Title

Permalink
https://escholarship.org/uc/item/2936c7dj

Journal
InterActions: UCLA Journal of Education and Information Studies, 13(2)

ISSN
1548-3320

Author
Cirell, Anna Montana

Publication Date
2017

Peer reviewed
Abstract

This critical synthesis re-evaluates a widely cited sample of peer-reviewed empirical journal articles supporting electronic storybooks, or e-storybooks, as a promising medium for literacy learning among preschoolers of low socio-economic status (SES). From a critical re-assessment of these experimental studies across quality variables indicative of good research design (e.g., participant description, treatment detail, treatment fidelity, operationalized measures, measure reliability, internal and external validity, and clarity of causal inference), synthesis results indicate that all studies contained methodological weaknesses in one of more of these areas. The most serious flaws observed were failure to utilize proven instrumentation, inadequately described samples, small samples, and non-uniform treatment conditions. Evidence suggests that higher quality studies are needed before conclusive statements can be made on the efficacy of e-storybooks for vulnerable populations. Implications and recommendations for future e-storybook research are discussed.

*Keywords:* electronic storybook, emergent literacy skills, low SES, critical synthesis

Children from low socio-economic status (SES) enter kindergarten significantly behind their higher income peers in reading abilities, and despite government remediation this problematic achievement gap remains fixed throughout later grades (Reardon, 2013). By age three, a 30-million-word gap exists between U.S. children of low and high SES (Hart & Risely, 2003). Entering first grade, the typical middle-class child has accumulated 1,000-1,700 hours of storybook reading time, while the representative child of low SES averages just 25 hours of reading time\(^1\) (Hoff, 2013; Rowe, 2008). By the time students reach fourth grade, only 46% of children of low SES read at or above the basic level compared to 77% of their middle-class peers according to the Department of Education’s 2014 Nation’s Report Card. Given this, America’s current literacy crisis begins in early childhood (Neuman, 2008).

Reading abilities are highly dependent upon home literacy variables (Goncu & Gauvain, 2012; McIntyre, 2010), and the literature has found that children from low SES have less access to home literacy and verbal dialogue with caregivers than higher income peers (Hernandez, 2011; Lareau, 2011; Neuman & Celano, 2001). Recent United States Census data show a correlation between

\(^1\) While no exact definition of low, middle, and high SES exists, SES considers income, education, and occupation.
increased parent-child home literacy activities from 1993 to 2007 and an upsurge in preschooler’s “school readiness” (U. S. Report of Children with Parental Involvement in Home Literacy Activities, 2012). However, according to this data, the largest jump in home literacy practices and resulting school readiness follows from those families above the poverty threshold with higher parent education levels. Developing timely and responsive early reading interventions will have a significant impact on the reading abilities and future academic achievement of America’s most vulnerable student populations (Hernandez, 2011; Neuman & Celano, 2012). Furthermore, as current worldwide social and economic circumstances continue to decline and children of low SES entering American public schools becomes the new standard, few would question the urgency of developing effective early literacy interventions to stem the cyclical tide of disadvantage (Neuman, 2008; Neuman & Celano, 2012; Reardon, 2013).

Speaking to the state of literacy and social class worldwide, where wealth gaps in developed countries are the widest in decades (Carter & Reardon, 2014), our next generation’s ever-changing and technologically-mediated world will collectively depend on the 21st century literacy skills that best enable interaction among societies.

**Purpose of Critical Synthesis**

In conjunction with marketing forces and government initiatives, literacy researchers have advanced the claim that new digital technologies can support
literacy learning in children of low SES to remedy the achievement gap between socioeconomic classes (Benders, 2011; Fish et al., 2008; Li, Atkins & Stanton, 2006). In particular, electronic storybooks, or e-storybooks, are proposed as a promising medium for literacy learning among children of low SES. In their minimal format, these e-storybooks complement written text with oral narration, while more advanced models capitalize on multimedia interactive features (e.g., background music, animations, sound effects) and/or hypermedia via embedded activities (Zucker et al., 2009). Despite the growing popularity of e-storybooks, the literature lacks a careful critical synthesis to evaluate the evidence base from the most widely-cited experimental studies (Bus & Neuman, 2014).

Herein, my purpose is to re-evaluate studies through a critical approach (i.e., appraising strengths and weakness of the experimental evidence) to locate and describe themes across studies (Sandelowski & Barroso, 2006). Given this, I utilize a critical synthesis approach over meta-analytic methods that simply calculate an average effect size. Such an integrative and timely critical synthesis sheds light on the theoretical frameworks, methodological approaches, and implications of the research over the past 15 years, illuminates gaps and inconsistencies, and aligns policy recommendations with evidence-based technology tools (Biancarosa & Griffiths, 2012).

