tests were statistically significant. A one-way multivariate analysis of variance (MANOVA) was conducted to test the hypothesis that there would be one or more mean differences between PAT students ($n = 60$) and comparisons ($n = 229$) on AZELLA assessment outcomes in reading, writing, and total combined scores in the last three years (2015, 2016, and 2017). Prior to conducting the MANOVA, a series of Pearson correlations were performed between the dependent variables to test the assumption that the dependent variables would be correlated in the moderate range (i.e., 0.20–0.60; Meyers et al., 2006). A meaningful pattern of correlations was observed (i.e., correlations ranged from 0.17 to 0.80; additional results available upon request) among most of the dependent variables, suggesting the appropriateness of a MANOVA. Additionally, the Box's M value of 113.95 was associated with a *p* value of 0.006, which was interpreted as non-significant based on Huberty and Petoskey's (2000) guideline (i.e., $p < .005$). Thus, the covariance matrices between the groups were assumed to be equal for the purposes of the MANOVA.

The homogeneity of variance assumption was tested for all nine dependent variable scales using a series of Levene's *F* tests. The assumption was considered satisfied, even though one of the nine Levene's *F* tests were statistically significant ($p < .05$). Specifically, although the variance associated with the AZELLA Writing (2016) scale was not homogenous, an examination of the standard deviations revealed that none of the largest standard deviations were more than four

Table 2
One-way F tests with AZELLA reading, writing, and total scores as dependent variables and study group as the independent variable, 2015–2017.

AZELLA	Levene's		*F* - test		partial η2	PAT-Group		Non-PAT Group	
	F	*p*	*F*	*p*		*M*	*SD*	*M*	*SD*
2015									
Read	0.068	0.795	6.54	0.011	0.022	234.53	29.63	225.59	22.47
Write	2.12	0.146	6.93	0.009	0.024	233.65	26.88	222.41	30.04
Total	0.871	0.351	6.58	0.011	0.022	2397.63	53.40	2375.27	61.73
2016									
Read	0.001	0.997	2.82	0.094	0.010	233.05	22.89	227.71	21.70
Write	4.13	0.043	11.12	0.001	0.037	241.87	19.95	229.80	26.09
Total	0.802	0.371	6.70	0.010	0.023	2427.43	42.78	2409.58	48.75
2017									
Read	0.091	0.764	0.653	0.420	0.002	236.27	17.11	233.86	21.34
Write	0.967	0.326	4.03	0.046	0.014	240.98	29.89	233.92	22.60
Total	1.91	0.168	3.69	0.056	0.013	2455.72	43.80	2441.80	51.41

Note. PAT-Group (n = 60); Non-PAT Group (n = 229), p = .032.

times the size of the corresponding smallest, suggesting that the *F*-test would be robust in this case (Howell, 2007). Six of the nine *F*-tests were statistically significant. Effect sizes (partial η2) were small ranging from 0.002 (reading 2017) to a high of 0.037 (Writing 2016). PAT students showed consistently larger means than comparisons and a statistically significant MANOVA effect was obtained, Wilks' Lambda = 0.94, $F(9, 279) = 2.08$, p = .032. The multivariate effect size was estimated at 0.063 (small effect size; Cohen, 1988)., which implies that 6.3% of the variance in the canonically derived dependent variable was accounted for by the group variable (PAT participation / comparisons; See Table 2).

3.6. Absenteeism and suspension outcomes

Results suggest that PAT participants had lower rates of absenteeism than comparison students, as analysis showed a statistically significant chi square (independence) tests in favor of PAT participants across absence rate categories for SY2015, 2016, and 2017. A four by two chi square test of independence (four categories of absence rate categories by two groups) was calculated comparing the observed counts and percentages between PAT and non-PAT students by categories of absence rates for three school years.

As shown in Table 3, statistically significant chi square (independence) tests were observed in comparing PAT participants and comparisons across absence rate categories for SY2015 ($\chi^2 = 88.38$, $p < .001$, df = 3, n = 5898); SY2016 ($\chi^2 = 32.16$, $p < .001$, df = 3, n = 5898); and SY2017 ($\chi^2 = 59.75$, $p < .001$, df = 3, n = 5898). Cramer's *V* (shown in the far-right column) is reported showing the strength of the association between the two variables.

