mcg michael cohen group LLC
research, evaluation, consulting

Effects of Interaction with *Pocoyo* Playsets on Preschool (Spanish) ELL Children's
English Language Learning: A Randomized Controlled Trial

*Prepared by:*

mcg

Evaluation Team, **Early Learning Collaborative**

*Prepared for:*

Early Learning Collaborative and

U.S. Department of Education Office of Innovation & Improvement

July 26th, 2013

Report #3-6 ELD

Research undertaken as part of the evaluation activities of the

**U.S. Department of Education Ready To Learn Program**

**Background**

The Hispanic Information and Telecommunications Network (HITN) was awarded one of the 2010-2015 Ready to Learn (RTL) grants. Under RTL, HITN has assembled the Early Learning Collaborative (ELC). ELC is producing a transmedia suite of products for three- to five-year-old children that supports: 1) English language development (ELD) for children whose primary language is Spanish; 2) early literacy; and 3) early math skills in preschool-age children.

As the independent evaluator for this RTL initiative, MCG conducts three kinds of research: formative research to inform the development of the products or properties, secondary research to support the development process, and summative evaluation to determine their educational effectiveness.

One of the key components of the transmedia sets are the *Pocoyo PlaySets*, digital activities and games for touch screen devices featuring Pocoyo, the popular children's media character. The first two PlaySets, *At the Seashore* and *Things that Go*, target ELD. MCG has conducted a summative evaluation of these first two PlaySets that is reported on here.

This report describes results of a randomized controlled trial (RCT) assessing the effects of use of the *Pocoyo PlaySets* by preschool Spanish-speaking ELL children on their learning of English. This study was conducted as the summative evaluation for ELC, HITN and Ready to Learn. We compared gains in English language fluency, learning of target English words, and attitudes toward English language learning among ELL preschool children assigned to use either the *Pocoyo PlaySets* or comparator materials in the form of two commercially available apps designed to teach English to ELL children (*Bilingual Child* and *LinguPenguin*) in a three-week placement period.

**Potential Value of Using Media as Part of Support Offered to ELL Children**

English Language Learners (ELL) are fast becoming the largest "minority" demographic group in U.S. schools. Students whose home language is other than English comprise more than 20% of the nation's young children (Miller and Garcia, 2008) and are projected by the U.S. Census Bureau to be 40 percent of the school-age population by the 2030s (U.S. Census Bureau, 2011), possibly sooner if present demographic trends continue.  This dramatic increase has spurred educators and policymakers to revisit research and evidence-based practices that support long term student achievement and success for this specific population of children. Recent national student achievement data highlights the need to identify instructional models and teaching strategies that will benefit young ELLs and result in improved academic performance at every educational level (Magruder, 2012).

Multimedia language-learning properties and different technology platforms have been investigated as tools to support ELL Learners in a number of interventions and studies (Istifci, 2011; White & Gillard (2011); Verhallen, Bus, & deJong (2006); Nomass, (2013); Cooper, (2005) and Godwin-Jones (2011).  However, these studies are all focused on interventions designed for older elementary and college-age ELL students.  They highlight the benefit of mobility (playable at home, at school or elsewhere) as well as engagement. A number of these studies, varying in design and sample size, do report positive findings or increases in English language learning.

There are a few relatively small studies of ELL preschool and kindergarten age children in interventions that assess the effects of technological mediation on vocabulary learning. In one study (Leacox & Wood Jackson, 2012), 4- to 6-year-old ELL students learned more vocabulary from an e-book with Spanish-bridging vocabulary (relating new English words to

Spanish vocabulary and definitions) than from traditional group reading. In another study, preschool to second-grade ELL children learned more vocabulary when traditional instruction was accompanied by multimedia with vocabulary-relevant content (Silverman & Hines, 2009). Silverman and Hines speculate that these gains are due to ELL children using meaningful visual content to support vocabulary learning. Interestingly, these benefits were only found for ELL children, suggesting that this kind of technological mediation of the learning process may be uniquely beneficial for that group.

Touch screen devices provide a form of technological mediation of educational activity that is especially appropriate for young children. Researchers have shown that preschoolers can select and move objects more easily and accurately on a touch screen compared to using a mouse or keyboard (Battenberg and Merbler, 1989; Lu and Frye, 1992). These advantages are relevant because of the increasing availability of touch screen devices in the classroom as well as in children's homes, and the rapidly growing supply of apps for preschool learning. According to Brian (2012), over 1.5 million iPads are used in educational programs, with access to more than 20,000 educational apps. A study from the same year reported that 80% of the top selling educational apps were targeted at children, and almost three-quarters (72%) of these target preschool or elementary school children (Shuler, 2012).

Despite the abundance of available educational apps, little research exists that directly assesses the educational outcomes of using touch screen apps. As a medium, the touch screen affords diverse possibilities for user interaction. The educational potential of the touch screen medium is therefore largely a function of the software (apps) or content that the medium hosts. Despite this certain affordances of the touch screen interface may be particularly

important for educational purposes. Software that allows a user to interact with the device via meaningful gestures is an important example.

Gestures play an important role in conceptual learning by providing an additional source of information that aids in understanding linguistic, mathematic and other concepts. Goldin-Meadow, Cook and Mitchell (2009) report that children picked up information about novel words that was implicit in gestures that accompanied word use. Another study showed that children who were instructed to gesture when recalling an event were able to remember more information about the event (Stevanoni & Salmon, 2005). Children who were told to use gestures when talking about solving a math problem remembered more about solving the problem, and had an enhanced ability to learn more when taught in the future (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007).

This research has important implications for the use of touch screen devices in educational settings. Touch screens afford gesturing (unlike a keyboard) and do so in a way that offers more versatility and direct expressiveness than a mouse. Importantly, a touch screen device can also provide feedback on a user's gestures, which may enhance the educational benefits described above. Yet, without research, the educational effects of these features, and their potential to enhance curriculum-based learning, remain unclear.

In addition, there is a growing body of evidence from RCT studies that well designed educational media for young children can facilitate early literacy in a child's native language. An example of this was the evaluation of an educational television show—*World World*[1]—carried

---

[1] The *Word World* summative study was submitted to the U.S. Department of Education as part of the Ready to Learn initiative, and reviewed and approved by a representative of the Institute of Education Sciences, as well as an expert panel convened by the Department of Education's Office of Innovation and Improvement.

out by MCG for the 2005-2010 RTL initiative and reviewed by the U.S. Department of Education's Institute of Education Sciences (IES). MCG used a RCT study to measure gains in the literacy skills of kindergartners who watched the program, compared to a control group. The study found that children who watched the program showed significant gains in learning and recognition of words featured in the show (Michael Cohen Group, 2008).

Previous RCT studies demonstrate the viability of using research to measure the benefits of educational media. Yet, only a small number of these studies have been done. Consequently, very little conclusive evidence exists about the effectiveness of different types of media or different platforms. The current study adds to the small body of randomized controlled trials testing the effects of curriculum-based interactive technology, games and apps on early learning. It is, to the best of our knowledge, the first study to be conducted with interactive educational media designed to support English language learning among young children whose native language is not English.

**Objectives**

The overall objective of this RCT study is to evaluate the educational effectiveness of the *Pocoyo PlaySets* in supporting English language learning among primarily Spanish-speaking preschool children. The specific objectives of this study are to investigate the effects of being exposed to and playing with two *Pocoyo PlaySets* on: (1) preschool-age ELL children's increases in English language fluency as measured by a standardized assessment of English language fluency (Pre-IPT), (2) participating children's learning of target vocabulary through a customized picture recognition task, and, (3) changes in children's attitudes towards learning English.

The analysis of the findings from this study also included an examination of whether English language learning with *Pocoyo PlaySets* or other ELL learning apps is moderated by children's initial level of English fluency, familiarity with the *Pocoyo* character and other demographic variables (e.g., age or gender of child, city, language that parents speak to child, parent education, etc.).

This RCT study tests a primary hypothesis that exposure to and interaction with *Pocoyo PlaySets* in a non-mediated intervention is efficacious in promoting English language learning. More specifically, we tested whether preschool ELL children who played with the *PlaySets* showed greater gains in learning of vocabulary than children in a comparator condition on program-specific measures, as well as more generalized learning of English as assessed by a standardized test of English language fluency.

## Methodology

**Design**

The effects of playing with *Pocoyo PlaySets* on learning of English vocabulary concepts was tested in a cluster-randomized trial, with randomization stratified by school and age of children in classrooms. Within each participating child care center, classrooms serving 3-, 4- or 5-year-old children were randomly assigned to use either *Pocoyo PlaySets* or comparator app properties, *Bilingual Child* and *LinguPinguin*, in addition to receiving their regular classroom instruction.