**Background Information**

This critical synthesis paper is divided into five sections. The first section
reviews the relevant scholarship guiding this synthesis. With this critical background information established, I then state the need for a significant and critical contribution that evaluates the study claims on the affordances of e-storybooks over the adult reader in the second section. In the third section, I introduce the methods taken to critically synthesize the body of evidence and its ability to address a literacy crisis as complex and current as the one in schools today. In the fourth and fifth section, the results and discussion are presented.

**Emergent literacy skills and dialogic readings**

Emergent literacy refers to the developmental antecedents of formal reading, covering literacy-related behaviors that occur prior to formal schooling (Lonigan, Dickinson, & Neuman, 2006; Whitehurst & Lonigan, 2001). Two interdependent domains of literacy key to children’s emergent literacy are outside-in and inside-out skills (Van Den Broek, Kendeou & White, 2014; Whitehurst & Lonigan, 2001). The outside-in domain refers to information that exists outside printed text that supports print understanding (e.g., vocabulary, semantic and conceptual knowledge, contextual understanding, and story schemas), while inside-out domains refer to information within text that aids in decoding print into sound and sound into language, such as letter knowledge or phonological awareness (Van Den Broek, Kendeou & White, 2014). Beyond isolating skills, the ability to integrate information from both domains contributes to successful reading (Van Den Broek, Kendeou & White, 2014; Whitehurst & Lonigan, 2001).
Dialogic readings meet the Institute of Education Sciences (IES) What Works Clearinghouse (WWC) evidence criterion for best practices and stand out as the “gold standard” for early literacy learning (IES, 2007). Likewise, educators have long heralded dialogic reading as “the single most important activity for developing the knowledge required for eventual success in reading” (Commission on Reading, 1985, p. 23; see also Pearson & Hiebert, 2010). When compared against teacher-child readings at school, parent-child dialogic readings at home have shown the most powerful gains in young children’s emergent literacy skills (Arnold & Whitehurst, 1994; Lonigan, Shanahan & Cunningham, 2008).

Briefly described, dialogic reading practices gradually shift the storytelling role from the adult reader to the child through various dialogue-promoting scaffolding techniques (e.g., open-ended questions, repetition, modeling) (Lonigan, Shanahan & Cunningham, 2008). Scaffolding refers to adults’ guidance as they structure the activity, build bridges from what children currently know to their goal potential, and transfer the responsibility of independent problem solving when children are ready (Goncu & Gauvain, 2012). Because family is the first context for young children’s learning, scholars believe parents can fulfill a pivotal role in modeling appropriate behaviors and calibrating responsive support to improve their children’s learning (Wenger, 1998).

**Socioeconomic status and home literacy variables**
Many parents of low SES want to support their child’s literacy learning, but may have limited access to appropriate children’s literature (Hernandez, 2011; Neuman & Celano, 2001, 2012). Over the course of ten years, Neuman and Celano documented drastic differences in both print access and digital literacy variables between rich and poor communities in urban Philadelphia, dispelling the myth of equal educational opportunities for children and their families. Wealthy neighborhoods in their study had abundant public spaces suitable for reading, plentiful bookstores selling thousands of types of superior children’s literature, and quality digital learning opportunities (i.e., public libraries, afterschool programs) with more computers and better-trained technology staff. In comparison, the poorer areas, despite a higher population of young children, sold just 55 types of low-quality children’s literature from dollar stores and pharmacies, provided no public spaces for reading, hired day care staff less versed in conventional standards of early literacy development, and granted less access to quality digital learning opportunities (Neuman & Celano, 2001, 2012).

Other scholars have found that many parents of low SES may not know how to engage their children in quality shared reading activities (Ordoñez-Jasis, 2010; Oritz & Ordonez-Jasis, 2005). Parents of low SES may lack experience with scaffolding a dialogic literacy event to structure question-asking and stimulate the child’s ability to conceptualize, synthesize, reason, and infer (Gee, 2001). When reading, middle-class parents may be better positioned to enhance
their child’s higher-order thinking and language skills to parallel school discourse, or middle-class teacher–child dialogues (Gee, 2012).

Strong correlations exist between literacy skills and vocabulary knowledge (Mol & Bus, 2011); many report home literacy environments in low SES families to be less rich in vocabulary than higher SES families (Hart & Risely, 2003; Hoff, 2013; Neuman & Wright, 2013; Rowe, 2008). Because some mothers of low SES rarely use vocabulary beyond the 3,000 most common words, their children enter kindergarten with significantly fewer words at their disposal than their peers from higher SES families (Hoff, 2013). A six-year longitudinal study of 42 families of different SES found that by preschool, a 30-million-word gap exists between the cumulative language experiences of high- and low-SES children (Hart & Risely, 2003; Neuman & Wright, 2013). This vocabulary disparity only magnifies as children of low SES learn 750 versus 3,000 new words per year (Hoff, 2013).

