Apps As Learning Tools: A Systematic Review
Shayl F. Griffith, PhD, Mary B. Hagan, BA, Perrine Heymann, BPhil, Brynna H. Heflin, BA, Daniel M. Bagner, PhD

CONTEXT: Young children have increasing access to interactive applications (apps) at home and at school. Existing research is clear on the potential dangers of overuse of screens, but there is less clarity around the extent to which interactive apps may be helpful in supporting early learning.

OBJECTIVE: In this systematic review, we present a narrative synthesis of studies examining whether children <6 years can learn from interactive apps.

DATA SOURCES: The PsycInfo, PubMed, ACM Digital Library, and ERIC databases were searched.

STUDY SELECTION: Studies were included if the study design was randomized or nonrandomized controlled (quasi-experimental), the sample mean age was <6 years, the intervention involved children playing with an interactive app, and academic, cognitive, or social-emotional skill outcomes were measured.

DATA EXTRACTION: Of 1447 studies, 35 were included.

RESULTS: Evidence of a learning benefit of interactive app use for early academic skills was found across multiple studies, particularly for early mathematics learning in typically developing children. Researchers did not find evidence of an intervention effect for apps aiming to improve social communication skills in children with autism spectrum disorder.

LIMITATIONS: Risk of bias was unclear for many studies because of inadequate reporting. Studies were highly heterogenous in interventions, outcomes, and study design, making comparisons of results across studies difficult.

CONCLUSIONS: There is emerging evidence to suggest that interactive apps may be useful and accessible tools for supporting early academic development. More research is needed to evaluate both the potential of educational apps to support early learning, and their limitations.

DOI: https://doi.org/10.1542/peds.2019-1579
Accepted for publication Aug 30, 2019
Address correspondence to Daniel M. Bagner, PhD, Department of Psychology and Center for Children and Families, Florida International University, 11200 SW 8th St, Miami, FL 33199. E-mail: dbagner@fiu.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).
Copyright © 2020 by the American Academy of Pediatrics

Touchscreen mobile devices (eg, smartphones, tablets) have become ubiquitous for young children. Interactive applications or “apps” considered “educational” for young children have similarly gained in popularity and are increasingly being integrated into early childhood classrooms as learning tools because of perceived advantages for child engagement and active learning. The integration of interactive app technology into children's lives at home and school has outpaced research needed to inform comprehensive recommendations for its use. Recommendations have thus far focused on preventing overuse of screens rather than opportunities for maximizing learning. Research on whether young children can learn from interactive apps; the academic, cognitive, or social-emotional skill domains that may be best supported by interactive apps; and the conditions under which this learning may be maximized; is still emerging.

In 2016, the American Academy of Pediatrics (AAP) recommended screen time (including touchscreen) for 2- to 5-year-olds be limited to <1 hour per day of “high-quality educational programming” and that parents coview screens with children. Given the scarcity of studies on interactive screen media, the AAP statement largely referenced studies from the more-extensive literature on television, in which researchers documented the detrimental effects of overconsumption of television on sleep, obesity, and cognitive and language development but also showed that high-quality and developmentally appropriate programming (eg, Sesame Street) can support vocabulary, math, and social-emotional skills in children >30 months.

Theories on learning suggest children learn best when activities promote active “minds-on” learning, are appropriately scaffolded, and are engaging. Additionally, research shows that young children benefit from repeated and varied practice of foundational skills. Interactive apps have potential to support active engagement with learning material through embedding learning concepts into gamelike activities, scaffolding children’s learning through adaptive learning technology, providing feedback and rewards through gameplay, and promoting repeated practice of important foundational skills. Additionally, the ubiquity of touchscreen devices could facilitate more equitable and convenient access to educational materials in the home, in contrast to other types of technology (eg, computers), which still show marked income gaps in access. Research on whether interactive educational touchscreen media can support early learning is therefore essential for updating our understanding of “high-quality educational programming” for the new digital age.

Studies in diverse fields, including psychology, education, and computer science, have begun to investigate children’s learning from apps, with most studies published in the past 3 years. The findings of these studies have not yet been well synthesized. In existing review articles, authors have provided commentaries on children’s use of mobile technology, rather than systematic reviews of the literature. In one recent meta-analysis, authors reported a significant positive effect across studies of young children’s learning from touchscreen devices. However, there was significant clinical and methodological heterogeneity in the included studies, which encompassed varied types of interventions (eg, electronic books, interactive game apps), outcomes targeted, comparison groups, and study designs (randomized and nonrandomized, controlled and noncontrolled). Statistical heterogeneity in this study was high (I^2 >90%), limiting interpretability of a meta-analytic effect size. Instead, a narrative synthesis of results is the recommended approach when studies are highly variable in interventions, outcomes, study design, and risk of bias.

Thus, in this study, we aimed to systematically review the rapidly expanding but heterogeneous body of evidence for young children’s learning from touchscreen interactive apps. Given the popularity of interactive educational apps for young children; the importance of early learning and school readiness for later academic, behavioral, and emotional functioning; and the existing questions about whether young children can apply information learned on a screen to real life; we focused on children younger than 6 years in this review. The specific research questions were as follows:

1. Is there evidence that children <6 years can learn from using interactive touchscreen apps?
2. What content subjects or skills are particularly well suited to touchscreen app learning?
3. How does learning from touchscreen interactive app games compare with children’s learning in other instructional contexts, including in-person instruction or educational video?