School indicators were analyzed to examine the effect of PAT participation and number of suspension days (in-school and out-of-school). Analyzing three SYs of in-school and out-of-school suspension data, PAT participants showed a statistically significant lower number of suspension days than comparisons for one school year. There was a significant chi square (independence) test between PAT participants and comparisons across three categories of in-school suspension days for SY2015. All other categories of days suspended in SY2016 and SY2017 were not significant.

A three by two chi square test of independence (three categories of suspension days by two groups) was calculated comparing the observed counts and percentages between PAT and non-PAT students associated with in-school suspension days, and the same chi square test was also conducted on the same students with out-of-school suspensions - for three school years.

For in-school suspensions the frequencies and percentages cross tabulated in Table 4, show a significant chi square (independence) test between PAT participants and comparisons across categories of in-school suspension days for SY2015 ($\chi2 = 6.85$, p = .033, df = 2, n = 5898). Significance was not observed for SY2016 ($\chi2 = 4.89$, p = .087, df = 2, n = 5898) or SY2017 ($\chi2 = 2.43$, p = .297, df = 2, n = 5898).

For out-of-school suspensions, frequencies and percentages cross tabulated in Table 4, show there was a significant chi square (independence) test between PAT participants and comparisons across categories of out-of-school suspension days for SY2015 ($\chi2 = 7.03$, p = .030, df = 2, n = 5898). Significance was not observed for SY2016 ($\chi2 = 4.21$, p = .239, df = 2, n = 5898) or SY2017 ($\chi2 = 3.45$, p = .178, df = 2, n = 5898).

3.7. Parenting outcomes

To test the null hypothesis that pretest and posttest mean KIPS scores were equal among the same group of PAT parents, a dependent samples *t*-test was performed. The null hypothesis of equal means was

Table 3
Chi square test of independence on categories of absence rates by PAT and comparison students, SY2014–2017.

SY/group	0% Absences *n* (%)	0.01–10.0% *n* (%)	10.01–20.0% *n* (%)	> 20.0% *n* (%)	Total *n* (%)	ES *V*
SY2015						
PAT-Group	501 (51.0%)	141 (14.3%)	23 (2.3%)	318 (32.3%)	983 (100.0%)	0.12
Non-PAT Group	1773 (36.1%)	841 (17.1%)	317 (6.4%)	1984 (40.0%)	4915 (100.0%)	
SY2016						
PAT-Group	505 (51.4%)	169 (17.2%)	28 (2.8%)	281 (28.6%)	983 (100.0%)	0.07
Non-PAT Group	2106 (42.8%)	825 (16.8%)	237 (4.8%)	1690 (35.5%)	4915 (100.0%)	
SY2017						
PAT-Group	510 (51.9%)	203 (20.7%)	27 (2.7%)	243 (24.7%)	983 (100.0%)	0.10
Non-PAT Group	2186 (44.5%)	755 (15.4%)	248 (5.0%)	1726 (35.1%)	4915 (100.0%)	

Note. PAT-Group (n = 983); Non-PAT Group (n = 4915). ES = effect size for Chi Square is Cramer's *V*.

Table 4
Chi square test of independence on categories of days suspended by PAT and comparison students, SY2015–2017.

School year/group	In-school suspension days			Out-of-school suspension days		
	0 *n* (%)	1–3 *n* (%)	4 + *n* (%)	0 *n* (%)	1–3 *n* (%)	4 + *n* (%)
SY2015						
PAT-Group	973 (99.0%)	6 (0.6%)	4 (0.4%)	967 (98.4%)	14 (1.4%)	2 (0.2%)
Non-PAT Group	4805 (97.8%)	84 (1.7%)	26 (0.5%)	4796 (97.6%)	65 (1.3%)	54 (1.1%)
SY2016						
PAT-Group	967 (98.4%)	14 (1.4%)	2 (0.2%)	961 (97.8%)	12 (1.2%)	10 (1.0%)
Non-PAT Group	4780 (97.3%)	102 (2.1%)	33 (0.7%)	4745 (96.5%)	86 (1.7%)	84 (1.7%)
SY2017						
PAT-Group	957 (97.4%)	22 (2.2%)	4 (0.4%)	957 (97.4%)	18 (1.8%)	8 (0.8%)
Non-PAT Group	4736 (96.4%)	151 (3.1%)	28 (0.6%)	4736 (96.4%)	103 (2.1%)	76 (1.5%)

Table 5
Dependent *t* - test of PAT Parents' KIPS and PFS Pretest and Posttest Scores.