This study was designed to meet the criteria for "strong evidence" outlined by the U.S. Department of Education including:

a) Participants centrally randomly assigned to different treatment and control conditions;

b) Standardized and documented intervention procedures to enable replication;

c) Baseline data used to confirm equivalence across conditions prior to the intervention and to statistically control for any differences not removed by randomization;

d) Use of validated assessment measures when possible (assessment of the primary outcome in the present study required a new measure that specifically matched curriculum content);

e) Sufficiently large sample sizes will provide adequate power to detect treatment effects;

f) Research conducted in multiple, geographically diverse sites to maximize external validity;

g) Assertive follow-up procedures to limit attrition to less than 25%;

h) If there is substantial attrition, analyses are conducted on an intent-to-treat basis, using multiple imputation methods to estimate outcomes for subjects lost to follow-up;

i) Reporting of effect sizes and statistical significance;

j) Reporting of positive and negative findings; and

k) Examining process, setting and child factors that moderate or mediate treatment effects to further address external validity.

In addition, this study was designed to meet RTL Government Performance and Results Act (GPRA) measures for summative research studies, specifically the GPRA guideline to "demonstrate positive and statistically significant gains in math or literacy skills when RTL transmedia properties are compared to similar non-RTL-funded digital properties or to other more traditional educational materials." The design of the study reported on here includes a comparison of two RTL digital properties designed to support English language learning --

*Pocoyo PlaySets* -- with two comparator apps that are non-RTL-funded, commercially available digital properties also designed to support English language learning.

Analyses of child and setting factors that moderated treatment effects are included in this report. Analyses of process measures (specifically analytics data on app use) as potential mediators of intervention effects will be presented in a subsequent report.

This multi-site trial was conducted in four geographically diverse locations: New York, NY; Bridgeport, CT; Miami, FL; and Los Angeles, CA . A total of 608 ELL Spanish-speaking, preschool-age children were identified and recruited in 15 early childhood centers or schools that serve a high percentage of ELL children. Fifty-three (53) classrooms of three-, four- or five-year-old children in these centers were included as part of the sample for this study. Classrooms were included if they served a minimum of five students identified by the school as Spanish-speaking ELL children. Within each school, roughly half the classes were randomly assigned to use tablets containing *Pocoyo PlaySets*, and half were randomly assigned to use tablets containing two comparator products, *LinguPinguin* and *Bilingual Child* .

All 53 participating classrooms were provided with touch screen tablets (1 tablet for every 4 children) on which either *Pocoyo PlaySets* (Experimental) or comparator apps (*LinguPinguin* and *Bilingual Child)* had been downloaded, depending on random assignment of the individual class to the Experimental (26 classrooms) or Comparator (27 classrooms) groups. All assessments, permission slips and other study materials were made available in both English and Spanish.

**Sample**

This multi-site trial included total of 608 preschool age children attending 15 preschools or day care centers in four geographically diverse locations: Miami, FL; Los Angeles, CA; New York, NY, and Bridgeport, CT. Twenty-six classes (295 children) were randomly assigned to the *Pocoyo* Experimental condition and 27 classes (313 children) were randomly assigned to the *LinguPinguin / Bilingual Child* comparator condition (see Table 1).

**Subject retention.** All of the classes participated in the pre-test and post-test. Of 608 children recruited, 580 (96%) were retained at post-test. This extremely high retention rate far exceeded the 75% retention required for a "strong evidence" effectiveness study. Although there was a non-significant trend toward higher retention of subjects in the *Pocoyo* condition (n=286, 97%) than in the Comparator conditions (n = 294, 94%; chi square (1 df) = 3.2, p < .08), the extremely high retention rate means that there is little potential for biasing of results due to subject dropout. This makes imputation of missing data unnecessary, and enables us to base our on analyses on observed data from children who completed both assessments.

Table 1

*Classes Recruited*

|  | Pocoyo | LinguPinguin/ Bilingual Child | Total |
|---|---|---|---|
| Total | 26 | 27 | 53 |
| Miami, FL |  |  |  |
|    Site 1 | 3 | 2 | 5 |
|    Site 2 | 2 | 3 | 5 |
|    Site 3 | 1 | 1 | 2 |
|    Site 4 | 2 | 1 | 3 |
|    Site 5 | 2 | 2 | 4 |
|    Site 6 | 2 | 3 | 5 |
| Los Angeles, CA |  |  |  |
|    Site 12 | 2 | 2 | 4 |
|    Site 13 | 0 | 2 | 2 |
|    Site 14 | 2 | 1 | 3 |
|    Site 15 | 2 | 2 | 4 |
| New York, NY |  |  |  |
|    Site 7 | 2 | 2 | 4 |
| Bridgeport, CT |  |  |  |
|    Site 8 | 2 | 2 | 4 |
|    Site 9 | 2 | 1 | 3 |
|    Site 10 | 1 | 1 | 2 |
|    Site 11 | 1 | 2 | 3 |

**Participant characteristics.** Demographic and linguistic or English language fluency characteristics of participants are shown in Table 2. Twenty-two percent of participants were three-year-olds, 19% were early four-year-olds, 29% were late four-year-olds, and 30% were five-year olds. Participants were fairly evenly split between males (48%) and females (52%). The majority (56%) of parents reported that their annual household income was under $20,000, and most of the remainder (27%) had incomes between $20,000 and $39,999. Only 17% of parents reported incomes of $40,000 or more.

Eight-four percent of parents reported that Spanish was the primary language in the home, and another 10% reported speaking primarily Spanish and English (see Table 2). Only 6% of participating parents reported English was the primary language in the home.  Eight-eight percent of the children first spoke in Spanish. Ninety percent of the children were U.S. born; 10% had immigrated.  A plurality (45%) of children scored at the level of "Limited English Speaking," level (B) on the baseline Pre-IPT assessment of English language fluency. Fifteen percent of children scored at the "Non-English Speaking,"  (A) level. Only 18% scored in the Fluent English Speaking (E) range on the Pre-IPT.

Baseline testing showed that 77% of children had used a touch screen before. Nearly two-thirds of the children (62%) were familiar with *Pocoyo PlaySets*.

Table 2

*Participant Demographics*

|  | Total (580) | Pocoyo (280) | LinguPinguin/ Bilingual Child (294) | Chi-square | *p* |
|---|---|---|---|---|---|
| Age |  |  |  | 20.4 (3 df) | .001 |
|   Three (36-47 months) | 22% | 29% | 16% |  |  |
|   Early fours (48-53 months) | 19% | 21% | 16% |  |  |
|   Late fours (54-59 months) | 29% | 26% | 32% |  |  |
|   Five (60-71 months) | 30% | 24% | 36% |  |  |
| Gender |  |  |  | 1.0 (1 df) | .33 |
|   Male | 48% | 50% | 46% |  |  |
|   Female | 52% | 50% | 54% |  |  |
| City |  |  |  | 0.6 (3 df) | .89 |
|   Miami | 53% | 51% | 54% |  |  |
|   Los Angeles | 26% | 27% | 26% |  |  |
|   New York | 10% | 11% | 10% |  |  |
|   Bridgeport | 11% | 11% | 10% |  |  |
| Pre-IPT English Fluency (Pre-test) |  |  |  | 12.2 (4 df) | .02 |
|   A (Non-English Speaking) | 15% | 16% | 14% |  |  |
|   B (Learning English Speaking) | 45% | 51% | 40% |  |  |
|   C (Learning English Speaking) | 8% | 6% | 10% |  |  |
|   D (Learning English Speaking) | 13% | 10% | 16% |  |  |
|   E (Fluent English Speaking) | 18% | 17% | 20% |  |  |
| Primary Language Spoke at Home |  |  |  | 0.2 (2 df) | .90 |
|   Spanish | 84% | 85% | 84% |  |  |
|   Spanish and English | 10% | 9% | 10% |  |  |
|   English | 6% | 6% | 6% |  |  |
| Language Caregiver Speaks to Child |  |  |  | 0.3 (2 df) | .85 |
|   Spanish | 73% | 72% | 74% |  |  |
|   Spanish and English | 18% | 19% | 17% |  |  |
|   English | 9% | 9% | 9% |  |  |
| Language Child First Spoke |  |  |  | 2.2 (2 df) | .33 |
|   Spanish | 88% | 87% | 89% |  |  |
|   Spanish and English | 4% | 5% | 2% |  |  |
|   English | 8% | 8% | 9% |  |  |
| Child Length of Time in U.S. |  |  |  | 0.9 (2 df) | .63 |
|   Born in U.S. | 90% | 91% | 89% |  |  |
|   More than one year | 5% | 4% | 5% |  |  |
|   Less than one year | 5% | 5% | 6% |  |  |