Thus, children of low SES may not have equal opportunities for rich early literacy instruction and may not be equally equipped to profit from education. In the subsequent curriculum shift from “learning to read” to “reading to learn,” language skills and background knowledge may become limiting factors and place children of low SES at risk for special education referral (Artiles & Kozleski, 2010). Additionally, they may be less likely to enjoy reading, resulting in decreased motivation and fewer exposures to print, which may magnify literacy gaps over time (Neuman & Celano, 2001, 2012).
E-stroybook potential for children of low SES

Several studies have examined the role that e-storybooks may play in providing rich language and literacy learning opportunities in the lives of children from low SES (Korat & Shamir, 2007; Shamir, 2009; Van Dijken, Bus & de Jong, 2011). All e-storybooks provide basic read aloud features that digitally narrate the story at a child-led pace “in loco parentis,” or in the place of the parent. This can be of benefit to parents of low SES who, for reasons such as lack of time, fatigue, or their own level of literacy, may be less inclined to share books with their children or to initiate other literacy promoting activities (Van Dijken, Bus & de Jong, 2011). Additionally, the availability of e-storybooks from local public libraries allows families to download many titles without having to purchase the book (Van Dijken, Bus, & de Jong, 2011). Because e-storybooks commonly provide audio and text in multiple languages, such as Spanish, German, Dutch, French, Chinese, and Japanese (Tsai & Huang, 2014), they have international appeal and are engaging to a computer-literate global generation of learners (Park & Kim, 2015). Further, the research on the efficacy of e-storybooks comes from all corners of the globe and quickly dispels the myth that use of e-storybooks is limited to U.S. English speakers (Guernsey & Levine, 2015).

Children’s storybooks have a long history of technological enhancements. The earliest versions, called “sound books,” housed a simple sound effects bar along the book’s edge, where certain pictures in the story corresponded to specific
side bar sound buttons (Slimani, 2009). Later, picture books sold as “books on tape,” bundled with audio cassettes or CDs narrating the text. As computers became commonplace in schools around the late 90s, CD-ROM versions, or talking books, brought the text alive on screen as the viewer clicked through words and pictures in the storybooks (Robb, 2010). From this, multimedia innovations embedded highly interactive learning activities (e.g., animations, dictionaries, music, and sound effects) into the story (Smeets & Bus, 2013), allowing interactive e-storybooks to evolve. Still later versions of e-storybooks pushed interactivity beyond graphics and sound, enabling the learner to go beyond the story itself to “click and explore” through games or puzzles (NCIP, 2008). Living books, one type of interactive e-storybook, developed further and were marked by their inclusion of “play-through” videos, which are often live-action and differ from multimedia in that they link the text to non-linear hypermedia (Park & Kim, 2015).

**Multimodal learning implications on the digital divide**

Dialogic readings of text, whether taking place around e-storybooks or print, segue from the information processing view that understands reading as simply a “coldly cognitive” undertaking (Alexander, 2010). In addition to cognitive processing, the act of reading is multimodal and requires the orchestration of many elements, including motivational, neurophysiological, and socio-contextual strands (Alexander, 2010). Similarly, multimodal approaches to
literacy encourage young children to produce and interpret a variety of text, drawing on the multiple modes of expression and communication (e.g., visual, bodily, auditory), wherein each compensates for the other’s limitations (Kress, 2000). Likewise, the new Common Core State Standards emphasizes the reading of all types of media and digital texts (Strickland, 2013).

Calvert, Rideout, Woolard, Barr, and Strouse (2005) reported that children as young as two years old use the computer with the help of their parent, and three-year-olds already manage some computer functions on their own, such as controlling the handheld mouse. The researchers also found that young children’s daily hours of computer usage will continue to increase until the child enters kindergarten. However, the mere presence of new technologies will not lead to equally meaningful use. Though Rideout, Foehr, and Roberts’s (2010) study of young people’s home media ecology discovered no difference in the amount of media used between low- or high-SES families, differences in the type of media consumed persists.

The digital divide reflects the issue of low-educated families having access to computer programs but gaining fewer benefits derived from access (Robinson, 2009). Because technological competencies are “closely linked to students’ economic, cultural and social capital” (OECD, 2010, p. 3), a current divide exists among those who know how to use technology and those who do not. Therefore, attention must be paid to not only provide access to high quality
technologies, but also extend supportive learning opportunities for effective
technology use (Neuman & Celano, 2012; Toyama, 2015).

Young learners need experience with multiple communicative modes in a
variety of formats to increase their ability to navigate the changing landscape of
semiotic material and to understand the ever-more complex texts that shape their
literacies (Kress, 2000). Hence, there is a need to develop multimodal reading
skills in vulnerable populations of students who lack access and scaffolding to
many of the new digital technology tools (Kress & van Leeuwen, 2001).
Multimodal technology tools, such as e-storybooks, may provide an inexpensive
and accessible means of enhancing emergent literacy skills and meaning making
repertoires to prevent early grade retention and long-term academic struggles in
vulnerable student populations (de Jong & Bus, 2004).