Assessment	*N*	*M*	*SD*	Diff	*t*	*p*	Effect Size
KIPS Mean Scores							
Pretest	182	3.66	0.61	0.21	5.01	< 0.001	0.37
Posttest		3.87	0.71				
PFS Family Functioning Subscale							
Pretest	250	6.10	0.89	0.18	2.85	< 0.005	0.18
Posttest		6.29	0.87				
PFS Social Support Subscale							
Pretest	250	6.25	0.98	0.18	2.80	< 0.006	0.18
Posttest		6.42	0.83				
PFS Concrete support Subscale							
Pretest	248	4.01	2.05	0.91	5.31	< 0.001	0.34
Posttest		4.91	2.06				
PFS Nurturing and Attachment Subscale							
Pretest	250	6.69	0.38	0.05	1.70	0.091	0.11
Posttest		6.74	0.37				

Note. Effect Size = Cohen's d.

rejected $t(182) = 5.01$, $p < .001$. Thus, the post-program (posttest) KIPS mean score (3.87) was statistically significantly higher than the pre-program (pretest) KIPS mean score (3.66). Cohen's *d* was estimated at 0.37, which is a medium effect based on Cohen's (1992) guidelines (See Table 5).

A dependent samples *t*-test was also conducted on the four PFS subscale scores to test for equal mean scores of PAT parents from pretest to posttest. Again, the assumption of normally distributed difference scores was examined. The assumption was considered satisfied, as the skew and kurtosis levels were estimated at −0.087 and 0.982 respectively. The correlation between the two conditions was estimated at $r = 0.53$, $p < .001$, suggesting that the dependent samples *t*-test was appropriate in this case. The null hypothesis of equal means was rejected for three of the four subscales. PAT parents showed significant improvements over time in the subscale domains of family functioning $t(250) = 2.85$, $p < .005$, social support $t(250) = 2.80$, $p < .006$, and concrete support $t(248) = 5.31$, $p < .001$ from pretest to posttest. The null hypothesis of equal means was rejected as the post-program means for these three subscales was statistically significantly higher than the pre-program means. Cohen's *d* was estimated at 0.18 for family functioning and social support, which is a small effect; the effect size for the concrete support subscale was 0.34, which is a small to medium effect (See Table 5).

4. Discussion

Overall, findings indicated that the PAT home visiting program was associated with statistically significant improvements in child academic outcomes (e.g., ELA/reading achievement, math, ELL English proficiency), and parenting skills. The current study extended past research on PAT by also ascertaining that PAT participation was associated with statistically significant lower absenteeism and in- and out-of-school-suspensions.

In line with Ecological Systems Theory (Bronfrenbrenner, 1979), the PAT program focuses on improving the family Microsystem by increasing parent knowledge about child development as well as improving parenting skills, both of which help improve parent-child social interactions and ultimately result in better developmental outcomes for the child. In addition, PAT strengthens family ties to the larger community (including various other Microsystems and the cultural Macrosystem) through access to medical care, community events, and social activities. Rather than viewing child development as only connected to the individual child, PAT uses an ecological view to improve the child's family Microsystem, which greatly benefits the child and his/her development. By increasing parents' knowledge of child development and parenting skills, PAT helps parents become more sensitive to their children's developmental needs, which likely improves the quality of parent/child interactions and the overall family Microsystem. The PAT program also views families as interconnected systems where the behavior of each family member is impacted by the behavior of all other family members. Thus, changing parent behavior is a key mechanism to change child behavior and the PAT program capitalizes on this by educating parents and supporting them in improving their parenting skills. The findings of the current study suggest that this approach is effective not only in terms of increasing parenting skills, confidence, and knowledge about child development and improving overall family functioning (e.g., social support, attachment), but also in terms of improving childhood academic outcomes and school behavior.