Table 2

*Participant Demographics (continued)*

| | Total (580) | Pocoyo (280) | LinguPinguin/ Bilingual Child (294) | Chi-square | p |
|---|---|---|---|---|---|
| Parent Education | | | | 3.8[a] (1 df) | .051 |
| None or primary school | 1% | < 1% | 1% | | |
| Some high school | 27% | 32% | 23% | | |
| High school graduate | 34% | 33% | 34% | | |
| Some college | 20% | 19% | 21% | | |
| College graduate | 14% | 12% | 16% | | |
| Some graduate work | 4% | 4% | 5% | | |
| Annual Household Income | | | | 1.6[a] (1 df) | .20 |
| Under $20,000 | 56% | 60% | 52% | | |
| $20,000 - $39,999 | 27% | 24% | 30% | | |
| $40,000 - $49,999 | 10% | 9% | 10% | | |
| $50,000 - $69,999 | 4% | 5% | 4% | | |
| $70,000 - $99,999 | 2% | 1% | 3% | | |
| $100,000 or more | 1% | 1% | 1% | | |
| Familiar with Touch Screen | | | | 0.1 | .76 |
| Yes | 77% | 76% | 77% | | |
| No | 23% | 24% | 23% | | |
| Familiar with Pocoyo | | | | 0.6 | .44 |
| Yes | 62% | 60% | 63% | | |
| No | 38% | 40% | 37% | | |

[a] Chi-square for linear association

**Intervention**

Touch screen tablets were placed in participating classrooms (ratio of 1 tablet for every 4 children) with the stimuli to which particular classes had been randomly assigned downloaded on the tablets. Teachers were given both written and verbal descriptions of the apps / condition to which their class had been randomly assigned (*Pocoyo PlaySets* or *Bilingual Child* and *LinguPinguin*). Teachers were asked to have the tablets with their respective apps available as an activity for children three times a week for three weeks during free play, when

several alternative activities were offered, assuring that participating ELL children had access at these times. This amount of exposure was expected to be sufficient to produce changes on proximal outcomes (learning of vocabulary directly presented in the apps). However, it was not expected to produce changes on more distal outcomes, such as overall English fluency, as assessed by a standardized test.

The intention was to enable children to access the apps as they might at home or in school where tablets with apps were available, but not used as part of a formal intervention that included formal instruction or additional materials. In other words, the goal was to provide exposure to digital learning tools in an informal intervention. The teachers were instructed on the use of the tablets and on the specific apps to which their class had been assigned so that they could answer simple questions and understand what the children in their class were doing.

**Stimuli**

The selection of commercially available comparator apps for the Control or comparator condition was challenging. Some apps are available for English language speakers to learn basic Spanish. However, very few apps have been developed to support English language learning by Spanish-speaking preschoolers.  We also wanted an app or apps that included vocabulary (colors, vehicles, animals) that were similar, if not identical, to the content words presented in the *Pocoyo PlaySets*. The two comparator apps chosen – *Bilingual Child* and *LinguPinguin* – met the criteria of preschool usability and together covered much of the target vocabulary.  The activities on these apps were less varied and interactive than those on the *Pocoyo PlaySets*, but they were the best that could be found in the commercial marketplace. A full description of

both the *Pocoyo PlaySets* and the two comparator apps, including visuals and specific content can be found in Appendix B.

**Procedures**

Preschool childcare centers were recruited by calling directors of early childhood learning centers in neighborhoods where there was a high percentage of Spanish speaking residents (based on local census information). A description of the purpose and design of the study as well as its funding were presented in an introductory conversation. The Center Directors who agreed to participate were asked to sign a form that was then submitted as part of the IRB for the study. All participating centers were incentivized for their time and effort.

Participating childcare centers were visited by researchers prior to the beginning of the study in order to have an introductory meeting in which the study, the pre- and post-assessment process and the placement were fully presented. Written materials describing the study details were left for Directors and participating teachers. Specific instructions for the Teachers were made available along with a brief Study Protocol.

A Manual for Research Assistants in the field was used as the basis for training sessions for researchers at all sites. Research Assistants were trained in and had an opportunity to practice administering the assessments used in the pre- and post-testing. Teams of researchers made arrangements to begin assessment at each center or school. All participating children in both Experimental and Comparator classrooms were assessed individually in as private a space as possible. Once the pre-testing was completed the teachers were given additional hands-on instruction in the use of the tablets and apps. The placement period began at a given center or school when the pre-testing was completed.

In order to assure compliance, structured observations were conducted biweekly by researchers in all participating classrooms. This provided an opportunity for teachers to ask questions and for researchers to assure that the apps were being used as instructed. These structured observations also provided information on the details of children's use of the apps, their engagement, sources of frustration, or changes in play patterns over the course of the placement.

Teachers were asked to complete a teacher log on a biweekly basis during the three-week placement. They were incentivized for completing the log, which was primarily composed of closed-end ratings of usability, appeal, engagement, interactivity, and children's spontaneous repetition of words, songs or activities from the apps during class or free play, as well as providing space for teachers to write open-ended comments. (See Appendix A for Structured Classroom Observation Sheet and Teacher Log).

A brief overview of findings from these teacher logs is provided in Appendix C, Figure 7. These are supplementary to the RCT findings and provide some insight into the process behind the outcomes.

**Measures**

There were three core measures administered at pre- and post-test sessions, as well as additional measures of familiarity with the *Pocoyo* property, and familiarity with touch screen devices administered only at pre-test. These measures were piloted several months before fielding in a study with 30 children. All measures were available in both English and Spanish.

1. A standardized pre-school assessment of English language fluency, the Pre-IPT Oral English Test (Ballard & Tighe, 2010).

16

2. A customized picture recognition task of core vocabulary or words (those presented a minimum of three times in the *Pocoyo PlaySets*), with many target words related to colors and vehicles also represented in the comparator apps.

3. A brief questionnaire on Attitudes toward English Language Learning that was designed with 4-item likert scale (modeled on the ERAS, (McKenna & Kear, 1990)) with smiley faces indicating extent of happiness or unhappiness about learning English (in English and Spanish): happy (feliz), a little happy (un poquito feliz), OK (bien), or sad (triste). Inter-item reliability on this measure was low (alpha = .39), so instead of computing a summary score, we used responses to only the first item: "How do you feel about learning English? (*¿Cómo te sientes aprendiendo inglés?)".*

4. A simple familiarity with the *Pocoyo* property assessment was administered at pre-test in order to determine the impact of familiarity with the property and its characters. Pictures of the *Pocoyo* characters and screen shots from several other children's shows were presented and children were asked with which ones they were familiar.

5. A brief familiarity with touch screen devices assessment was administered at pre-test.

**Analysis Plan**

Outcomes analyses are based on the 580 students who completed both the pre-test and post-test. We had been prepared to use multiple imputations to estimate post-test values for subjects who did not complete the study.  However, because 95% of subjects were retained, there was negligible potential for bias due to drop out. This made multiple imputation of missing data was unnecessary and we simply used the observed data for children who completed both assessments.

**Equivalence of conditions at pre-test.** Although classes were randomly assigned to

conditions, preliminary analyses indicated there were some differences by condition at pre-test

(see Table 2).  Children in the *Pocoyo* condition tended to be younger (chi-square (1 df) = 20.4,

p < .001), and to score lower on the baseline Pre-IPT test of verbal English skills (chi-square (1

df) = 12.2, p < .02). Parental education was also lower in the *Pocoyo* condition than in the

comparator condition (linear association chi-square (1 df) = 3.8, p < .051) Subsequent analyses

for the effect of intervention condition on outcomes controlled for the effects of age, Pre-IPT

score, and (in most analyses) parent education.

**Main effect of intervention condition on outcomes.** All outcome analyses were

conducted using the SPSS 19.0 mixed models procedure. Mixed modeling can account for

observations being clustered within groups, and is appropriate for studies where entire classes

rather than individuals are randomly assigned to treatment conditions. Mixed models can

include predictors that we want to model explicitly (fixed effects) as well as variables that we

want to control for without developing specific models of their effects (random effects).