METHODS

Data Sources

An electronic database search was conducted in 4 databases popular to the fields of psychology, pediatrics, education, and computer science (PsycINFO, PubMed, ERIC, and ACM Digital Library) by using the following Boolean search phrase: (tablets OR iPads OR touchscreen OR apps OR interactive media) AND (learning OR literacy OR mathematics OR skills OR development OR functioning OR outcomes OR achievement) AND (toddlers OR young children OR...
preschoolers). Articles published before January 1, 2008 (the year Apple and Android app stores opened), and written in languages other than English were excluded. The database searches were conducted on January 22, 2019. Additionally, the reference sections of included articles and relevant review articles were examined to identify any other studies meeting inclusion criteria (n = 2).

**Study Selection**

Study inclusion and exclusion criteria are presented in Table 1. The search initially identified 1447 unique records. After initial screening, 115 records were subject to full text review by 2 independent coders from a group of 4 (authors S.F.G., B.H.H., M.B.H., P.H.). Disagreements were discussed, and consensus was reached. Randolph’s $k^2$ was 0.84, indicating excellent interrater reliability. Thirty-four articles, describing 35 studies, were included (see Fig 1 for Preferred Reporting Items for Systematic Reviews and Meta-analyses diagram).

**Data Extraction and Synthesis**

The following information was extracted from each study: design, sample size, location of study, language of participants, length of intervention, type of control group, name of app(s), study context, targeted subject(s) of intervention, outcome measures, and results. Given the number of articles (n = 12) for which there was not sufficient information to calculate an effect size and the clinical, methodological, and statistical heterogeneity of the included studies in terms of design, outcomes measured, and comparison group used, formal meta-analytic strategies were not used to summarize overall effects. Instead, we presented a narrative synthesis of the included studies according to the targeted subject of the intervention. Additionally, to provide context as to clinical significance of study findings, where possible, we presented ranges of effect sizes (using Cohen’s $d$).

**TABLE 1**

<table>
<thead>
<tr>
<th>Category</th>
<th>Inclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Children 5 y and under (mean age of sample &lt;72 mo or identifiable subgroup with mean &lt;72 mo). Studies conducted in any country, with samples of any language background, were included, as long as the article was published in English. Clinical and nonclinical samples were included.</td>
</tr>
<tr>
<td>Intervention</td>
<td>Intervention involves use of an interactive game app on a touchscreen. “Interactive” suggests that children’s interaction with the screen has an effect on the material presented. Studies were excluded if they involved only interventions with electronic books, computers, noninteractive video, or multimedia or transmedia interventions. Studies in which a touchscreen device or app was used only as an assistive technology (eg, iPad version of a Picture Exchange Communication System) were excluded.</td>
</tr>
<tr>
<td>Comparison</td>
<td>Any comparison group type (eg, treatment or instruction as usual, no treatment, attention control, waitlist control, noninteractive video control, alternative app condition control, etc).</td>
</tr>
<tr>
<td>Outcome</td>
<td>Academic achievement, executive functioning skills, socio-emotional skills, and task-specific learning (eg, learning nonsense names, learning to find an object, learning facts). Studies were excluded if only process outcomes (eg, engagement) were measured.</td>
</tr>
<tr>
<td>Study design</td>
<td>Randomized controlled designs (including laboratory experiments and field trials) and quasi-experimental designs. Case studies, reviews, multiple baseline or multiple probe studies, correlational studies, and descriptive studies were excluded.</td>
</tr>
</tbody>
</table>
representing post-test differences between the app intervention and control groups for the randomized studies.

Assessment of Risk of Bias

Because most included studies had a randomized design, risk of bias was assessed with the Cochrane’s risk of bias tool for randomized trials²⁴ that was adapted to include only categories relevant to psychological and educational research. Studies were rated in the following categories: randomization (selection bias), blinding of outcome assessment (detection bias), missing data (attrition bias), and selective reporting (reporting bias). Quasi-experimental studies, for which assignment to group was not random (eg, administrative convenience; n = 5), were also evaluated by using this tool to capture risk of bias due to nonrandomized assignment in the randomization category. Ratings of “high,” “low,” or “unclear” risk were assigned for each domain. Each study was coded for risk of bias independently by 2 coders from a group of 4 (coders stated above). Kappa values were 0.91, 0.61, 0.51, and 0.96 for randomization, missing data, outcome assessment, and selective reporting, respectively, indicating moderate to excellent agreement.²⁵ All disagreements were discussed, and a consensus was reached between coders.

RESULTS

Description of Studies

The total number of participants across studies was 4639. Eighteen studies were conducted in the United States, 5 in Australia, 4 in the United Kingdom, 2 in Canada, and 1 in each of the United Arab Emirates, Italy, Greece, Croatia, the Netherlands, and Germany. The primary language of most of the samples (n = 30) was English (for studies conducted in the United States, Australia, the United Kingdom, Canada, and the United Arab Emirates). The 5 remaining studies included samples in which the primary language spoken was not English (Italian, Greek, Croatian, Dutch, and German, respectively); in each case, the interventions and assessments were delivered in the primary language of the sample. Samples included typically developing children, except for the studies in which researchers examined social communication (n = 3), which all consisted of children diagnosed with autism spectrum disorder (ASD). In 12 studies, researchers used apps developed by the study authors (noted in Table 2); in 4 of these 12 studies, the app was commercially available, suggesting potential financial profit for the author.²⁶–²⁹ In the remaining studies, researchers used existing apps developed by nonrelated companies.