Specifically, PAT participants displayed statistically significant higher ELA/reading achievement scores relative to the non-PAT group as measured by the AzMERIT (Hedges' $g = 0.10$); although this effect size is considered small (Cohen, 1988), it was the equivalent of moving a child from the 50th percentile to the 54th percentile in ELA/reading achievement. In terms of math, PAT participation was associated with statistically significant higher math gain scores for the PAT group as compared to the non-PAT group as measured the by AzMERIT (Hedges' $g = 0.28$). These findings suggest that the PAT program played a role in the improvements in ELA/reading and math achievement displayed by the PAT student group. Because a mediational relationship between parenting skills/knowledge and academic outcomes was not tested, we cannot say for sure that improved parenting was the mechanism of change that caused the increased ELA/reading and math achievement. However, we can surmise that because PAT focuses mainly on increasing parental knowledge of child development and on improving parenting skills, that these improvements played a role in the improved academic achievement outcomes displayed by PAT participants. Future research should examine the mechanisms of change of the PAT

program, but such analysis was beyond the scope of the current study.

Given the well documented benefits of reading to children, it follows that PAT participation was associated with increased ELA/reading achievement. More specifically, past research indicates that PAT participation was associated with increased reading for fun in participating families (Carroll et al., 2015), suggesting that one of the benefits of PAT is increased reading exposure for young children. Reading to young children is beneficial and leads to increased future reading acquisition, vocabulary, and reading ability (See Mol, Bus, de Jong, & Smeets, 2008 for a review; See Senechal & Young, 2008 for a review). An in-depth analysis of the text of 100 children's books found that books offer more unique words compared to child directed speech, indicating that reading to children exposes them to more word diversity, resulting in language and reading improvements (Montag, Jones, & Smith, 2015). The benefits of reading to young children are especially pronounced when dialogic reading is used (e.g., parents encourage children to discuss the pictures and involve the child in the story rather than just reading to them; See Mol et al., 2008 for a review). PAT likely encourages such dialogic reading by teaching parents how to engage their children in developmentally appropriate conversations in general and also related to books. Taken together, this research on the benefits of reading and the fact that PAT participation was associated with increased reading, provides an explanation for the higher ELA/reading achievement scores demonstrated by PAT students in the current study. It is conceivable that as PAT parents read more to their children and engaged in more developmentally appropriate conversations with their children, that they also focused on other academic areas such as math. Perhaps math achievement increased in PAT students because PAT parents were overall more focused on the importance of academics in the home.

It is interesting and somewhat enigmatic that although PAT participants demonstrated significantly higher AzMERIT ELA/reading achievement scores relative to the non-PAT group, the increase in RAP-360 scores did not reach statistical significance. PAT focuses on increasing parenting skills and abilities, which certainly includes spending more time reading to young children, which would result in improved overall reading achievement as measured by the AzMERIT. However, PAT does not instruct parents on how to teach their children to read or how to improve the specific reading skills measured by the RAPS-360. Perhaps if the PAT program were slightly modified to include more parent information on attainment of specific reading acquisition skills, the PAT program would have had more of an effect on the RAPS-360 scores. This is an important area for future research.

It is also noteworthy that the PAT group had statistically significant higher scores in English language ability for ELL students as measured by the AZELLA as compared to the non-PAT group. ELL students are a particularly high-risk group as they face many unique stressors including immigration, family separations, poverty, discrimination, violence, and cultural conflicts that increase the likelihood for negative academic outcomes (Suarez-Orozco, Suarez-Orozco, & Todorova, 2008). For example, ELL students often have lower levels of academic achievement compared to their English-Proficient (EP) classmates and score lower on reading and math proficiency tests (Abedi & Lord, 2001; Ballantyne, Sanderman, & Levy, 2008; National Assessment of Educational Progress [NAEP], 2009a, 2009b). This decreased academic achievement appears to erode ELL students' academic confidence as one study found that this vulnerable group rated themselves significantly lower on academic efficacy compared to their non-ELL counterparts (LeClair, Doll, Osborn, & Jones, 2009). This research highlights the importance of implementing supports for ELL students that can bolster their academic achievement and current findings indicate that PAT is one such program. Providing PAT for ELL students appears to be particularly important given their increased risk for negative academic outcomes. Providing PAT to this vulnerable group was associated with higher scores on a test of English language ability, which could serve to keep ELL students engaged in school, a vital protective factor for future positive developmental outcomes.

