American Psychological Association

Journal of Educational Psychology

© 2018 American Psychological Association
0022-0663/19/$12.00

2019, Vol. 111, No. 1, 32–44
http://dx.doi.org/10.1037/edu0000278

Learning Vocabulary From Educational Media: The Role of Pedagogical Supports for Low-Income Preschoolers

Susan B. Neuman and Kevin M. Wong
New York University

Rachel Flynn
Northwestern University

Tanya Kaefer
Lakehead University

This article reports on two studies designed to examine the landscape of online streamed videos, and the features that may support vocabulary learning for low-income preschoolers. In Study 1, we report on a content analysis of 100 top language- and literacy-focused educational media programs streamed from five streaming platforms. Randomly selecting two episodes from each program, we identified the prevalence of vocabulary opportunities, and the pedagogical supports—techniques or features in these media that are designed to orient children to specific vocabulary words. In over the 2,000 scenes coded, we identified two overriding categories of supports: ostensive cues, designed to provide definitional information to children; and attention-directing cues, designed to signal children's attention to a target word. In Study 2, we use eye-tracking technology to examine which of these pedagogical supports might predict children's ability to identify program-specific vocabulary. Results indicated that although ostensive cues predicted overall attention to scenes, attention-directing cues were most effective in directing children to target words and their subsequent word identification. Children with higher language scores were more likely to use these cues to their advantage than their lower language peers. These results may have important implications for designing digital media to enhance children's opportunity to learn vocabulary.

Educational Impact and Implications Statement

Screen media use on mobile devices for children ages 8 and under has risen rapidly in recent years to an average of 48-minutes day. Recognizing its potential to engage children's interest, this study examines the current landscape of educational media programs for children's word learning and vocabulary development. Our study shows both the prevalence and wide variation of word learning opportunities and highlights the production cues that appeared to differentially elicit children's attention to words and media content. These results could support a more intentional approach to media design to enhance children's opportunity to learn vocabulary.

Keywords: educational media, vocabulary, early childhood, early literacy

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The "Digital Wild West" might be a most apt metaphor for the burgeoning educational media marketplace in early literacy development for young children (Guernsey, Levine, Chiong, & Severns, 2012). Characterized by often-confusing claims about the educational benefits of screen media, its quality, and developmentally appropriateness for young children, parents and educators have had to navigate this relatively new terrain on their own. Scanning the marketplace across various platforms and top apps—researchers at the Joan Ganz Cooney Center found a serious mismatch in what developers were producing (e.g., featured e-books, websites, apps), and what young children were likely to need. Over 70% of the apps reviewed, for example, featured competitive or testing-based activities in games, puzzles, or quizzes contrary to deeper knowledge-building opportunities that might include vocabulary and comprehension (Vaala, Ly, & Levine, 2015).

No doubt this state of affairs represents an opportunity lost, especially given young children's interest and increasing use of media. According to the most recent survey, over 72% of children age 8 and under are using mobile devices for playing games, watching videos, and apps, up from 38% just 2 years before (Rideout, 2017). In this same time frame, the average time spent on

This article was published Online First April 26, 2018.

Susan B. Neuman and Kevin M. Wong, Steinhardt School, Department of Teaching and Learning, New York University; Rachel Flynn, Medical Social Sciences, Feinberg School of Medicine, Northwestern University; Tanya Kaefer, Faculty of Education, Lakehead University.

Correspondence concerning this article should be addressed to Susan B. Neuman, Steinhardt School, Department of Teaching and Learning, New York University, 239 Greene Street, New York, NY 10014. E-mail: sbneuman@nyu.edu

*This research was conducted as part of an award R305A150143 from the Institute of Education Sciences, U.S. Department of Education to New York University. The views expressed are those of the authors.

media activity has tripled, to more than 2 hr a day. Although TV still commands almost an hour of that time, streaming video services on mobile devices have rapidly gained ground, rising from 5 min a day in 2011 to 48 min a day in 2017. Moreover, a recent study (Kabali et al., 2015) reported that a staggering 97% of U.S. children under the age of four now own and use mobile devices regardless of family income, representing almost universal exposure. Despite recommendations to avoid digital media use for children ages 18 to 24 months of age from the prestigious American Academy of Pediatrics (2016), even our youngest children under the age of 2 have become regular users.

Consequently, although imposing restrictions on media use is certainly advisable, it may be more profitable to work toward improving the quality of the screen media children are likely to access. For early childhood in particular, this might mean a greater attention to the foundational language and vocabulary skills that lie at the heart of later reading comprehension. Even in infancy (birth to age 2), pediatric ratings of language milestones predict later reading achievement and the magnitude of the longer-term correlations between preschool language abilities and school outcomes is larger than any corresponding individual skill (Paris, 2005; Scarborough, 2001). In short, children's oral language skills when they enter kindergarten not only predict their later literacy skills in elementary school but later school success even through high school (Cunningham & Stanovich, 1997; Storch & Whitehurst, 2002).