Twenty-nine studies had randomized designs, including 18 randomized controlled trials (RCTs) with app interventions delivered over an extended period (ranging from 5 days to 1 academic year) and 11 laboratory-based experimental studies, which included interventions (ie, learning phases) delivered over a short time span (ranging from 5 to 20 minutes). For simplicity, studies with randomized designs and interventions delivered over an extended time are referred to as RCTs. Laboratory-based experiments (all of which also have randomized designs but a short-term intervention phase) are referred to as experimental studies. The remaining 6 studies included in this review were quasi-experimental (ie, nonrandomized controlled) trials. Comparison groups varied across studies, including in-person instruction (eg, classroom-based, in-person, using manipulatives), video instruction, or other non-app-based instruction (eg, pen and paper practice); attention controls receiving apps not targeting the identified outcome; or groups receiving no treatment. Five studies included interventions delivered at home, 18 included interventions delivered in the school, and 12 were delivered in a laboratory. The studies had a variety of outcomes in targeted educational domains, including mathematics (n = 15), language arts (LA) (n = 11), executive functioning (n = 5), health and science knowledge (n = 3), and social communication (n = 3). One study included targeted outcomes in both mathematics and science,⁶² and another included math and LA outcomes.⁶⁰ Each of the studies examined potential learning or skill benefits of interactive apps. Although not specifically excluded from the search strategy, no identified study was designed to examine harmful effects of interactive app use on functioning. See Table 2 for a summary of studies.

Risk of Bias Within Studies

Overall, 19 of the 35 studies (54%) were at high risk of bias in at least 1 domain. High risk of bias ratings were given for nonrandom assignment to groups, substantial missing data and data not missing at random (eg, imbalances in missing data between groups), and lack of blinding of outcome assessors. Additionally, risk of bias was judged to be unclear due to inadequate reporting of methods in at least 2 domains for 19 of the 35 studies. Only 3 studies had low risk of bias across the randomization, missing data, and outcome assessment domains, and no studies were given low risk ratings across all 4 domains. The results of the studies should be interpreted with caution given the literature lacks rigorously designed RCTs.

Narrative Synthesis of Studies by Learning Outcomes

Early Math Learning

In 12 randomized studies (9 RCTs) and 3 quasi-experimental studies, researchers examined children's math
**TABLE 2 Studies Included in the Systematic Review**