PAT participation was also associated with statistically significantly lower absenteeism as well as in- and out-of-school-suspensions for one school year (SY2015). This lower rate of absences could indicate that in the first year of the PAT parents were more aware of the importance of education as a result of PAT participation and therefore were more inclined to ensure that their children attended school. Indeed, past research indicates that PAT parents were significantly more likely to enroll their children in preschool (Pfannenstiel et al., 2002; Ziegler et al., 2008), suggesting PAT helped parents understand the importance of education. However, as program participation extended over two more years, it is possible that parents become overwhelmed with getting their child to school and thus the lower absenteeism was not sustained over the three-year program. The lower rates of school suspensions suggests that PAT participation was associated with improved child behavior. This finding highlights the notion that altering and improving parenting behavior and skills could in fact result in improved child behavior as well. However, caution must be warranted in interpreting these results as mediational analysis was not conducted and further research on the causal links between improved PAT parenting behavior and child behavior is needed.

Finally, findings of the current study confirm that PAT participation was associated with statistically significantly higher parenting quality per the KIPS mean scores, and family functioning, social support, and concrete support, as measured by the PFS subscales. The effect sizes were moderate (KIPS [ES = 0.48] PFS [ES = 0.18, 0.18, 0.34]; Cohen, 1988), but quite impressive for intervention research. This finding highlights the success of the PAT curriculum in improving parenting skills, knowledge, and confidence, which is the heart of the PAT program. Based on these results, it appears that the PAT curriculum provides parents with the necessary skills and knowledge to actually improve their parenting and the overall family Microsystem. Future research should investigate and confirm that this improved parenting is actually the mechanism of change that is responsible for improvements in child academic and behavioral outcomes.

5. Limitations

Findings of the current study must be understood in light of specific limitations. First, the lack of random assignment among treatments and comparisons leads to non-equivalent test groups which can reduce internal validity, and conclusions about causality are less definitive in quasi-experimental designs. Straight comparisons of outcomes between PAT participants and a comparison group do not meet the requirements for making a clear causal inference. For these reasons, it was imperative that the two groups were well-matched on key variables of gender, ethnicity, special education status, ELL and free/reduced lunch status. This is important because the more similar the two groups are at baseline, the more observed differences between the two groups could be attributed to the intervention. Yet, even though Propensity Score matching was successfully applied in identifying similar groups on key demographics mentioned above, the inability to rule out the possibility that outcomes are unrelated to program participation still exist.

Second, the initial data on children as PAT participants for this study were contained in two files: a file of 1443 students that were former PAT participants, with birthdates ranging from September 1999 through October of 2010, and a file of 642 children pre-Kindergarten through 3rd grade with birthdates ranging from September 2008 to June 2017, who are receiving or had participated in PAT services during early childhood. The overlap in birthdates from these two files with participants born from 9/2008 through 10/2010, captured a very small number of PAT participants in the group variable. This effected the ability to provide PAT dosage analysis, as data on the frequency of home visits are available for children who participated in the PAT intervention after 2011, but not before. A within group analysis - within the PAT group only, would have allowed us to address dosage-related

research questions such as the relationship between the duration of PAT participation with their school achievement outcomes, or the relationship between frequency of home visits and various school outcomes. Future research should examine the impact of program dosage on child and parent outcomes.

Third, it would have been ideal to have a control group of non-PAT parents in order to compare changes in KIPS and PFS scores to the PAT parent group. Additionally, the KIPS measure was completed by PAT staff without assessment of interrater reliability and may include bias as staff make observations of families on a rating scale and are likely to believe their efforts are leading to positive outcomes.