Summary of Findings and Gaps from the Literature

A growing body of work in best practices calls for timely detection and
intervention to reduce the literacy gap early on and emphasizes the use of
technology for improving children’s emergent literacy skills (National Reading
Panel, 2000). Thus, educators are working to provide high-quality evidence-based
emergent literacy technology tools for the critical developmental years of early
childhood, especially for those children of low SES most at risk (Bus, Takacs &
Kegel, 2015). Given the extensive marketing of e-storybooks, it is not surprising
that within the last decade, e-storybook popularity has grown significantly among
parents and teachers of young children. A national survey targeting American children ranging from birth to six years of age, revealed that 20% of three- to four-year-olds and 10% of five- to six-year-olds used e-storybooks daily and on average, spent 36 minutes and 47 minutes respectively interacting with them (Vandewater et al., 2007). Additionally, over 60% of parents indicated in the same survey that educational interactive technologies, such as e-storybooks, were “very important.”

Despite the growing use of e-storybooks among young children, the extent to which e-storybooks promote emergent literacy skills in vulnerable populations is highly variable among research results (Van Daal & Sandvik, 2013). Thus, it is important for parents, teachers, and reading specialists to be informed consumers. Thoroughly reviewing the literature and re-evaluating its most widely-cited experimental studies will help identify the strongest evidence to support not only how to access high quality, developmentally appropriate e-storybooks, but also how to use e-storybooks in an optimal way.

Lastly, in furthering connections from research to policy, it is imperative that policymakers and the general public become aware of the most valid and conclusive evidence from which to base recommendations for early reading intervention. Most acknowledge the importance of the early years of schooling and view funding for early literacy learning tools as a strategic investment—especially in low SES populations (Biancarosa & Griffiths, 2012; Strickland,
2013). Despite budget cuts, states continue to support and even increase funding to close early literacy gaps, and as researchers, we must thoroughly critique the established literature to provide the best evidence for the just allocation of funds.

**Methods**

As indicated previously, I took a critical approach to a review of research, engaging in an evaluation of research reports as persuasive texts. Given my narrow selection criteria, locating the articles was an extensive process. With the aid of a coding sheet, I critically analyzed each of my nine select articles across Troia’s (1999) select quality criteria to consider other possible elements that the authors may have concealed or overlooked.

**Search Strategy**

I first searched computerized catalogued databases, including ERIC, JSTOR, and Google Scholar, to locate studies published between 1995 to the present. I established this broad 18-year range to explore whether the considerable innovations in e-storybook format and delivery equated progress in improving children’s literacy skills. Various search terms and Boolean operators were initially used: (Electronic books) OR (storybooks) AND (print books) OR (literacy) OR (emergent language) AND (young children) OR (preschoolers) AND (low SES) OR (low-income). Additional descriptors or root forms of those descriptors (e-book, read*, comprehend*, technology, tech*, poverty) were used in various combinations to capture the greatest possible number of articles.
The initial search criteria yielded 969 matches, of which 22 studies met the criteria for full review. I searched the reference lists of relevant articles and/or seminal studies to find subsequent articles of interest. Beyond this, I conducted a hand search of the six top tiered journals in the fields of literacy and technology from 1995 to present. Journals examined in this hand search were the following: *Journal of Educational Psychology, Computers and Education, Learning and Instruction, Journal of Computer Assisted Learning, Journal of Research in Reading,* and *Reading Research Quarterly.* The additional bibliography and hand searches yielded 14 additional studies. A total of 36 articles were reviewed in full, yet only nine experimental studies remained and are the focus of this subsequent critical synthesis.

These selected nine studies met important research design quality criteria. Studies were accepted only when their experimental research designs clearly indicated the successful random assignment of child participants into acceptable e-storybook intervention. Experimental design was a primary inclusion criterion, because random assignment of participants establishes internal validity. With this level of experimental control, one can apply rules governing sampling variation to make statistical inferences in confidence that the magnitude of difference is from experimental manipulation and not from chance or other extraneous factors biasing study results (Huck & Cormier, 1996). Lastly, acceptable intervention criteria consisted of any type of e-storybook treatment, including talking books,
interactive books, living books, or a combination of these.

Chosen quantitative studies were all peer-reviewed and published in top journals in the field of technology and literacy. Meta-analyses, government documents, independent research lab reports, policy briefs, and corporate-funded and/or self-published e-storybook studies were omitted from this synthesis. Observational or qualitative studies were also excluded; researchers must have quantitatively analyzed the empirical effectiveness of e-storybooks on emergent literacy skills.

Finally, all selected studies had to also adhere to specific demographic and literacy learning criteria. Only studies measuring differences across literacy learning variables were included. The three- to four-year-old age range was selected because it represents the most critical period for developing reading competence: the National Institute for Literacy (2008) found that alphabet knowledge, phonological awareness, and oral language of emergent readers (ages three to four) are, in fact, among the best predictors of later success in literacy achievement. Though the definition of low SES varied across the corpus of international studies, if authors justifiably indicated their sample was representative of low SES, they were selected. The disclosure of at least two or more factors from total SES factors (e.g., parent’s occupation, education, family size or income) was mandatory.