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Study Population</th>
<th>Study Design</th>
<th>Length of Intervention</th>
<th>App Intervention Condition and Name of App</th>
<th>Control Condition</th>
<th>Outcome (Measure)</th>
<th>Intervention Condition Outperformed Control Condition?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning domain:</strong> early math learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aladé et al,30 2016a</td>
<td>N = 60, mean age = 58.06 mo, 42% male</td>
<td>Experimental</td>
<td>3 trials</td>
<td>Measure That Animal</td>
<td>Noninteractive video of gameplay in Measure that Animal</td>
<td>Measurement (measuring tasks similar and less similar to app gameplay)</td>
<td>Yes (for measuring task similar to app)</td>
</tr>
<tr>
<td>Han,31 2018</td>
<td>N = 40, mean age = 42.08, 50% male</td>
<td>Experimental</td>
<td>12 training trials</td>
<td>Math app developed for study</td>
<td>No treatment</td>
<td>Quantity knowledge (quantity recognition and Give-N tasks)</td>
<td>No</td>
</tr>
<tr>
<td>Kooko and Ferdig,32 2016a</td>
<td>N = 73, mean age = 51.54 mo</td>
<td>RCT</td>
<td>3 wk</td>
<td>Zorbit's Math</td>
<td>No treatment</td>
<td>Broad math skills (19-item measure created for study)</td>
<td>No</td>
</tr>
<tr>
<td>Mattoon et al,33 2015a</td>
<td>N = 24, range = 4–5 y, 62.5% male</td>
<td>RCT</td>
<td>6 wk</td>
<td>DinoKids – Math Lite, Hungry Fish, Motion Math, Math Express, Maths Age 3–5, Park Math</td>
<td>Instruction with traditional math manipulatives</td>
<td>Broad math skills (TEMA-3)</td>
<td>No</td>
</tr>
<tr>
<td>Miller,34 2018a</td>
<td>N = 13, range = 4–5 y, 53.8% male</td>
<td>RCT</td>
<td>10 d</td>
<td>Package of 15 math apps</td>
<td>Classroom instruction</td>
<td>Broad math skills (22-item measure created for study)</td>
<td>No</td>
</tr>
<tr>
<td>Outhwaite et al,35 2019a</td>
<td>N = 389, mean age = 60.64 mo, 49.6% male</td>
<td>RCT</td>
<td>12 wk</td>
<td>Maths 3–5, Maths 4–6</td>
<td>Classroom instruction</td>
<td>Broad math skills (Progress Test in Math Level 5)</td>
<td>Yes</td>
</tr>
<tr>
<td>Outhwaite et al,36 2017a</td>
<td>Range = 50–61 mo, 48.1% male</td>
<td>Quasi-experimental</td>
<td>16 wk</td>
<td>Maths 3–5, Maths 4–6</td>
<td>Classroom instruction</td>
<td>Broad math skills (48-item and 50-item measures created for study)</td>
<td>Yes</td>
</tr>
<tr>
<td>Papadakis et al,37 2018a</td>
<td>N = 365, mean age = 62.0 mo, 48.5% male</td>
<td>RCT</td>
<td>14 wk</td>
<td>16 math apps designed for the study</td>
<td>Classroom instruction</td>
<td>Broad math skills (TEMA-3)</td>
<td>Yes</td>
</tr>
<tr>
<td>Park et al,38 2016a</td>
<td>N = 103, mean age = 4.87 y, 51.5% male</td>
<td>RCT</td>
<td>10 sessions over 2–3 wk</td>
<td>Math app designed for study</td>
<td>Memory game app</td>
<td>Broad math skills (TEMA-3)</td>
<td>Yes</td>
</tr>
<tr>
<td>Schacter and Jo,26 2016a</td>
<td>N = 227, mean age = 56.00 mo, 46.5% male</td>
<td>Quasi-experimental</td>
<td>15 wk</td>
<td>Math Shelf</td>
<td>Classroom instruction</td>
<td>Broad math skills (iPad assessment created for the study)</td>
<td>Yes</td>
</tr>
<tr>
<td>Schacter and Jo,27 2017a</td>
<td>N = 433, mean age = 4.55 y, 50.8% male</td>
<td>RCT</td>
<td>22 wk</td>
<td>Math Shelf</td>
<td>Classroom instruction</td>
<td>Broad math skills (iPad assessment created for the study)</td>
<td>Yes</td>
</tr>
<tr>
<td>Schacter et al,39 2016a</td>
<td>N = 100, mean age = 55.85 mo, 48% male</td>
<td>RCT</td>
<td>6 wk</td>
<td>Math Shelf</td>
<td>Top 5 math apps in the app store</td>
<td>No sense abilities (iPad assessment created for the study)</td>
<td>Yes</td>
</tr>
<tr>
<td>Blackwell,40 2015</td>
<td>Range = 4–5 y</td>
<td>RCT</td>
<td>5 d</td>
<td>Know Number Free</td>
<td>Classroom instruction</td>
<td>Numeracy (correctly written No. 1–10)</td>
<td>Yes</td>
</tr>
<tr>
<td>Blackwell,41 2013</td>
<td>N = 352, mean age = 5.56 y, 51.1% male</td>
<td>Quasi-experimental</td>
<td>1 y</td>
<td>Shared iPad in classroom with set of 70 apps</td>
<td>Classroom instruction</td>
<td>Early literacy skills (Star Early Literacy Assessment)</td>
<td>Yes</td>
</tr>
<tr>
<td>Brown and Harmon,42 2013a</td>
<td>N = 20, range = 48–59 mo</td>
<td>RCT</td>
<td>1 wk</td>
<td>3 educational iPad apps targeting upper and lowercase alphabet knowledge, matching, and No. concepts</td>
<td>Classroom instruction</td>
<td>Alphabet knowledge (PALS Pre-K Upper and Lower Case Alphabet Knowledge subtests and measures created for study)</td>
<td>No</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Population</td>
<td>Study Design</td>
<td>Length of Intervention</td>
<td>App Intervention Condition and Name of App</td>
<td>Control Condition</td>
<td>Outcome (Measure)</td>
<td>Intervention Condition Outperformed Control Condition?</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------</td>
<td>--------------</td>
<td>------------------------</td>
<td>-------------------------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Cubelic and Larwin, 2014</td>
<td>N = 281, kindergarten, “relative balance of boys and girls”</td>
<td>Quasi-experimental</td>
<td>One academic year</td>
<td>Apps targeting first sound fluency (ABC Phonics, Build a Word, Magic Penny, Starfall), phoneme segmentation fluency (First Word Sampler, Word Wall IQ, Pocket Phonics, Skills Builder Spelling, Phonics Monster 1), and nonsense word fluency (ABC Touch and Learn, Bee Sees, Kindergarten Lite, Starfall, Super Why)</td>
<td>Classroom instruction</td>
<td>Phonological awareness (DIBELS Next Assessment Phoneme Segmentation Fluency and Nonsense Fluency subtests)</td>
<td>Yes for Phoneme Segmentation Fluency and Nonsense Fluency</td>
</tr>
<tr>
<td>Gilliver et al, 2016</td>
<td>N = 30, mean age = 36.5 mo, 50% Male</td>
<td>RCT</td>
<td>6 wk</td>
<td>Phonological awareness app created for the study</td>
<td>Vocabulary app created for the study</td>
<td>Phonological awareness and vocabulary (measure created for study)</td>
<td>Yes</td>
</tr>
<tr>
<td>Neumann, 2018</td>
<td>N = 48, mean age = 45.19 mo, 52.1% male</td>
<td>RCT</td>
<td>9 wk</td>
<td>Endless Alphabet, Letter School, the Draw App</td>
<td>Classroom instruction</td>
<td>Letter name and sound knowledge, name writing (measure created for study)</td>
<td>Yes</td>
</tr>
<tr>
<td>Patchan and Puranik, 2016</td>
<td>N = 170, mean age = 51.9 mo, 45% male</td>
<td>RCT</td>
<td>8 wk</td>
<td>Writing practice with iPad app (finger drawing)</td>
<td>Writing practice with pen and paper</td>
<td>Alphabet knowledge (naming 26 letters and writing 26 uppercase letters)</td>