We ran separate models to test the effect of the intervention condition on five post-test

outcomes:  target vocabulary taught in *Pocoyo* (primary outcome); PlaySet 1 and PlaySet 2

Vocabulary (subscales of the primary outcome); Pre-IPT score; and attitude toward learning

English.  Our models assessing the effects of Pocoyo vs. comparator products on outcomes

included the following predictors as fixed effects: intervention condition (Pocoyo vs.

comparator), pre-test score on the outcome measure, gender, age, Pre-IPT score, and parent

education. Class and city were included as random effects. Class was included as a random

effect to control for the extent to which children within the same class tend to have similar

18

scores. Because one city (New York) had only two classes per condition, it was hard to develop

reliable estimates of the effect of each city: the effect of city was therefore controlled as a

random rather than fixed effect. For our main effect analyses, if the main effect of **condition**

was significant, this would indicate that post-test scores differed between the *Pocoyo* and

*LinguPinguin/Bilingual Child* condition after controlling for the effects of pre-test scores and

other covariates.

      **Moderator analyses.** We ran additional models testing whether various factors

moderated the effects of the *Pocoyo* apps on our primary outcome. Potential moderators we

examined included city, child gender, child age, child's English fluency (Pre-IPT score) at pre-

test, child's pre-test Attitude toward Learning English, child's prior use of touch screens, child's

familiarity with *Pocoyo,* whether the child was born in the U.S., whether English was one of the

child's first languages, what language the child's primary caregiver spoke to the child, parent

education, and parent income. If the effect of an interaction term (e.g., intervention condition

x language spoken by primary caregiver) on post-test vocabulary was significant, this would

indicate that the latter variable moderated the effects of *Pocoyo  PlaySets* on outcomes. This

would mean that the effects of *Pocoyo* relative to the comparator apps were stronger among

some subgroups of children (e.g., children whose primary caregiver spoke only Spanish) than

among others (children whose primary caregiver spoke English or both Spanish and English).

### Results

**Effects of Intervention on Learning of Target English Vocabulary**

      **Children in classes randomly assigned to use the *Pocoyo PlaySets* showed significantly**

**greater improvements in vocabulary than did children in classes randomly assigned to use the**

*LinguPinguin* and *Bilingual Child* apps.  Mixed model results are shown in Table 4.  Controlling

for the effects of vocabulary score at pre-test, age, Pre-IPT score, and parent education, the

(non-significant) effect of gender, and with city and class as random effects, children in the

*Pocoyo* condition knew an average of 1.7 more words at post-test than did children in the

comparator condition (F (1, 46 df) = 18.3, p < .001).  If we look at subscale scores for vocabulary

taught in each PlaySet, we see that children in the *Pocoyo* condition showed significantly

greater gains than children in the comparator condition for both vocabulary taught in *PlaySet 1*

(F  (1, 45 df) = 12.6, p < .001) and vocabulary taught in *PlaySet 2* (F  (1, 45 df) = 18.9, p < .001).

Controlling for covariates, children in the *Pocoyo* condition knew an average of 0.7 more words

from *PlaySet 1* and an average of 1.1 more words from *PlaySet 2* at post-test than did children

in the comparator condition (see Table 3).

Table 3

*Main Effect of Intervention Condition on Vocabulary Taught in Pocoyo*

| | Total Vocabulary Score | | PlaySet 1 Vocabulary Score | | PlaySet 2 Vocabulary Score | |
|---|---|---|---|---|---|---|
| | b-weight | *p* | b-weight | *p* | b-weight | *p* |
| Pocoyo (reference = comparator condition) | 1.71 | .001 | 0.66 | .001 | 1.12 | .001 |
| Pre-test score on outcome | 0.78 | .001 | 0.65 | .001 | 0.73 | .001 |
| Age (months) | 0.08 | .004 | 0.03 | .01 | 0.05 | .002 |
| Pre-test Pre-IPT category | 0.66 | .001 | 0.30 | .001 | 0.58 | .001 |
| Male  (reference = female) | 0.42 | .25 | 0.17 | .34 | 0.10 | .66 |
| Parent education | 0.31 | .05 | [a] | [a] | [a] | [a] |

[a] Not included because mixed model would not converge with parent education in the model.

These results are shown graphically in Figure 1.  Total vocabulary total scores and subscale scores increased over time among children in both conditions. However these scores increased significantly more among children in the *Pocoyo PlaySets* condition than among comparators.
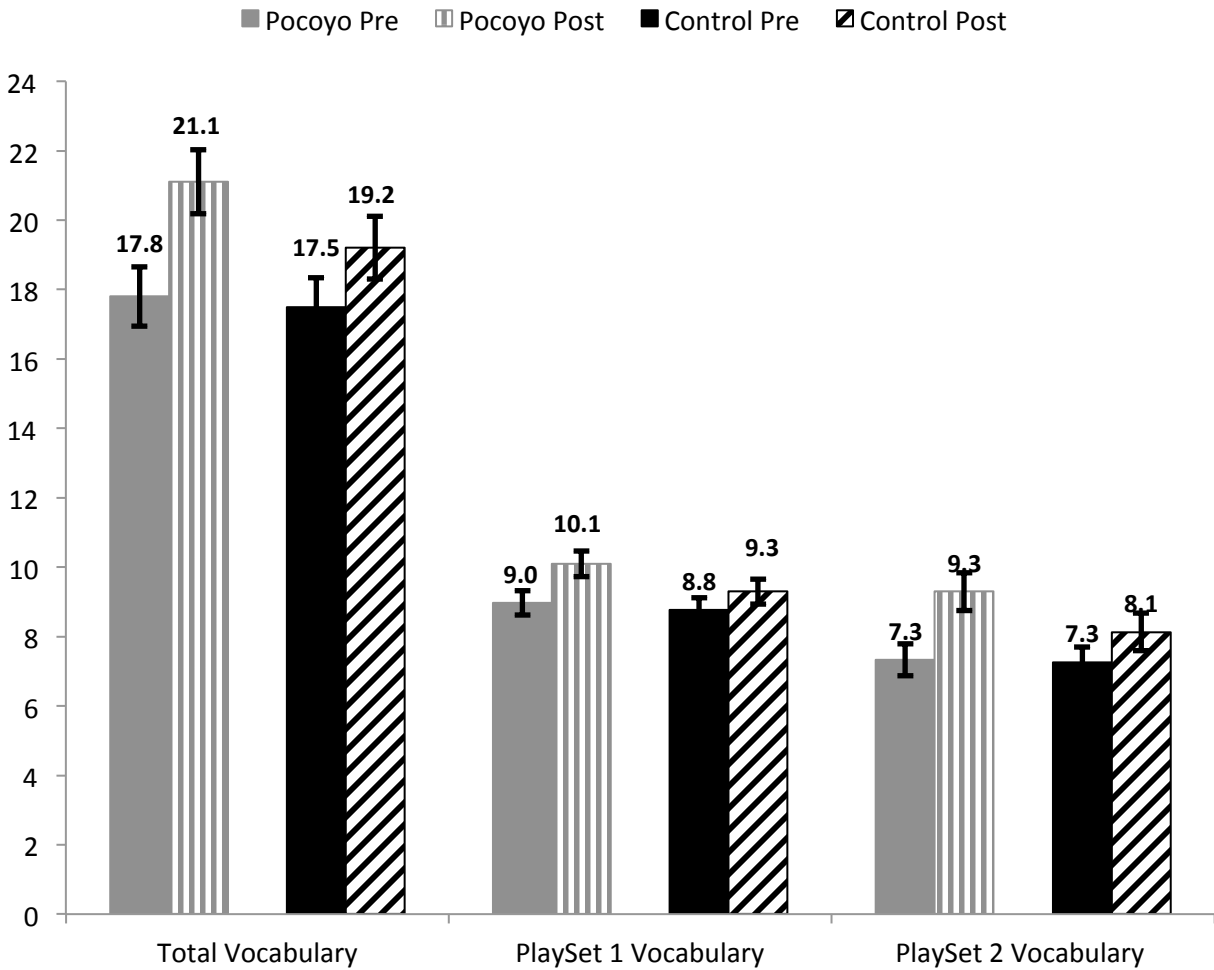


*Figure 1.* Pre-test and Post-test Vocabulary Scores by Intervention Condition.