**Coding**
A coding sheet helped to systematically organize information from each of the nine studies and cross-compare selected features. The coding sheet was based on elements specified in the What Works Clearinghouse Design and Implementation Assessment Device (Institute of Education Sciences, 2003), a widely-cited document used to evaluate the quality of studies. In addition to this document, other high-quality coding sheets, such as one developed by Mamlin, Harris and Case (2001), guided the development of code categories.

The final coding sheet included key information regarding participants (e.g., number, gender, familiarity with technology, SES measure used, attrition rates), study design (e.g., number of conditions, assignment to condition, fidelity to implementation, rigorous instrumentation), specifications about conditions (e.g., intervention, comparison, treatment duration, site of study), clarity of causal inference (e.g., plausibility of intervention contaminants, measurement inconsistencies), and reported findings. Once this first coding stage had been completed, I summarized and aggregated the studies to assess and compare each study’s methodological rigor as per Troia’s (1999) select quality criteria, such as internal and external validity for randomized design, control group intervention, sufficient participant description, sufficient treatment detail, treatment fidelity ensured, operationalized measure, measure reliability reported, and valid effect size reported.

Results
Though a sizable body of educational research literature linking emergent literacy outcomes to e-storybook use among children of low SES exists, only nine select experimental studies were within the scope of this synthesis. According to Troia's (1999) eight criteria of internal and external validity for establishing methodological rigor, seven out of nine articles demonstrated strong research designs, and two out of the nine demonstrated only moderate strength (see Table 1). Hence, only seven studies exhibited qualities considered essential for asserting causality or generalizability of the findings to other populations. Results are further presented across four dominant trends in e-storybook studies, as their multimodal enhancements and digital formats advance from the early 1997 versions to the more recent 2009 versions. Each trend evaluates a varying combination of multimedia-enhanced e-storybook formats (e.g., talking books, interactive books, living books, and e-storybook co-viewing) against a comparison condition (e.g., adult reading of print book, “business as usual” control, static e-storybook narration, etc.). Within each trend, the participant selection, design instrumentation, and treatment conditions are critically assessed.

<table>
<thead>
<tr>
<th>Study title</th>
<th>Randomization</th>
<th>Control group</th>
<th>Sufficient participant description</th>
<th>Sufficient treatment detail</th>
<th>Treatment fidelity ensured</th>
<th>Operation alized measure</th>
<th>Measure reliability reported</th>
<th>Effect size reported</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johnston (1997)</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>moderate (5)</td>
</tr>
<tr>
<td>Study</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Talley et al. (1997)</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>de Jong &amp; Bus (2002)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>de Jong &amp; Bus (2004)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Verhallen et al. (2006)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Korat &amp; Shamir (2007)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Korat &amp; Shamir (2008)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Shamir (2009)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Korat, Segal-Drori &amp; Klein (2009)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

Note. Adapted from Troia’s (1999) eight criteria of internal and external validity.

**Talking Book Interventions (1997)**

Two early experimental studies from 1997 compared a “business as usual” control group with talking e-storybooks that contained minimal multimedia interactivity (Johnson, 1997; Talley, Lancy & Lee, 1997). Delivered from a classroom computer, both studies’ talking books differed little from typical print books. Individual students listened to the narrated story with headphones via a read-only track, as they clicked to the next screen with a mouse. Both studies found a significant improvement in the treatment groups’ emergent reading skills.
over that of the control group. Table 2 indicates strengths and weaknesses across these studies.
Table 2: Strengths and Weaknesses of Talking Book Trend

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>--Intervention delivered for adequate duration (average of 7.5 weeks)</td>
<td>--Failed to describe source of training given to facilitators of IV</td>
</tr>
<tr>
<td></td>
<td>--Operationalized emergent literacy outcomes across standardized measures of print awareness, print concepts, and phonological awareness</td>
<td>--The reliability and/or construct validity of DV measures not reported.</td>
</tr>
<tr>
<td></td>
<td>--Talley et al. provided sufficient description of participants.</td>
<td>--Johnston provided no detail of participants.</td>
</tr>
<tr>
<td></td>
<td>--Successfully randomized participants &amp; provided informed consent</td>
<td>--None provided screening measures prior to randomization to control for ceiling effect or learning disability</td>
</tr>
</tbody>
</table>

Interactive Book Interventions (2002-2006)

Three experimental studies from 2002-2006 used an interactive e-storybook format that enhanced the story with interactive multimedia, dynamic visuals, and “hotspots” (i.e., mouse-click activated embedded links that dramatize the story with sound effects, music, animations, interactive games, and puzzles).