<td>Yes</td>
</tr>
<tr>
<td>Russo-Johnson et al, 2017</td>
<td>N = 71, mean age = 41.05 mo, 48.2% male</td>
<td>Experimental</td>
<td>6 1-min trials</td>
<td>Word learning app designed for the study</td>
<td>Noninteractive video of gameplay in word learning app</td>
<td>Word learning (No correct responses in object naming)</td>
<td>No</td>
</tr>
<tr>
<td>Teepe et al, 2017</td>
<td>N = 71, mean age = 40.05 mo, 40.0% male</td>
<td>RCT</td>
<td>2 10-min intervention sessions over 2 wk</td>
<td>Jeffy’s Journey</td>
<td>No treatment</td>
<td>Expressive and receptive vocabulary (measure created for study)</td>
<td>No</td>
</tr>
<tr>
<td>Vatalaro et al, 2018</td>
<td>N = 65, range = 3–5 y, 52.5% male</td>
<td>Quasi-experimental</td>
<td>8 wk</td>
<td>Endless Alphabet, Noodle Words, Goodnight ABC, ABC GO, Beck and Bo, Draw and Tell, Don’t let the Pigeon Run this App, Alien Assignment</td>
<td>Apps chosen and used by the study preschool (Letter School, Gazzili Science, Yumilo Rainbow Power, Faces I Make, Counting Bear)</td>
<td>Receptive vocabulary skills (PPVT) and expressive vocabulary skills (EVT)</td>
<td>No</td>
</tr>
<tr>
<td>Walter-Laager et al, 2017</td>
<td>N = 98, mean age = 27.3 mo, 53.1% male</td>
<td>Experimental</td>
<td>20 min</td>
<td>Lingua Kidz used with an adult</td>
<td>Vocabulary picture cards, used with an adult</td>
<td>Language development (abridged version of the standardized “Language and Development Test for 2-Year-Olds”)</td>
<td>No</td>
</tr>
<tr>
<td>Learning domain: science and health</td>
<td>Kwok et al, 2016</td>
<td>Experimental</td>
<td>8 training trials</td>
<td></td>
<td></td>
<td>No correct answers on 4 animal facts</td>
<td>No</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Population</td>
<td>Study Design</td>
<td>Length of Intervention</td>
<td>App Intervention Condition and Name of App</td>
<td>Control Condition</td>
<td>Outcome (Measure)</td>
<td>Intervention Condition Outperformed Control Condition?</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>--------------</td>
<td>------------------------</td>
<td>--------------------------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Putnam et al, 2018a</td>
<td>N = 86, mean age = 66.93 mo, 62.8%</td>
<td>Other</td>
<td>5 d</td>
<td>App developed for study to teach animal facts</td>
<td>Face-to-face instruction on animal facts</td>
<td>No treatment</td>
<td>Recall (No times correctly names healthy and unhealthy foods)</td>
</tr>
<tr>
<td>Learning domain: math and ELA</td>
<td>Bebell and Pedulla, 2015a</td>
<td>Quasi-experimental</td>
<td>1 y</td>
<td>Set of commercially available apps targeting ELA and math used in the classroom (during school-wide iPad implementation year)</td>
<td>Classroom instruction as usual (in the year before the iPad implementation)</td>
<td>Early literacy and broad math skills (CPM 61 ELA and Math Subtests)</td>
<td>Yes</td>
</tr>
<tr>
<td>Learning domain: math and science</td>
<td>Schroeder and Kirkorian, 2016a</td>
<td>Experimental</td>
<td>4 trials</td>
<td>Two mini games in the app Mesozoic Math Adventures (Don's Collections [Math] and Life Cycles [Science])</td>
<td>Noninteractive video of gameplay in word learning app</td>
<td>Quantity, knowledge of growth cycle (measures created with items similar to app gameplay)</td>
<td>Yes (for knowledge of the growth cycle for older preschoolers); no (for quantity task for younger preschoolers)</td>
</tr>
<tr>
<td>Learning domain: executive functioning</td>
<td>Choi and Kirkorian, 2016a</td>
<td>Experimental</td>
<td>4 trials</td>
<td>App that required children to tap on specific location to reveal the hiding place of a character</td>
<td>App that showed a noncontingent video of the hiding place of a character</td>
<td>Responses on object retrieval trials</td>
<td>Yes (for younger toddlers)</td>
</tr>
<tr>
<td>Huber et al, 2016a</td>
<td>Study 1: N = 18, mean age = 5.4 y, 50% male; study 2: N = 50, mean age = 5.1 y, 46% male</td>
<td>Experimental</td>
<td>4 min trials</td>
<td>Practice with the app Extra Tower of Hanoi</td>
<td>Practice with 3D Tower of Hanoi Model</td>
<td>No extra moves on 30 Tower of Hanoi task</td>
<td>No</td>
</tr>
<tr>
<td>Huber et al, 2018a</td>
<td>N = 96, mean age = 36.6 mo, 56.5% male</td>
<td>Experimental</td>
<td>9 min of screen time</td>
<td>Early math app (Shiny Party)</td>
<td>Cartoon (Penguins of Madagascar)</td>
<td>Delayed gratification (gift-delay task) and working memory (working memory task)</td>
<td>Yes</td>
</tr>
<tr>
<td>Tarasuik et al, 2017a</td>
<td>N = 49, mean age = 4.9 y, 45% male</td>
<td>Experimental</td>
<td>4 trials</td>
<td>Practice with the app Extra Tower of Hanoi</td>
<td>Practice with 3D Tower of Hanoi Model</td>
<td>No extra moves on 30 Tower of Hanoi task</td>
<td>No</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Study Population</td>
<td>Study Design</td>
<td>Length of Intervention</td>
<td>App Intervention Condition and Name of App</td>
<td>Control Condition</td>
<td>Outcome (Measure)</td>
<td>Intervention Condition Outperformed Control Condition?</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>------------------------</td>
<td>---------------------------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td><strong>Learning domain:</strong> social communication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esposito et al., 2017a</td>
<td>N = 30, mean age = 47 mo, 90% male</td>
<td>RCTd</td>
<td>4 wk</td>
<td>Standard ABA + 3 apps created for the study to promote attention, imitation, and vocabulary</td>
<td>Standard ABA</td>
<td>Attention, vocabulary, and imitation (mastered targets in standard ABA)</td>
<td>No</td>
</tr>
<tr>
<td>Fletcher-Watson et al., 2016a</td>
<td>N = 54, mean age = 49.63 mo, 79.6% male</td>
<td>RCTd</td>
<td>2 mo</td>
<td>App targeting social communication skills, attending to people and following social cues (FindMe)</td>
<td>Treatment as usual</td>
<td>Social communication skills (MCDI, CSBS-DP, BOSCC), and presence of Autism (ADOS-2)</td>
<td>No</td>
</tr>
<tr>
<td>Whitehouse et al., 2018a</td>
<td>N = 80, mean age = 34.83 mo, 79.6% male</td>
<td>RCTd</td>
<td>6 mo</td>
<td>Community therapy + App designed to enhance ABA therapy social communication outcomes (TherapyOutcomes by You)</td>
<td>Community therapy</td>
<td>Social communication skills (MCDI, CSBS-DP), restricted and repetitive behaviors (RBS-R), behavioral flexibility (BFRS), developmental skills (ATEC), and adaptive skills (VABS-2)</td>
<td>No</td>
</tr>
</tbody>
</table>