These results were also expressed in terms of standardized effect sizes (Cohen's d = mean difference between conditions at post-test/standard deviation at post-test). The effect of *Pocoyo* intervention condition (relative to the *LinguPinguin/Bilingual Child* condition) on total vocabulary and the vocabulary subscales were d = .20 to .23, typically considered small-sized effects (see Figure 2).
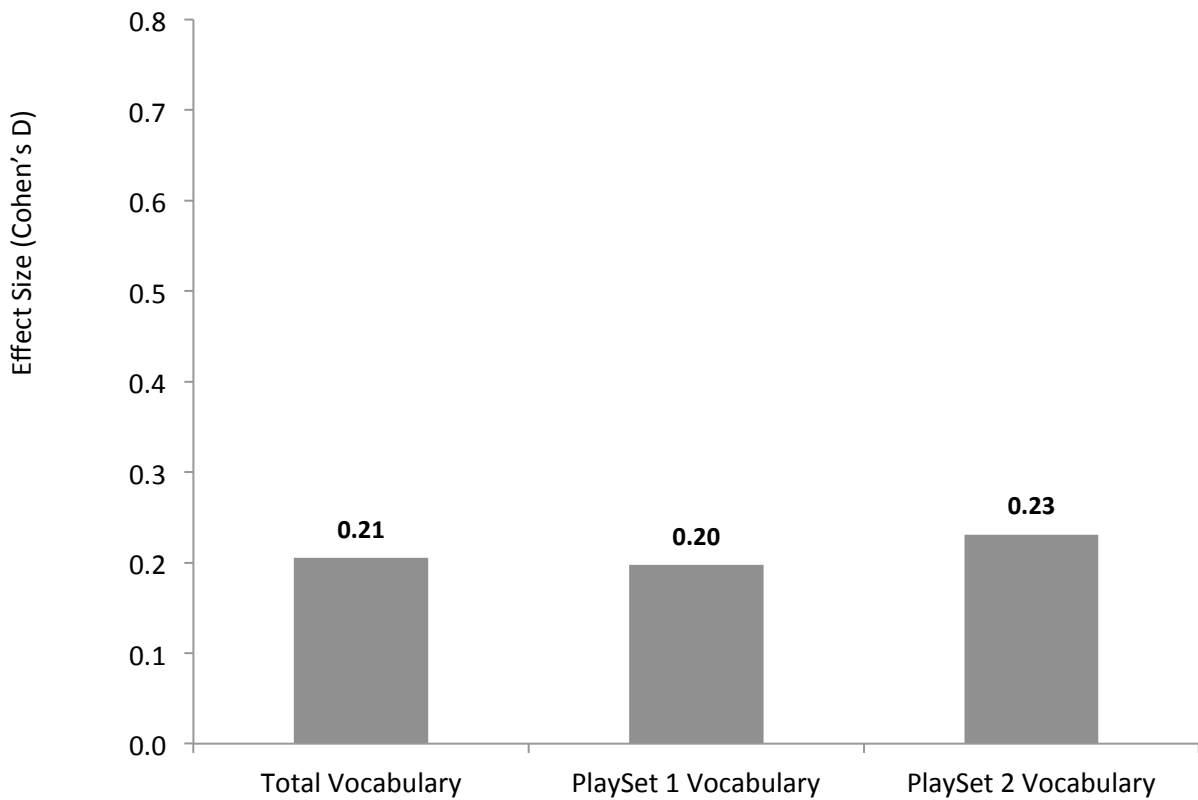


*Figure 2.* Effect Size for *Pocoyo PlaySets* vs. Comparators on Vocabulary at Post-test.

**Effects of Intervention on Secondary Outcomes**

There was no significant effect of *Pocoyo* vs. comparator apps condition on children's overall English fluency category in the Pre-IPT at post-test ($F$ (1, 53 df) = 0.1, $p < .76$; see Table 4). Intervention condition also did not have a significant effect on children's reported attitude towards learning English at post-test ($F$ (1, 43 df) = 0.0, $p < .86$; see Table 4).

Table 4

*Main Effect of Intervention Condition on Secondary Outcomes*

| | English Fluency (Pre-IPT Category) | | Attitude Towards Learning English | |
|---|---|---|---|---|
| | b-weight | *p* | b-weight | *p* |
| Pocoyo condition (reference = comparator) | -0.04 | .76 | -.02 | .86 |
| Pre-test IPT score | 0.60 | .001 | .02 | .58 |
| Pre-test attitude toward learning English | -- | -- | .10 | .03 |
| Age (months) | 0.02 | .002 | .01 | .06 |
| Male  (reference = female) | -0.07 | .37 | .04 | .62 |
| Parent education | 0.01 | .72 | -.03 | .46 |

**Moderation by City**

As noted above, our ability to compare intervention outcomes across cities was limited by the fact that one site, New York, had only two classes of children in each treatment condition. We assessed for moderation effects by city on vocabulary outcomes using a model that included pre-test vocabulary score, condition, city, age, and condition x city (we were unable to control for pre-IPT score or parent education as the model would not converge if we included either of these as covariates). We found a non-significant main effect for city ($F$ (3, 32 df) = 2.0, $p < .13$) but a significant condition x city interaction effect ($F$ 3, 30 df) = 3.1, $p < .04$),

indicating that the effects of *Pocoyo* varied by city. In three of the four cities, classes

randomized to the *Pocoyo* apps condition showed greater gains in vocabulary than did children

assigned to the comparator condition (see Figure 3).  However, in New York, children in the

comparator condition showed gains that were at least as large as those shown by children
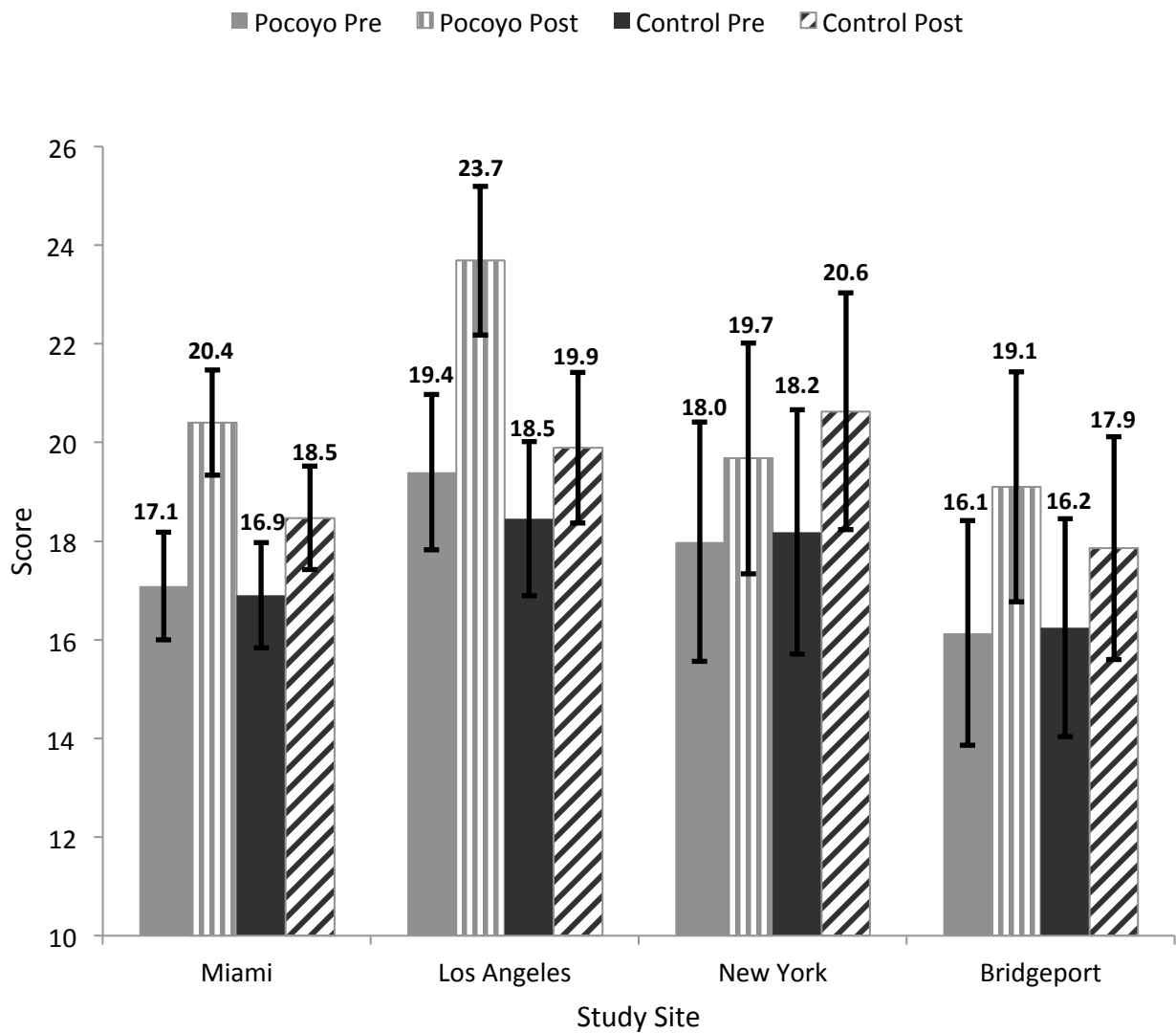
randomized to use the *Pocoyo* PlaySets.