All studies’ e-storybook treatments included an unrestricted treatment e-storybook (i.e., a read-and-play e-storybook) as well as other comparison conditions, such as a restricted or static narration-only track (i.e., a read-only e-storybook), a control condition and/or a print storybook dialogic reading. Studies evaluated the read-only and read-and-play treatments against two different groups, the no-treatment control and/or a print storybook condition (De Jong & Bus, 2002, 2004; Verhallen, Bus, & de Jong, 2006).
Though all three studies demonstrated adequate internal and external experimental control and reported effect sizes, their results were mixed. De Jong and Bus, in their 2002 study, found the read-and-play e-storybook hotspots to be incongruent to the story, provide no benefit to story comprehension, and result in cognitive overload, while the other two studies found no difference between the read-and-play e-storybook and the read-only e-storybook on the children’s emergent literacy skills. Relatedly, studies found that children took almost four times longer to read the book in the unrestricted read-and-play treatment with the choice to explore hotspots and move back-and-forth across the plot. When studies included the print-based comparison group (De Jong & Bus, 2002, 2004), these conditions reported the highest literacy gains. Table 3 indicates strengths and weaknesses of this trend.

Table 3: Strengths and Weaknesses of Interactive Book Trend

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent variable (IV)</td>
<td>--Provided treatment fidelity checks and sufficient intervention detail</td>
<td>--Failed to indicate length of intervention (possibly only 2.5 weeks long)</td>
</tr>
<tr>
<td></td>
<td>--Operationalized literacy measures</td>
<td>--No information given on literacy learning conditions of control group</td>
</tr>
<tr>
<td></td>
<td>--Reported inter-rater reliability</td>
<td>--Utilized researcher-developed non-standard DV measures</td>
</tr>
<tr>
<td>Dependent variable (DV)</td>
<td>--Provided good description of participants</td>
<td>--Lacked discussion of training of the independent coders for IV measures</td>
</tr>
<tr>
<td>Participant selection</td>
<td></td>
<td>--Small sample size (averaging 42)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--Failed to indicate participants’ familiarity with technology</td>
</tr>
</tbody>
</table>

In these two studies, Korat and Shamir (2007; 2008) evaluated one living e-storybook they had co-developed specific to the needs of their student population. The living books in these studies contained a rich film-like representation of the story and provided no interactivity with these “living” illustrations. Opportunities to obtain word pronunciations or definitions offer a minimal level of interactivity (Korat & Shamir, 2007, 2008). Read-with-dictionary mode supplemented an oral reading of the text with automatic definitions of difficult words. The read-with-minimal-play mode offered opportunities for interaction with the story (e.g., activating characters or objects on the page, etc.). However, the interactions in these modes could only be activated after the complete narration of the screen’s text.

Korat and Shamir (2007) first involved children from varying SES levels and evaluated two e-storybook conditions (e.g., read-with-dictionary track and read-with-minimal-play track) against an adult’s dialogic reading of a print storybook. Then in 2008, they used gain scores to test the effectiveness of the e-storybook’s three pre-programmed modes (e.g., read-only track, read-with-dictionary track, and read-with-minimal-play track) and excluded the print condition. In 2007, they found that the minimally interactive e-storybook modes improved children’s emergent literacy skills over the static read-only mode, but it did not significantly improve reading gains when compared against the adult
reader condition. Both studies took randomized block designs, blocking on the variable of SES, to find that the children from low SES made greater gains than the students from the middle SES levels. Studies strengthened their results by reporting effect sizes. Table 4 indicates strengths and weaknesses of these studies.

**Table 4: Strengths and Weaknesses of Living Book Trend**

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Independent variable (IV)</strong></td>
<td>--Interventions pre-programmed technology restrictions increased treatment fidelity</td>
<td>--Intervention lasted only 3 weeks</td>
</tr>
<tr>
<td></td>
<td>--Various comparison treatments adequately described</td>
<td></td>
</tr>
<tr>
<td></td>
<td>--Did not utilize “business as usual” control group</td>
<td></td>
</tr>
<tr>
<td><strong>Dependent variable (DV)</strong></td>
<td>--Operationalized their emergent literacy measures</td>
<td>--Pretest didn’t include reading comprehensions</td>
</tr>
<tr>
<td></td>
<td>--Reported inter-rater reliability</td>
<td>--Utilized researcher-developed non-standard DV measures</td>
</tr>
<tr>
<td></td>
<td>--Provided sufficient description of participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>--Sufficient sample size (averaging 138)</td>
<td>--Using 8 different kindergarten sites may confound treatment fidelity</td>
</tr>
<tr>
<td><strong>Participant selection</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**E-Storybook Co-Viewing (2009)**

The studies in this trend used the same researcher-developed living book from the previous 2007-2008 studies, but each study incorporated a collaborative component to their e-storybook treatments (Korat, Segal-Drori, & Klien, 2009; Shamir, 2009). Because many living e-storybooks include features designed to replicate what an adult reader may do in a print reading, the idea of co-viewing takes on a different meaning when involving living e-storybooks. To evaluate the efficacy of co-viewing with e-storybooks, Korat, Segal-Drori, and Klien (2009) randomly assigned kindergardeners from low SES into one of four groups: read...
the e-storybook independently without adult instruction (EB), read the e-
storybook with adult instruction (EBI), read the printed book with adult
instruction (PBI), and a control group which received the regular kindergarten
program, or 20 minutes of basic emergent literacy instruction. For both studies,
the three co-viewing treatment groups were taken out of their kindergarten class
during reading time for multiple reading sessions that lasted no longer than 20
minutes over an average of 3 weeks. In both studies’ treatment conditions, pairs of
students or adult-child dyads were given instructions on how to view the e-
storybook completely and offered the choice in viewing modes (e.g., read-with-
dictionary track or read-with-minimal-play track).