Furthermore, studies (Hirsh-Pasek et al., 2015) suggest that the early years may represent an optimal time to promote oral vocabulary knowledge. Using the Children of the National Longitudinal Survey of Youth (NLSY79) national sample, for example, Farkas and Beron (2004) examined the monthly growth trajectory of vocabulary knowledge from ages 36 to 156 months. They reported that the highest rate of vocabulary growth occurred during the preschool years, with the rate declining for each subsequent age period. However, they also noted a troubling gap in vocabulary knowledge by race and class. For each race group, social class significantly affected vocabulary, with striking differences between low-and high-income families early on. These results are consistent with analyses that have shown the very early onset of group differentials by socioeconomic class (Halle et al., 2009). In fact, Fernald, Marchman, and Weisleder (2013) found socioeconomic status (SES) differences in vocabulary development as early as 18 months.

At the same time, there is evidence that early intervention in vocabulary development can work to mitigate these gaps (Wasik, Bond, & Hindman, 2006). For example, Marulis and Neuman (2010) in their meta-analysis of 67 studies reported an overall average effect size of .88, representing gains of nearly one standard deviation. Subsequent narrative analyses suggest that interventions which used multimedia meaningfully in instruction were among those demonstrating the largest vocabulary gains (Wright & Cervetti, 2016). When stories were accompanied by visual or other nonverbal information, the vocabulary words were retained better than if conveyed alone.

Therefore, screen media that apply what we know from developmental science could potentially enhance children's vocabulary and also take advantage of children's interest in educational media. For example, Takacs, Swart, and Bus (2015) meta-analysis of 29 studies found that multimedia features such as animated illustrations, background music and sound effects were linked to improvements in comprehension ($g+ = 0.40$) and vocabulary ($g+ = 0.30$). Given that digital devices have become almost ubiquitous in homes (Rideout, 2017), with nine out of 10 low-income families now owning smart phones, language-rich screen media with such enhancements as images, music, and sounds, could provide important educational opportunities for those who live in resource-poor neighborhoods. A recent study by Rideout and Katz (2016), for example, found that low-income families felt largely positive about media, and that children and parents frequently learn with and about the technology together.

In this article, we report on two studies designed to examine the landscape of literacy-related streamed video, and the features they may support vocabulary learning. In Study 1 we conduct a content analysis of online videos from popular streaming platforms and identify the pedagogical supports—techniques or features that are designed to orient children to specific vocabulary words. In Study 2, we use eye-tracking technology to examine which of these pedagogical supports might predict children's ability to identify program-specific vocabulary. Together, our goal is to derive certain principles of instructional design that might enhance children's opportunity to learn vocabulary from digital media.

Theoretical Foundation for the Research

Our research is based on two complementary theoretical assumptions. The first is *dual coding* (Paivio, 2008), the assumption that humans possess separate information processing channels, one devoted to processing verbal information (such as speech), while the other, to processing nonverbal information (such as visual images). Since two is better than one, information encoded both verbally and nonverbally is likely to be represented more fully in memory than information encoded through a single channel. Studies have demonstrated that adding nonverbal information to stories either read or heard enhances children's ability to figure out unknown words (Verhallen & Bus, 2010). In this respect, educational media has the potential to serve as a worthwhile scaffold for children's vocabulary acquisition by simultaneously providing both verbal and nonverbal information (i.e., speech accompanied by dynamic visual content).

In addition, the *synergy assumption* (Neuman, 1992, 2009) proposes that multimedia presentations can help children organize a more robust mental representation of content. For example, a book may explain that sharks swim through water, while a video dynamically demonstrates how it happens. Studies (Meringoff, 1980; Meringoff et al., 1983) suggest that children can recall actions more readily from video, while they can recall aspects of characterization more readily from text. This may be especially important for low-income children who may not possess the necessary background knowledge to make constructive use of new information presented in a single format (Fisch, 2000; Linebarger & Piotrowski, 2010).

According to both theories, educational screen media may support low-income preschoolers' vocabulary acquisition by offering multiple sources and types of information on the same topic (Bus, Takacs, & Kegel, 2015). In this way, watching educational media may help children develop more multidimensional and extensive understandings of new words and their meanings. This affordance, along with the potential to orient attention (Anderson & Pempek,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

2005), reduce cognitive demands (Sharp et al., 1995), and motivate knowledge-seeking (Kamil, Inrator, & Kim, 2000), suggests that educational screen media may be an especially powerful mechanism for encouraging vocabulary development and oral language comprehension for low-income children in the early childhood years (Verhallen, Bus, & deJong, 2006).