*Figure 3.* Vocabulary at Pre-test and Post-test by Condition and Study Site.

**Moderation by Language Spoken by Primary Caregiver**

The effects of interacting with *Pocoyo PlaySets* on vocabulary were stronger among children whose primary caregiver spoke to them only in Spanish (n = 400) than among children whose caregiver spoke to them in both Spanish and English (n = 98) or in English (n = 48).[2] Controlling for the main effects of intervention condition, language spoken by primary caregiver, child age, gender, baseline pre-IPT score, and vocabulary score at pre-test, there was a significant interaction between condition and caregiver language in predicting children's vocabulary score at post-test (F 1, 525 df) = 3.9, p < .05).

Not surprisingly, comparator children whose caregiver did not speak to them in English showed the least improvement in vocabulary taught in *Pocoyo* (see Figure 4). Among children whose caregiver spoke Spanish to them, there was a substantial difference in the mean pre-post improvement in vocabulary among children in the *Pocoyo* condition (3.4 words) and comparator condition (1.2 words). This difference corresponds to an effect size of d = .25 (see Figure 5). If children's primary caregiver spoke at least some English to them, there was smaller difference (effect size d = .05) in mean improvement shown by children in the *Pocoyo* condition (3.0 points) and comparator condition (2.5 points; see Figures 4). Among children whose caregivers spoke some English to them, the effect size for *Pocoyo PlaySets* vs. comparator apps on vocabulary was small (d = .05) and not statistically significant (see Figure 5).

---

[2] Only 546 parents reported the language spoken by the primary caregiver.
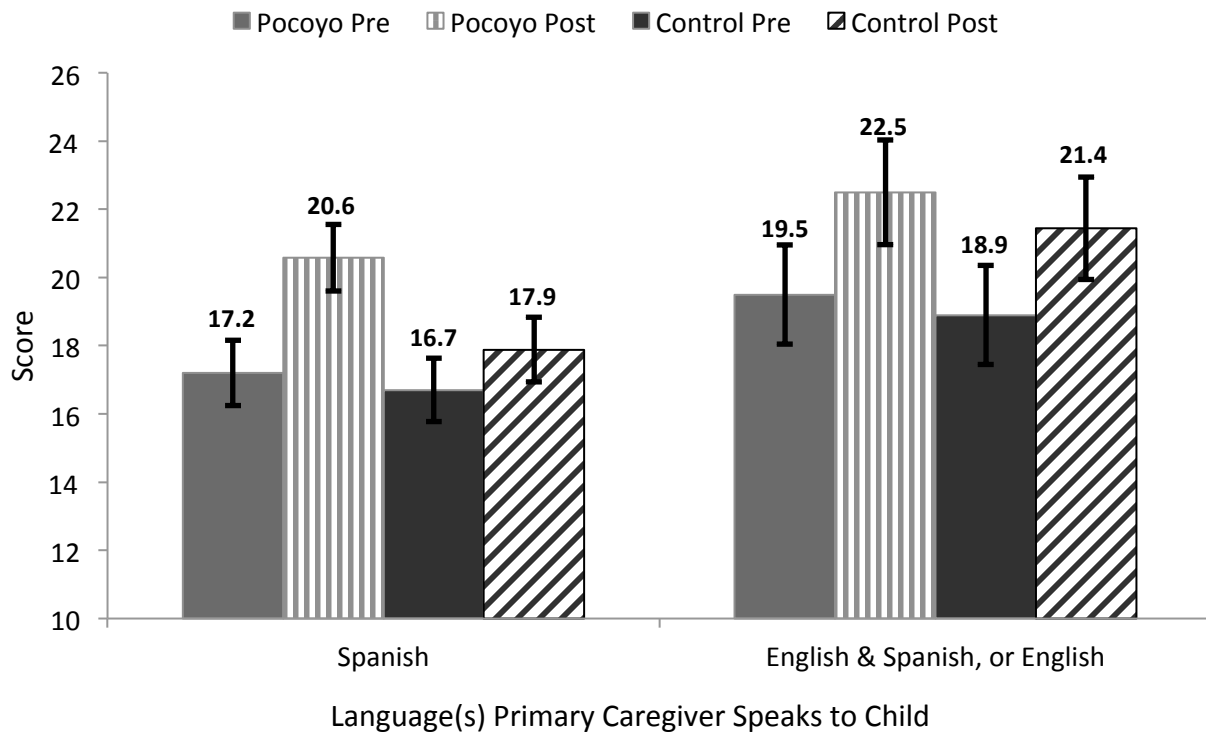
*Figure 4*. Vocabulary at Pre-test and Post-test by Condition and Language Spoken by Primary Caregiver.
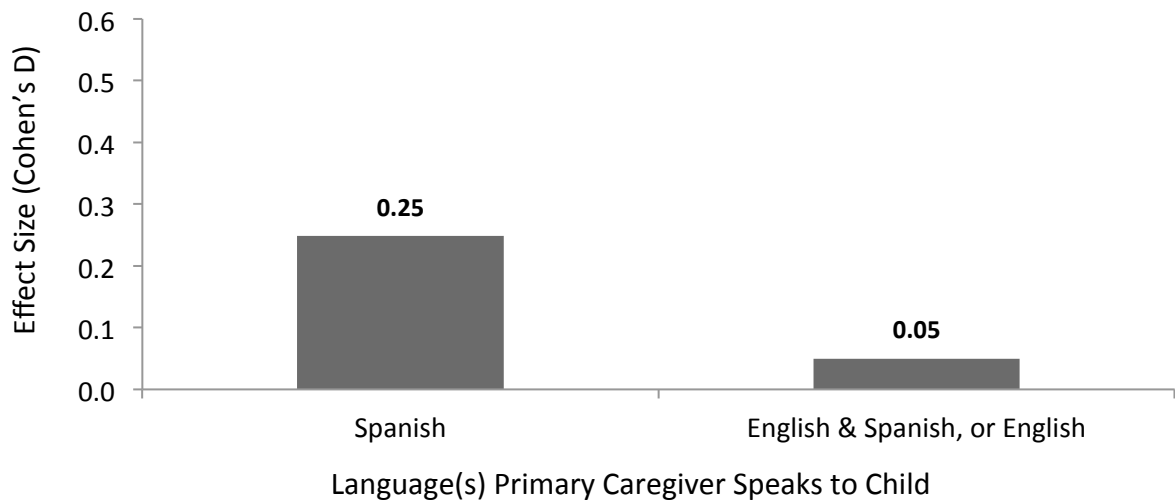


*Figure 5.* Intervention Effect Size on Vocabulary Post-test by Language Spoken by Primary Caregiver.

**Moderation by Age and Baseline Pre-IPT English Fluency**

The effects of interacting with *Pocoyo PlaySets* relative to the comparator apps on learning of vocabulary was moderated by age of child (3, younger 4, older 4, or 5) and English fluency at pre-test (Pre-IPT levels A-D vs. Pre-IPT level E). Because Pre-IPT score was somewhat confounded with age (they were correlated $r = .47$, $p < .001$), moderator effects only emerged if the model accounted for a complex three-way interaction of condition x age x pre-IPT score. In this full model, we found significant moderation of intervention effects by child age (condition x age interaction $F_{(3, 472\ df)} = 3.9$, $p < .01$), by Pre-IPT score (condition x Pre-IPT score interaction $F_{(3, 549\ df)} = 10.4$, $p < .001$), and condition x age x Pre-IPT score ($F_{(3, 548)} = 3.0$, $p < .03$).

As shown in Figure 6, children in the *Pocoyo c*ondition generally showed greater gains than children in the comparator condition. The only *exception* was among younger children (ages 36-53 months) who scored in the Fluent English Speaking range on the baseline Pre-IPT. Because Pre-IPT scores were correlated with age, the few young children ($n = 19$) who scored in the Fluent English Speaking range were probably particularly proficient in English, with little opportunity for *Pocoyo PlaySets* to improve their vocabulary.