Shamir (2009) explored the effectiveness of students’ collaborative
learning around e-storybooks and randomly assigned kindergarteners of low SES
to either the treatment condition of joint learning around e-storybook, or the
control group of kindergarten classroom instruction only. All studies reported
effect sizes and outcome measures to demonstrate that the joint learning around
either the e-storybook or the print book improved the children’s measures of
emergent literacy over the control or read e-storybook alone conditions. Table 5
indicates strengths and weaknesses across this trend.
### Table: Strengths and Weaknesses of Co-Viewing Trend

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent variable (IV)</td>
<td>-- Various comparison treatments adequately described</td>
<td>-- No information given on literacy learning conditions of control group</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-- Irregularities between student pairs may have confounded results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-- Non-standardized conditions in Korat, Segal-Drori &amp; Klein study because implemented 2 different books</td>
</tr>
<tr>
<td>Dependent variable (DV)</td>
<td>-- Operationalized their emergent literacy measures</td>
<td>-- Utilized researcher-developed non-standard DV measures</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-- DV measures had poor reliability measures</td>
</tr>
<tr>
<td>Participant selection</td>
<td>-- Provided sufficient description of participants</td>
<td>-- Using 4 different kindergarten sites may confound treatment fidelity</td>
</tr>
<tr>
<td></td>
<td>-- Sufficient sample size (averaging 112)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-- Included advanced screening measures prior to randomization to ensure participants met inclusion criteria</td>
<td></td>
</tr>
</tbody>
</table>

## Conclusions Across E-Storybook Trends

In sum, nine peer-reviewed experimental studies met the inclusion criteria for this critical synthesis. Though the majority of the e-storybook experimental studies had strong designs, there were many conceptual, psychometric, and analytical issues observed within the studies. Methodological flaws within all studies and the fact that no e-storybook treatment showed more meaningful advantage over adult-led print readings critically questions the claim that e-storybooks can effectively support literacy learning in children of low SES to remedy the achievement gap between socioeconomic classes. Outlining the
most serious flaws (e.g., problematic reliability of non-standard instrumentation, the limited detail and small number of child participants, the insufficient consideration of an equal control condition, variations/lack of fidelity in e-storybook treatments), the following paragraphs reveal how the promise of e-storybook use among vulnerable populations of readers has yet to be empirically grounded.

**Instrumentation: Non-standardized measures**

Across all widely-cited studies, the greatest concern lay in their non-standardized reading comprehension instrumentation, with no proven psychometric validity. To further complicate a synthesis of the available experimental research, each study employed a different non-established measure (e.g., not psychometrically validated and/or researcher-developed measures of emergent literacy skills). Using standardized measures would have controlled for instrumentation validity across test-retest reliability, internal consistency reliability, criterion-related validity, and construct validity of the measure. A low reliability (less than .7) complicates a tests’ consistent ability to measure what it is purporting to score, or its construct validity, and any effect size related to unreliable measures is thus attenuated (Henson, 2001). Similarly, modern psychometric testing theory places construct validity as the overarching concern of validity research, subsuming all other types of validity evidence (Cronbach & Meehl, 1955; Messick, 1995). Therefore, all inferences following from measures
with no established construct validity will fail to hold as empirical evidence (AERA, 2014).
Sampling: Missing details and small size

Most quantitative researchers study only samples from which they make inferences about the populations sampled (Gay et al., 2012). The international scope of the literature may have limited the extent of comparison. However, the inclusion of international studies was grounded on each study’s ability to address the shared global goal to use e-storybook technologies to remedy the seemingly pandemic literacy achievement gap between rich and poor children. Additionally, another such limitation across many of these studies is the authors’ neglect in accounting for children’s familiarity with technology or any diagnosis of certain learning disabilities; both of these variables could have influenced outcomes in various conditions. Screening measures prior to randomization would have controlled for any incidence of learning disability and/or minimized potential ceiling effects within the measure (i.e., when the participants’ pre-existing measures are too high and resulting variability in performance cannot be accurately captured by the instrument).