Fourth, our entire sample came from Arizona and the majority of the sample was Latino/Hispanic. These unique sample characteristics indicate that findings should be generalized to other geographic regions and races with caution. Future PAT research should use more geographically and ethnically/racially diverse samples in order to better understand how the program might operate differently across populations.

Finally, outcome data on children's academic performance in elementary school is limited until more PAT children enter the school system and take part in assessments. This lack of connection (from limited passage of time) between dosage data (post-2011) and PAT children assessed in the public-school system produced very little overlap between PAT participation as late as 2011 and, having been in the school district long enough to establish records on their academic measures. Yet, the improvements in data collection and maintenance starting in 2011 by PAT staff create a promising opportunity for future analyses, as PAT students move through the school district and complete their academic assessments.

6. Conclusion

Findings of the current study suggest that participation in the PAT program was associated with statistically significantly higher scores in student level outcomes including ELA/reading achievement, math achievement, and ELL student English language proficiency as well as statistically significantly lower absenteeism and school suspensions. For parents, PAT participation was associated with statistically significantly higher parenting knowledge of child development, parent nurturance, parent-child attachment, family functioning, family social support, parent confidence, and parenting skills. These findings highlight the utility of the PAT home-visitation program in improving the family Microsystem, which likely helped improve children's academic achievement outcomes. Overall, current findings reinforce the importance of early and ongoing supportive interventions in terms of improving family functioning and child developmental outcomes.

Funding

This work was supported by the Parents As Teachers National Program. The sponsor did not have any role in the study design; no role in the collection, analysis and interpretation of data; nor in the writing of the report. The sponsor did not participate in the decisions as to where and or how to submit the article for publication.

Declaration of interest

LeCroy & Milligan Associates, Inc. was hired as a consultant to conduct the current research and this research was funded by Parents as Teachers. The sponsor did not have any role in the study design; no role in the collection, analysis, or interpretation of data; nor in the writing of the report. The sponsor did not participate in the decisions as to where and/or how to submit the article for publication.

Declaration of interest statement

LeCroy & Milligan Associates, Inc. was hired as a consultant to conduct the current research.

Role of the funding source statement

The current study was funded by Parents as Teachers. The sponsor did not have any role in the study design; no role in the collection, analysis, or interpretation of data; nor in the writing of the report. The sponsor did not participate in the decisions as to where and/or how to submit the article for publication.

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*, 219–234. https://doi.org/10.1207/S15324818AME1403_2.

Arizona Department of Education (2014). *Arizona english language learner assessment (AZELLA): Evidence-based standard setting review report kindergarten placement test.* Phoenix, AZ: Author.

Arizona Department of Education (2016). *Arizona english language learner assessment: 2016 technical report.* Phoenix, AZ: Author.

Arizona Department of Education, & & Harcourt Assessment, Inc. (2007). *Arizona english language learner assessment: Technical manual.* Phoenix, AZ: Author.

Arizona Department of Education, & American Institutes for Research (2017). *Annual technical report: arizona statewide assessment in english language arts and math 2015–2016 school year.* Phoenix, AZ: Author.

Azzi-Lessing, L. (2013). Serving highly vulnerable families in home-visitation programs. *Infant Mental Health Journal, 34*(5), 376–390. https://doi.org/10.1002/imhj.21399.

Ballantyne, K., Sanderman, A., & Levy, J. (2008). *Educating English Language Learners: Building teacher capacity.* Washington, DC: National Clearinghouse for English Language Acquisition.

Bilukha, O., Hahn, R. A., Crosby, A., Fullilove, M. T., Liberman, A., Mosicicki, E., ... Task Force on Community Preventive Services (2005). The effectiveness of early childhood home visitation in preventing violence. *American Journal of Preventive Medicine, 28*(2S1), 11–39. https://doi.org/10.1016/j.amepre.2004.10.004.

Bronfrenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Carroll, L. N., Smith, S. A., & Thomson, N. R. (2015). Parents as teachers health literacy demonstration project: Integrating an empowerment model of health literacy promotion into home-based parent education. *Health Promotion Practice, 16*(2), 282–290. https://doi.org/10.1177/1524839914538968.