Although vocabulary scores appear to actually decrease over time among 36- to 53-month-olds in the *Pocoyo PlaySet* condition who scored in the Fluent range at pre-test, this sub-sample sample size is so small that this change is not reliable and is not statistically significant.
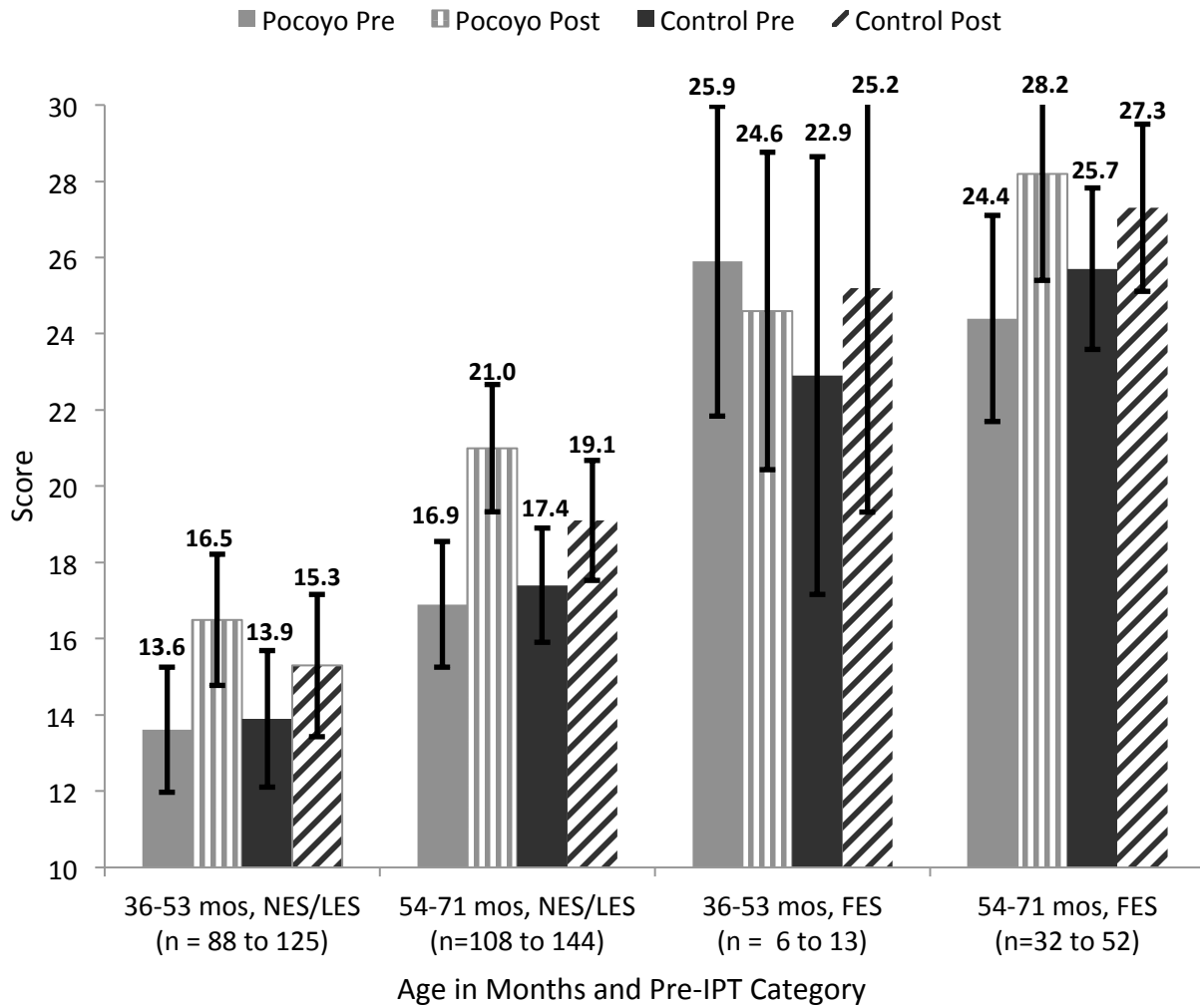
*Figure 6.* Vocabulary at Pre-test and Post-test by Condition, Age, and Pre-test Score on the Pre-IPT.

Note: NES/LES = Non-English Speaking (Pre-IPT = A) or Limited English Speaking (Pre-IPT = B to D). FES = Fluent English Speaking (Pre-IPT = E).

**Variables That Did Not Moderate Outcomes**

It is equally important to note several variables that did <u>not</u> moderate outcomes – these suggest that different subgroups of children can all benefit from using the *Pocoyo PlaySets*. The effects of interaction with *Pocoyo PlaySets* relative to *LinguPinguin* and *Bilingual Child* did not vary by gender: girls and boys were equally likely to benefit from the *Pocoyo PlaySets.* Intervention effects were not moderated by parent education or household income. Although the language the primary caregiver spoke was a significant moderator of the effect of *Pocoyo* on vocabulary acquisition, this outcome was not moderated by whether the children were born within or outside the U.S. (90% of subjects were reported to be US-born), or by whether the child first spoke in English. The effects of *Pocoyo* were also not moderated by children's pre-test responses to the Learning English Attitudes test. Finally, children's ability to benefit from interacting with the *Pocoyo PlaySets* was not moderated by whether they showed familiarity with using a touch screen at pre-test, or by whether they already were familiar with *Pocoyo.*

**Discussion**

This study provided evidence about the effects of preschool ELL children playing on an educational digital application on their learning of English as a second language. The findings presented here indicate that **preschool ELL children randomly assigned to use the *Pocoyo PlaySets* showed significantly greater gains in target vocabulary learned (main effect). Playing in an informal intervention for three weeks (three times per week) with *Pocoyo PlaySe*ts produced gains in learning for vocabulary content covered in both *Pocoyo PlaySet* 1 and in *Pocoyo PlaySet* 2, when compared to other commercially available apps designed for young ELL children.** There were similar effect sizes for subscales of both PlaySets. These gains in

learning are uniquely attributable to the use of the *Pocoyo PlaySets*, not to the children's usual experience or instruction in preschool (which was comparable in both conditions). The effect size for these findings is d = .21, which is on a par with the effects observed in randomized trials of several other educational media interventions for preschool age children, including the summative evaluations of the *Word World*[3] television series (d = .28) and the *Duck's Alphabet* online game (d = .21) (both reports were prepared by MCG for the Ready to Learn initiative and reviewed by IES).  This study provides strong evidence that free play with *Pocoyo PlaySets* results in significantly greater acquisition of target English language vocabulary when compared to learning after use of other commercially available English language learning apps.

While use of *Pocoyo PlaySets* or comparator apps did not produce gains in overall English language fluency assessed with the Pre-IPT, achieving change on a more distal, standardized outcome measure would likely require a longer exposure to more PlaySets with a wider range of vocabulary and content. However, these results do clearly indicate that it is feasible to use educational digital media as an informal intervention for young, preschool-age children. Even the 3-year-old children were, with some self-directed exploration and practice, able to use and learn from engagement with these properties. Further, Spanish-speaking preschool ELL children can independently interact with and learn from interactive digital media designed to support English language learning. This is noteworthy given that even young children, age 2 – 8, including children from lower income or Hispanic families, are increasingly likely to have access to touch screen devices, both at home and in school (Pew, 2012 and

---

[3] The *Word World* summative study was submitted to the U.S. Department of Education as part of the Ready to Learn initiative, and reviewed and approved by a representative of the Institute of Education Sciences, as well as an expert panel convened by the Department of Education's Office of Innovation and Improvement.

Michael Cohen Group, 2013). These devices are accessible for children in ways that traditional computers are not. They are portable and can be held flexibly in various positions, and the touch interface de-emphasizes the role of the keyboard and written commands, providing a more accessible entry point for pre-literate children. It should also be noted that previous experience with a touch screen device was not a significant moderating variable.

Findings also indicate that **use of the _Pocoyo PlaySets_ was especially beneficial for children who had less exposure to English and greater need for developing English-language proficiency.** The effects of the _Pocoyo_ condition (relative to the comparator property) on vocabulary acquisition was greatest (d=.25) among children whose primary caregiver did not speak to them in English; the effect size was negligible (d=.05) among children whose primary caregiver spoke at least some English to them. Similarly, although interacting with _Pocoyo PlaySets_ produced gains in learning among the majority of participants, it did not benefit the few younger children who already scored at the fluent English speaker level on the baseline Pre-IPT.

A complex relationship was found between the relative benefits of using the _Pocoyo Playsets,_ age of child, and level of English language fluency. Neither age of child, or levels of English fluency at pre-test were found to be significant moderators. However, age of child and English fluency were themselves correlated. When both age and fluency in combination were included as moderators, they moderated the extent to which using the _Pocoyo Playsets_ was more likely to result in vocabulary learning relative to the comparator apps. The findings showed that use of _Pocoyo_ produced greater gains in vocabulary in non-fluent children and even in the older (54-71 months) fluent children (who were less fluent than most children their

age). By contrast, younger (36-53 months) children who were fluent did not show gains. This may result from the anomalous nature of these children, who were fluent at an unusually young age within a context where fluency itself was unusual. The number of children in this group was also small, so the standard deviation (SD) was large. Therefore it was difficult to interpret the lack of significant effects in this small subgroup of young fluent children.