Across studies, sample sizes were generally small, which limits the accuracy of the analytical findings and subsequent effect size estimation (Shadish, Cook & Campbell, 2002). Implementing multiple e-storybook treatments at once (i.e., read-and-play, read-only) calls for a larger sample size to power the statistical analysis, as an underpowered study may result in false causal findings (Shadish, Cook, & Campbell, 2002). Calculated prior to estimate minimal sample
size, a power analysis would have justified small sample sizes and thus strengthened all studies’ experimental rigor (Cohen, 1992). Given this, more practical limitations justifying the use of small samples included the convenience and financial feasibility of using a single classroom as a research space, which lowered the possibility for disruption (Hackshaw, 2008). At the same time, including more classrooms, such as the studies of Shamir (2009) and Korat, Segal-Drori and Klien (2009), could have possibly overcome the challenges of large samples.

**Variations in e-storybook treatments**

Another limitation of most studies in this critical synthesis is the wide variation in the treatment interventions, which makes it difficult to generalize the results beyond the conditions and settings of the particular study’s intervention. Many studies also used the kindergarten classroom curriculum as a control group, without addressing the extent to which this “business as usual” classroom was exposed to comparable print or multimedia materials throughout the study. Inclusion of an equivalent treatment group and clarifying its nature and mode of delivery would have better controlled for any discrepant learning happening in either group and further minimized any “John Henry” effect (i.e., when the control group discovers their neglected status and takes pains to outperform the intervention group) (Gay, Mill, & Airasian, 2012).

Invariably, almost all interventions lasted no longer than 3 weeks, with the
average intervention length at about 3.5 weeks. Longer interventions could have minimized testing effects and helped to justify that the treatments’ posttest improvement was due to the intervention and not from information they retained from taking the pretest. Further, these short treatments hardly fulfill the recommended duration to which e-storybooks are to be used by children of low SES. Thus, the ecological validity of many of the studies is compromised, and it becomes difficult to generalize these conclusions beyond the duration of the e-storybook use. Likewise, concerning the ecological legitimacy of the 2009 co-viewing studies, Chiong, Ree, Takeuchi and Erickson (2012) found that most interactive e-storybooks are not designed for shared readings with adults.

A significant number of the studies did not ensure treatment fidelity or provide sufficient detail surrounding frequencies of repeated readings, student/facilitator ratio, special qualifications of facilitators, or their familiarity with the student population. Stating these precautions taken to control treatment fidelity would have improved the internal validity of the findings. In addition, it would have removed chance of error through variation in e-storybook treatment intensity, and improved opportunities for replication in other studies and application to classroom settings. Though the fidelity to treatment was somewhat strengthened with pre-programed restrictions placed on e-storybook software modes, disclosure of fidelity checks would have better supported study claims.
Limitations

The basic aim of this critical synthesis was to carefully re-evaluate the evidence base promoting e-storybooks as a promising medium for literacy learning among children of low SES. Practicing caution with seemingly conclusive statistical facts and figures would reduce the chances of vulnerable populations falling further behind. However, challenges confronted lay in the narrow scope of quantitative randomized experimental design and its limited sample size of nine articles.

Implications for Future Research

Given the popularization of all forms of digital media, the timing for this paper is auspicious. Yet, no recommendations regarding increased literacy skills for children of low SES can be derived from the extant e-storybook literature without considering the above study limitations of non-standardized instrumentation, narrow demographic, and non-uniform e-storybook treatments. Lack of scientific rigor has reduced the value of many of these studies. Beyond these issues, adult mediation during e-storybook use needs to be better addressed, as adult assistance has been associated with improved outcomes for emergent literacy in all populations of readers.

To practice caution when working with vulnerable populations of readers, future educators and policymakers would do well to understand their place as choice architects, nudging older children to “read up” or potentially choosing for
young children who may prefer exciting multimedia programs bereft of literacy learning opportunities (Bus & Neuman, 2009). Similarly, in early stages of book reading, children are drawn to adult interaction over the printed text, but they rely on adults to narrow the gap between their social world and the imaginary text-based one (Smeets & Bus, 2013). Without stressing the need for a strong adult-child dialogic model of shared readings built up over time, the attachment to the electronic book may never form (Smeets & Bus, 2013).

In addition, with the introduction of newer portable devices, such as tablet e-readers and iPads, the entire context of reading and learning to read has been transformed. Neuman and Strom (2015) have begun to disseminate the effects of print knowledge, phonological awareness, and letter sounds when using these more mobile, tactile e-storybook devices in seven low-SES Head Start classrooms. Given the immense gap in the limited body of experimental literature on e-storybook use among children of low SES from 2009 to present, the field looks forward to new evidence. In this ever-changing field of digital technology, the generalizability of platform-specific literacy gains decays; all findings become as obsolete or outmoded as the discontinued-storybook to which they are linked. Until then, this exhaustive critical synthesis provides only minimal evidence that e-storybooks can be used to support emergent literacy skills in children of low SES.
References


Benders, D. S. (2011). The Influence of Technology on the Academic Achievement of Students with Low Socioeconomic Status. *Available at SSRN 1968560*.


2 Asterisks (*) next to references signify their inclusion within the sample of articles chosen for this critical synthesis.