ABA, Applied Behavior Analysis; ADOS-2, Autism Diagnostic Observation Schedule; ATEC, Autism Treatment Evaluation Checklist; BFRS, Behavior Flexibility Rating Scale; BOSCC, Brief Observation of Social Communication Change; CPAA, Children’s Progress Academic Assessment; CSBS-DP, Communication and Symbolic Behavior Scale, Developmental Profile; DIBELS, Dynamic Indicators of Basic Early Literacy Skills; ELA, English Language Arts; EVT, Expressive Vocabulary Test; MCDI, MacArthur Communicative Development Inventory; PALS Pre-K, Phonological Awareness Literacy Screener, Preschool; PPVT, Peabody Picture Vocabulary Test; RBS-R, Repetitive Behavior Scale, Revised; TEMA-3, Test of Early Mathematics Ability, Third Edition; VABS-2, Vineland Adaptive Behavior Scales, Second Edition.

a Peer-reviewed.

b Study context is laboratory.

c App was developed by study authors; italic script indicates the app is commercially available.

d Study context is home.

e Study context is school.
outcomes after using math interactive apps. In 11 of the 15 studies, researchers examined school-based interventions; a home-based intervention was examined in 1; and 3 were laboratory-based experiments. Outcomes included measures of early math skills (including number recognition, number naming, and basic addition and subtraction). School and home interventions ranged from 5 days to 1 academic year. Overall, 7 of the 12 randomized studies (including 1 of the 3 laboratory-based experiments) and all 3 quasi-experimental studies showed better outcomes in the math app intervention group compared with the control group used.

Specifically, of the 9 studies (6 RCTs) that compared a math app intervention to in-person classroom instruction, 7 (4 RCTs) reported significantly greater math learning gains on early math measures (n = 8) and a number-writing task (n = 1) for groups using interactive math apps, whereas the remaining 2 RCTs reported statistically equivalent gains between the groups.33,35 Of the 4 studies (3 RCTs) in which researchers compared a math app intervention to control groups using apps not targeting math (ie, attention controls) or not receiving treatment, 3 reported better outcomes on early math measures for app intervention groups compared with controls, whereas 1 reported no significant differences between groups.32 Effect sizes, when reported (n = 9, from 9 RCTs with intervention periods between 5 days and 22 weeks), ranged from 0.04 to 0.94.

In 2 experimental studies, researchers examined math learning from interactive apps compared with noninteractive video, with mixed results. Younger preschoolers’ (3–4.5 years) performance on a quantity math task improved after watching an instructional video, but not after playing a math app, whereas older preschoolers (4.5–5.5 years) improved in neither condition.62 In the other study, preschoolers’ (3.5–5.5 years) performance on a measuring task improved in both interactive app and noninteractive video conditions.30

**LA Learning**

In 7 randomized studies (5 RCTs) and 4 quasi-experimental studies, researchers examined children’s LA outcomes (letter knowledge, phonological awareness, letter writing, or vocabulary). Eight of the 11 studies were based in schools, whereas 3 were laboratory-based experiments. Interventions ranged from 1 week to 1 academic year, and most used commercially available apps. Overall, 3 of the 7 randomized studies (including 0 of the 3 experimental studies) and 3 of the 4 quasi-experimental studies showed better outcomes in the app intervention group compared with the control group, whereas the remainder showed statistically equivalent outcomes between groups.