The similarity of gains in vocabulary acquisition regardless of children's gender, family income, region in which they live, parents' education or country of birth, all indicate that a wide range of children can benefit from interacting with the *PlaySets,* regardless of differences in many demographics or circumstances. As noted elsewhere, not even a child's previous experience with a touch screen device, or familiarity with the *Pocoyo* property was a significant mediator of the main effect.

These findings are promising for young ELL children, but they also open up the potential for utilizing touch screen applications for young children with a broader range of educational curriculum.  As any number of caregivers will attest, young children enjoy and are engaged by interaction with well-designed touch screen devices and apps. Summary reports from the Teacher Logs (collected as part of this study to assure compliance and available in Appendix C) provide clear trends in the response provided by the 53 teachers from each participating classroom. While these results do not provide rigorous quantitative findings, they do indicate that children enjoyed using the tablets, and showed sustained engagement while using them. Teachers also noted, particularly in the case of the *Pocoyo PlaySets,* that children sang the songs, used the vocabulary and dance moves from the apps in class. The extent of the apps' appeal, potential for engagement, interactivity and repetition of activity in the classroom was

consistently rated by teachers as higher for the *Pocoyo PlaySets* than the comparator apps. The only scale on which teachers rated the *PlaySets* a bit lower was usability by younger children.

It is not surprising, given the number of activities and complexity of the PlaySets, that younger children took time to explore and master them. Teachers' open-ended comments regarding the comparator apps frequently indicated that the children became bored after initial use, or that they were "not interactive." These reports from teachers are particularly important for contextualizing the results of the study. The learning that resulted from children's use of the *Pocoyo PlaySets* occurred in the context of using an engaging and enjoyable device.

The RCT reported here is an important step in providing empirical support for the benefits of using touch screen devices in educational settings with young children. The unique benefits of the *Pocoyo Playsets* build upon the inherent benefits of the touch screen medium. This type of research can play an important role in the development of a new medium.

These findings are promising, and this study raises additional questions to be explored in future research. First, efficacy of the *Pocoyo Playsets* may be affected by other variables, such as the context in which the *PlaySet* is used. For example, the benefits of the *PlaySet* may be optimized when it is used within a complementary curriculum and/or materials, as well as with more active teacher or parent involvement, or for a longer period of time.

Attitude towards learning English as a second language seemed important, based on our initial experiences in day care centers on the East Coast, where ELL children were typically in a minority and there was a potential stigma about being a Spanish speaker. It was with this concern in mind that we developed the brief Attitudes Towards Learning English scale. However, when we traveled to Miami, FL and Los Angeles, CA to work with centers in these

areas, the situation was very different. In many locations the majority if not all of the children and teachers spoke Spanish and the minority were English speakers.  Therefore, it is not a surprise that there were inconsistent results and no effect for Attitudes Towards Learning English. In addition, it suggests that children and teachers have different second language learning needs, depending on the setting and the local culture.

The findings reported here address the efficacy of *Pocoyo PlaySets* used by individual preschool children in an informal placement, without integration into a larger classroom curriculum or use in conjunction with related curriculum materials. These findings cannot address how ELL children's learning of English would be affected by having their teacher actively involved with their use of the *PlaySets* or in teaching the content (English vocabulary concepts related to colors, shapes, vehicles, etc.) as part of the class curriculum.  In addition, it would be interesting to know more about how exposure to educational materials and/or media could reinforce or amplify the curriculum content of the *PlaySets* on other platforms (e.g., books, web-based games, videos, etc.) and increase English language learning. Future studies are needed that include a formal classroom intervention as well as additional, complementary transmedia properties to answer these questions. Such studies should be considered as "next steps" in the evaluation of the potential of educational media on multiple platforms for supporting preschool ELL children's learning of English.

Finally, the relationship between children's *PlaySet* usage patterns and learning outcomes will be addressed in a subsequent report. The analytics from the *Pocoyo PlaySets* are currently being integrated and analyzed. These data will enable us to examine questions about how the use of different *PlaySet* activities contributes to children's learning. These patterns will

provide a way to enrich understanding of the current findings. Specifically, this data may be

useful in making sense of, and possibly moderating and/or mediating the effects reported here.

References

Battenberg, J. K., & Merbler, J. B. (1989). Touch Screen versus Keyboard: A Comparison of Task

Performance of Young Children. *Journal of Special Education Technology*, *10*(2), 24–28.

Brian, M. (2012, January 19). Apple: 1.5 million iPads are used in educational programs, with

over 20,000 education apps. *The Next Web*. Retrieved from

http://thenextweb.com/apple/2012/01/19/apple-1-5-million-ipads-in-use-in-educational-

programs-offering-over-20000-education-apps/

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture

brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology:*

*General*, *136*(4), 539–550.

Cooper, L. Z. (2005). Developmentally appropriate digital environments for young children.

*Library trends*, *54*(2), 286–302.

Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language*

*Learning & Technology*, *15*(2), 2–11.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing Gives Children New Ideas

About Math. *Psychological Science*, *20*(3), 267–272.

Istifci, I. (2011). Opinions of Elementary Level EFL Learners on the Use of Weblogs. *Turkish*

*Online Journal of Distance Education*, *12*(1), 212–222.

Leacox, L., & Jackson, C. W. (2012). Spanish Vocabulary-Bridging Technology-Enhanced

Instruction for Young English Language Learners' Word learning. *Journal of Early*

*Childhood Literacy*, *13*(2).

Lu, C., & Frye, D. (1992). Mastering the machine: A comparison of the mouse and touch screen

for children's use of computers. In I. Tomek (Ed.), *Computer Assisted Learning* (pp. 417–

427). Springer Berlin Heidelberg. Retrieved from

http://link.springer.com/chapter/10.1007/3-540-55578-1_88

Magruder, E. S. (2012). *Research Review: Effective Approaches to Language Development and

Academic Achievement for Young English Language Learners* (Research review conducted

for Ready to Learn Early Learning Collaborative ELL project).

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for

teachers. *The Reading Teacher*, *43*(9), 626–639.

Michael Cohen Group. (2008, October) Effects of *Word World* Viewing on Preschool Children's

Pre-Literacy Skills: A Cluster Randomized Controlled Trial. Research undertaken as part of

the evaluation activities of the U.S. Department of Education Ready to Learn Program and

presented to *Word World*, New York, NY.

Michael Cohen Group. (2013). *Survey on Ownership and Usage of Touch Screen Devices Among

Parents of Young Children* (Report of research conducted as part of ELC initiative within

the Ready to Learn program). New York, NY.

Miller, L.S., & García, E. (2008). A reading-focused early childhood education research and

strategy    development agenda for African Americans and Hispanics at all social class

levels who are English    language speakers or English language learners.  Tempe, AZ:

Arizona State University, Office of Vice President for Education Partnerships.

Nomass, B. B. (2013). The Impact of Using Technology in Teaching English as a Second

Language. *English Language and Literature Studies*, *3*(1), p. 1–11.

Shuler, C. (2012). iLearn II: An Analysis of the Education Category on Apple's App Store. In *Joan Ganz Cooney Center at Sesame Workshop. www. j oanganzcooneycenter. org/Reports-33. html (retrieved January 29, 2012)*. Retrieved from http://www.joanganzcooneycenter.org/Reports-33.html

Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology*, *101*(2), 305–314.

Stevanoni, E., & Salmon, K. (2005). Giving Memory a Hand: Instructing Children to Gesture Enhances their Event Recall. *Journal of Nonverbal Behavior*, *29*(4), 217–233.

U.S. Census Bureau, retrieved December, 2011. Fact Sheet for a Race, Ethnic, or Ancestry Group.  Washington, DC : U.S. Census Bureau; http://factfinder.census.gov/servlet/SAFFFactsCharIteration?_submenuId=factsheet_2&_sse=on.

Verhallen, M. J. A. J., Bus, A. G., & de Jong, M. T. (2006). The promise of multimedia stories for kindergarten children at risk. *Journal of Educational Psychology*, *98*(2), 410–419.

White, E. L., & Gillard, S. (2011). Technology-based literacy instruction for English language learners. *Journal of College Teaching & Learning (TLC)*, *8*(6), 1–6.