Casillas, K. L., Fauchier, A., Derkash, B. T., & Garrido, E. F. (2016). Implementation of evidence-based home visiting programs aimed at reducing child maltreatment: A meta-analytic review. *Child Abuse & Neglect, 53*, 64–80. https://doi.org/10.1016/j.chiabu.2015.10.009.

Chaiyachati, B. H., Gaither, J. R., Hughes, M., Foley-Schain, K., & Leventhal, J. M. (2018). Preventing child maltreatment: Examination of an established statewide home-visiting program. *Child Abuse & Neglect, 79*, 476–484. https://doi.org/10.1016/j.chiabu.2018.02.019.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 12*, 155–159.

Comfort, M., & Gordon, P. R. (2006). The Keys to Interactive Parenting Scale (KIPS): A practical observational assessment of parenting behavior. *NHSA Dialog: A Research-To-Practice Journal for the Early Intervention Field, 9*(1), 22–48. https://doi.org/10.1207/s19309325nhsa0901_4.

Comfort, M., Gordon, P. R., English, B., Hacker, K., Hembree, R., Knight, C. R., & Miller, C. (2010). Keys to Interactive Parenting Scale: KIPS shows how parents grow. *Zero to Three Journal, 30*(4), 33–39. Retrieved from https://www.zerotothree.org/resources/series/zero-to-three-journal.

Comfort, M., Gordon, P. R., & Naples, D. (2011). KIPS: An evidence-based tool for assessing parenting strengths and needs in diverse families. *Infants & Young Children, 24*(1), 56–74. https://doi.org/10.1097/IYC.0b013e3182001bd3.

Comfort, M., Gordon, P. R., & Unger, D. G. (2006). Keys to Interactive Parenting Scale: A window into many facets of parenting. *Zero to Three Journal, 26*(5), 37–44. Retrieved from https://www.zerotothree.org/resources/series/zero-to-three-journal.

Counts, J. M., Buffington, E. S., Chang-Rios, K., Rasmussen, H. N., & Preacher, K. J. (2010). The development and validation of the protective factors survey: A self-report measure of protective factors against child maltreatment. *Child Abuse and Neglect, 34*(10), 762–772. https://doi.org/10.1016/j.chiabu.2010.03.003.

Ellis, P. D. (2010). *The Essential Guide to effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research results.* New York: Cambridge University Press9–11.

Filene, J. H., Kaminski, J. W., Valle, L. A., & Cachat, P. (2013). Components associated with home visiting program outcomes: A meta-analysis. *Pediatrics, 132*(2), S100–S109. https://doi.org/10.1542/peds.2013-1021H.

Guastaferro, K., Lai, B. S., Miller, K., Chathama, J. S., Whitaker, D. J., Self-Brown, S., ... Lutzker, J. R. (2018). Braiding two evidence-based programs for families at risk: Results of a cluster randomized trial. *Journal of Child and Family Studies, 27*, 535–546. https://doi.org/10.1007/s10826-017-0886-2.

Guterman, N. (1999). Enrollment strategies in early home visitation to prevent physical child abuse and neglect and the "universal versus targeted" debate: A meta-analysis of population-based and screening-based programs. *Child Abuse & Neglect, 23*(9), 863–890. https://doi.org/10.1016/S0145-2134(99)00062-9.

Guterman, N. (2001). *Stopping child maltreatment before it starts: Emerging horizons in early home visitation services.* Thousand Oaks, CA: Sage.

Haskin, R., Paxons, C., & Brooks-Gunn, J. (2009). Social science rising: A tale of evidence shaping public policy. *The Future of Children, Fall,* 1–7. Retrieved from https://futureofchildren.princeton.edu/publications.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*(3), 199–236. https://doi.org/10.1093/pan/mpl013.

HomVEE (2017). *The Administration for Children and Families (ACF), Office of Planning, Research, and Evaluation (OPRE), part of the U.S. Department of Health and Human Services (HHS).* OPRE Report #2018 - 114 https://homvee.acf.hhs.gov/homevee_executive_summary_brief.pdf#Brief1.