Specifically, in 3 studies (2 RCTs), researchers examined alphabet knowledge and phonological awareness after children used apps targeting these skills, compared with apps not targeting these skills. No significant differences in learning were found between groups for 2 studies,44,53 whereas the other found children’s phonological awareness and vocabulary skills improved more after using a phonological awareness app compared with a vocabulary app.48 In 5 studies (2 RCTs), researchers found that children’s letter naming and writing skills46,66,77,78 and phonological awareness46,66,78 improved after using a set of LA apps compared with a classroom or in-person instruction control group. In 2 experimental studies, researchers examined children’s vocabulary learning using an app compared with nonapp instruction (noninteractive video and picture card instruction) and found no differences in learning.51,56 In 1 final study, researchers found that children’s receptive and expressive vocabulary did not improve after 2 10-minute sessions using a storytelling app with their parents compared with no treatment.52 Effect sizes for the RCTs, when reported (n = 12, from 3 school-based interventions over 1–9 weeks), ranged from 0.07 to 0.94. Effect sizes for the experimental studies, when reported (n = 3, from 2 studies), ranged from 0.41 to 0.60.

**Health and Science Facts**

In 3 randomized studies (1 home-based RCT and 2 experimental studies), researchers examined children’s learning of health or science concepts. In the studies, researchers examined children’s learning of healthy versus unhealthy foods, the growth cycle, and facts about animals. Overall, the 3 studies indicated children learned from the apps, with some exceptions. In 1 study, researchers found no difference in learning facts about animals between an interactive app group and a group receiving in-person instruction,58 whereas researchers in another found the app group learned more about healthy foods compared with a no-treatment control.59 In a third study, researchers found that older preschoolers (4.5–5.5 years) learned about the growth cycle better from interactive app play compared with noninteractive video instruction, but younger preschoolers (3–4.5 years) were not able to learn in either condition.62

**Executive Functioning**

In 4 experimental studies, researchers examined children’s performance on executive functioning tasks. In 2 studies, researchers found that children improved their performance on the traditional three-dimensional (3D) version of the
Tower of Hanoi task after practice trials with a touchscreen app version of the task, and this improvement was not significantly different from children who both practiced and were tested with the 3D version. In the third study, researchers found that young toddlers (mean age = 25 months) were better able to learn how to retrieve objects in a real-life situation when they received training with an interactive app compared with a noninteractive video; however, this effect was reversed for older toddlers (mean age = 33 months). In the fourth study, researchers found that children were more likely to pass a gift-delay task and performed better on a working memory task after playing an educational app game compared with watching noneducational television.

Social Communication

In 3 RCTs, researchers examined whether social communication skills improved after using apps designed to improve social skills, such as imitation, attention to social stimuli, and following social cues. Each study included a clinical sample of children with ASD; no studies with social development outcomes in typically developing children met inclusion criteria for this review. In these 3 studies, researchers compared outcomes in a treatment group receiving the app at home in addition to standard applied behavioral analysis or other community therapies to a control group receiving only treatment as usual. For 2 of the 3 studies, children and parents used the app together, whereas in the third study, children used the app alone. None of the 3 studies reported significant improvement in the primary social communication skills outcome measures for the app treatment group compared with the control group. Effect sizes for gains in the app groups on social communication outcomes ranged from 0 to 0.40.

DISCUSSION

In this article, we systematically reviewed 35 studies on young children’s learning from interactive educational apps. The studies were heterogeneous in targeted outcomes, length of intervention, context of intervention, and experimental design. Evidence for learning from touchscreen interactive app use was strongest for math-related app interventions and weakest for interventions targeting social communication skills. Study quality was mixed overall, with the risk of bias for many studies unclear in more than 1 domain, often from deficits in methods reporting.

Can Young Children Learn From Interactive Touchscreen Apps?

A majority of studies comparing math or LA (letter naming and writing, phonological awareness) outcomes in children using interactive apps to controls receiving equivalent regular classroom or small-group instruction found more favorable outcomes for the interactive app condition. In addition, although study effect sizes varied widely, large effects were reported for school-based studies of math and LA learning. For example, the largest effect sizes in math were from a school-based 22-week intervention using the app Math Shelf, whereas the largest effect sizes in LA were from a school-based 6-week intervention using a phonological awareness app. These findings are notable because they suggest the addition of interactive touchscreen technology may afford a learning advantage for early academic learning exceeding traditional modes of instruction and are consistent with the theory that interactive apps are well suited to support active, repeated, and varied practice of skills.

Results of other studies reviewed suggest touchscreen app games have the potential to support learning in other skill areas, including science knowledge and executive functioning. However, there were a small number of studies targeting these skill areas, and further research is needed to establish whether children’s understanding of science concepts could be promoted using interactive apps, and similarly, whether improvement in executive functioning task performance is task specific or would generalize to other measures of executive functioning.

Results from the 3 RCTs in which researchers examined app interventions targeting social communication skills for children with ASD suggest these skills may not be best supported by touchscreen apps. This finding is particularly important to note because touchscreen technology is popular in ASD treatment, including electronic augmentative exchange communication systems. Social communication skills may be less easily generalized to the real world, compared with foundational early math and preliteracy skills, which is consistent with research showing improvements in app gameplay but not real-world social communication skills. However, given the small number of studies examining this outcome, further research in this area is needed.

Finally, findings from laboratory experimental studies on children’s learning from interactive apps compared with noninteractive video were mixed. In general, researchers in these studies examined learning in younger children who have difficulty transferring learning from media. Although in some studies researchers found that task performance improved with an interactive learning
phase, others found the interactive component may have interfered with learning because of overtaxed attentional resources. Although other studies show older children may benefit from learning through interactive app play, further research is needed to establish the lower age limit of this potential benefit.

**Future Research and Clinical Implications**

Important gaps were identified and should be addressed in future research. First, more large-scale randomized trials of apps are needed, with focus on clarifying what app features and content may best support learning. Thousands of apps are added to app stores every month, so large-scale trials of each new educational app may not be feasible. Understanding features of interactive apps supporting learning in different content areas will allow clearer standards regarding apps to be deemed “educational” by app makers.

Second, most studies reviewed were conducted in schools and examined math and LA outcomes. Given touchscreen devices are widely present in homes and the AAP’s recommendation that screen time at home be educational, it will be important to examine educational benefits of touchscreen media use at home. More studies examining other learning content areas are also needed. Third, almost all app interventions reviewed involved child solo use of apps. Children learn best when co-viewing television with parents, and parent behaviors while co-using apps have been linked to child positive engagement and affect. Thus, examining children’s learning from app games with parental co-use is an important area for future research.

Fourth, all but 5 of the reviewed studies included app interventions delivered in English to English-speaking populations, and most included studies based in the United States, making it difficult to evaluate the impact of country or language on learning from apps. There was no apparent pattern to the results of studies according to country or language; for example, of the 4 studies in which researchers examined children’s math outcomes conducted outside of the United States, 3 showed better performance in the app condition, including the 1 study conducted in Greek. Further research is needed on the extent to which educational apps can support learning in a child’s nonpreferred language as well as on the generalizability to other languages of the findings presented on LA learning from apps.

Fifth, further research is needed to evaluate important moderators of treatment effects. There is emerging evidence from the studies reviewed that age may be an important moderating factor, and children below preschool age may have more difficulty learning from screen media. Future researchers should more closely examine moderating factors that may impact learning from touchscreen apps, including children’s pretest ability and demographic characteristics, such as sex, socioeconomic status, and previous experience and familiarity with technology.

Finally, in a separate body of research on television, researchers have identified negative effects of excessive screen use on children’s development. Although the authors of the articles in this review did not explicitly examine negative effects of use or overuse, as the popularity and diversity of digital media for children increases, it will be necessary for researchers and clinicians to balance the potential positives of interactive screen media for learning with the imperative to prevent overuse. Further research examining the relative advantages and disadvantages of various types of screen media (eg, television versus interactive touchscreen) will be important to update recommendations to better reflect the new digital context.

**Limitations**

There were some limitations to this systematic review. First, risk of bias was unclear for many studies, and many studies had incomplete reporting of study methods. Second, although the study included unpublished research, it is possible additional unpublished research was not identified. Unpublished studies may have fewer positive outcomes as a result of publication bias. Third, because of the nascent state of the literature, the summarized studies were highly heterogeneous. Thus, it was not possible to compare results across studies or investigate moderators of these effects.

**CONCLUSIONS**

In this systematic review, we examined the literature on educational app learning in young children. Evidence of learning benefits from educational touchscreen apps was found across multiple studies, particularly for early mathematics. Conversely, multiple studies did not find evidence of an intervention effect for apps aiming to improve social communication skills in children with ASD, providing some preliminary evidence that concrete academic skills may be better supported by interactive apps than more-complex social-emotional skills. Findings were mixed in studies in which researchers examined children’s short-term learning from interactive touchscreen media, particularly for toddlers. With the results of this review, we highlight both the potential for interactive apps...
to play a useful role in promoting some types of early skills as well as the substantial gaps in knowledge that still exist regarding their use. Continued research in this area will be critical to inform the debate around young children’s screen time, as clinicians and researchers try to strike a balance between taking advantage of the potential benefits of new technology, while encouraging limits in screen time.

ACKNOWLEDGMENTS

We thank David H. Arnold, PhD, for his helpful feedback on a draft of this article.

ABBREVIATIONS

AAP: American Academy of Pediatrics
App: application
ASD: autism spectrum disorder
LA: language arts
RCT: randomized controlled trial
3D: three-dimensional

REFERENCES


18. Fletcher J. What is heterogeneity and is it important? BMJ. 2007;334(7584):94–96


22. Randolph J. Free-marginal multirater kappa: an alternative to Fleiss’ fixed-
33. Mattoon C, Bates A, Shifflet R, Latham N, Ennis S. Examining computational skills in prekindergarteners: the effects of
34. Ginsburg H, Baroody A. Test of Early Mathematics Ability. Austin, TX: PRO-ED; 2003
58. Kwok K, Ghrear S, Li V, Haddock T, Coleman P, Birch SAJ. Children can learn new facts equally well from interactive media versus face to face instruction. Front Psychol. 2016;7:1603
61. NWEA. Children’s Progress Academic Assessment. 2009. Available at: https://www.nwea.org/cpaa/. Accessed November 12, 2019
62. Schroeder EL, Kirkorian HL. When seeing is better than doing: preschoolers’ transfer of STEM skills using touchscreen games. Front Psychol. 2016;7:1377
66. Tarasuk J, Demaria A, Kaufman J. Transfer of problem solving skills from touchscreen to 3D model by 3- to 6-year-olds. Front Psychol. 2017;8:1586