Howell, D. C. (2007). *Statistical methods for psychology.* Belmont, CA: Thompson Wadsworth.

Huberty, C. J., & Petoskey, M. D. (2000). Multivariate analysis of variance and covariance. In H. Tinsley, & S. Brown (Eds.). *Handbook of applied multivariate statistics and mathematical modeling.* New York: Academic Press.

LeClair, C., Doll, B., Osborn, A., & Jones, K. (2009). English Language Learners' and non-English Language Learners' perceptions of the classroom environment. *Psychology in the Schools, 46*, 568–577. https://doi.org/10.1002/pits.20398.

Meyers, L. S., Gamst, G., & Guarino, A. (2006). *Applied multivariate research: Design and interpretation.* Thousand Oaks, CA: Sage Publishers.

Mindplay & Methods and Solutions, Inc (2015). *Universal Screener RAPS 360: Reading Analysis & Prescription System.* Tucson, AZ: Author.

Mol, S. E., Bus, A. G., de Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic parent-child book readings: A meta-analysis. *Early Education and Development, 19*(1), 7–26. https://doi.org/10.1080/10409280701838603.

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science, 26*(9), 1489–1496. https://doi.org/10.1177/0956797615594361.

National Assessment of Educational Progress (2009a). The nation's report card: Mathematics. Retrieved from http://nationsreportcard.gov/math_2009/gr4_national.asp?subtab_id=Tab_7&tab_id=tab1#tabsContainer.

National Assessment of Educational Progress (2009b). The nation's report card: Reading. Retrieved on from http://nationsreportcard.gov/reading_2009/nat_g4.asp?subtab_id=Tab_7&tab_id=tab1#tabsContainer.

Parents as Teachers (2018). Parents as teachers. Retrieved from https://parentsasteachers.org.

Parents as Teachers National Center (2014). *Foundational 2 curriculum: 3 years through kindergarten.* St. Louis, MO: Author.

Parents as Teachers National Center (2015). *Parents as teachers foundational training guide.* St. Louis, MO: Author.

Parents as Teachers National Center (2017). *Foundational curriculum.* St. Louis, MO: Author.

Pfannenstiel, J. C., Seitz, V., & Zigler, E. (2002). Promoting school readiness: The role of the parents as teachers program. *NHSA Dialog: A Research-to-Practice Journal for Early Childhood Field, 6*(1), 71–86. https://doi.org/10.1207/s19309325nhsa0601_6.

Pfannenstiel, J. C., & Seltzer, D. A. (1989). New parents as teachers: Evaluation of an early parent education program. *Early Childhood Research Quarterly, 4*, 1–18. https://doi.org/10.1016/S0885-2006(89)90025-2.

Posten, H. O. (1984). Robustness of the two-sample t-test. In D. Rasch, & M. L. Tiku (Eds.). *Robustness of statistical methods and nonparametric statistics* (pp. 92–99). Dordrecht, Germany: Reidel.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41–55. https://doi.org/10.1093/biomet/70.1.41.

Senechal, M., & Young, L. (2008). The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review. *Review of Educational Research, 78*(4), 880–907. https://doi.org/10.3102/0034654308320319.

Suarez-Orozco, C., Suarez-Orozco, M., & Todorova, I. (2008). *Learning a new land: Immigrant students in American society.* Cambridge, MA: Harvard University Press.

Sweet, M. A., & Appelbaum, M. I. (2004). Is home visiting an effective strategy? A meta-analytic review of home visiting programs for families with young children. *Child Development, 75*(5), 1435–1456. https://doi.org/10.1111/j.1467-8624.2004.00750.x.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse (2017). WWC Procedures and Standards Handbook, v4. Available at: https://ies.ed.gov/ncee/wwc/Handbooks.

Ziegler, E., Pfannenstiel, J. C., & Seitz, V. (2008). The parents as Teachers program and school success: A replication and extension. *Journal of Primary Prevention, 29*, 103–120. https://doi.org/10.1007/s10935-008-0132-1.