Nevertheless, despite its potential to support the vocabulary development of children from low-SES backgrounds, not all educational screen media are created equal. For example, irrelevant or funny information may distract children's ability to acquire new words and understand essential content (DeJong & Bus, 2004), while certain production elements may be more or less supportive of children's learning, particularly as they get older (Miller & Warschauer, 2014). Over the past several decades, research has addressed how production techniques used in educational screen media may affect children's viewing behaviors (Huston & Wright, 1983; Kirkorian, Wartella, & Anderson, 2008). These indicators, or formal features, include various production elements such as editing techniques (e.g., zooms, pacing) and character features (e.g., puppets, humans). Formal features may guide children to look at the screen when they are likely to be rewarded by important, entertaining, and/or comprehensible content (Huston & Wright, 1983).

But while formal features may indicate which screen elements capture children's attention, they do not necessarily differentiate between what might be educationally relevant and what might be just a source of entertainment. More specifically, formal features may not explicitly indicate (a) when informative content is about to be presented, and (b) which content is important for children to learn. Because formal features are operationalized independent of content (Kirkorian et al., 2008), children's attention may sometimes be rewarded by relevant vocabulary or content, while other times it may be rewarded by less important, albeit entertaining, features. Moreover, while children engaged in the process of learning tend to be highly attentive (Anderson & Evans, 2001), simply orienting attention does not necessarily mean that they will learn important vocabulary or content.

Consequently, in this study, our goal was to examine features that may *uniquely* direct children's attention toward educational content, specifically vocabulary than traditional formal features alone. For example, cues supporting word learning in social-interactional contexts may also help support vocabulary acquisition from educational screen media. During social-interactional word learning (Akhtar & Tomasello, 2000), two individuals (e.g., child and adult) jointly interact over a third entity (e.g., unfamiliar object). During such interactions, adults typically teach new information by attracting and directing children's attention toward relevant or salient information. To do so, they may take advantage of a range of communicative and referential signals, such as using exaggerated prosody, calling the child's name, establishing joint attention, and overtly pointing. For example, a teacher may introduce a new word by explicitly looking at and pointing to a referent while labeling (e.g., "Look at this! It's a *triceratops*"). In addition, certain *pedagogical cues* may provide valuable signals for vocabulary acquisition, helping children to focus on new words and content related to the words' meanings (Csibra & Gergely, 2006, 2009). Instructional strategies such as explicit definitions along with repeated exposure of these words in multiple contexts are known to be associated with vocabulary learning (Marulis & Neuman, 2010).

In this study, we attempt to identify *screen-based pedagogical supports* that elicit children's attention and convey pedagogical intent. Like formal features, these cues may attract sustained visual attention. In addition, however, they may also be linked to content, helping children develop a more extensive understanding of new words and their meanings. These cues may be particularly helpful for our more vulnerable children whose performance is likely to depend more strongly on the quality of educational input than for others.

Study 1

Our first study was designed to identify screen-based pedagogical cues in educational media that might support preschoolers' vocabulary development. According to a recent market scan of language and literacy digital media (Nichols Linebarger, Brey, Fenstermacher, & Barr, 2017), the majority of award-winning programs claim to teach specific skills, with alphabet/letter sound knowledge and vocabulary development among the most common. Building on this research, our goal was to examine *how* they might do so, moving beyond educational claims to actual pedagogical practices. Previous content analyses of infant-directed media, for example, had reported a relatively large amount of general language-related content (e.g., nearly a quarter of the scenes in a typical video) but relatively few instances of explicit vocabulary definitions (e.g., less than 1%; Vaala et al., 2010), although as shown in a subsequent content analysis of Sesame Street's Word on the Street initiative (Larson & Rahn, 2015), certain cues could be employed to promote vocabulary instruction in educational media (Neuman, Wong, & Kaefer, 2017). Therefore, in this study, we sought to conduct a more comprehensive and exhaustive market scan of streaming video, played via an "app" on a mobile device, to examine the following questions: (a) To what extent do online streamed videos focus on vocabulary development?; and (b) What are the pedagogical cues used to teach vocabulary?

Method

Sample. For this study, we defined educational screen media as programs that are deliberately designed and specifically marketed to educate children in school readiness skills such as language and early literacy. We began with an Internet search where young children are likely to have the greatest exposure, experience, and access. These sources included online streamed videos from Amazon Prime, HBO Now, Hulu, Google Play, and Netflix. From each of these streaming platforms, we selected the top 20 child/family educational media programs representing the most common in the media marketplace, which were (a) intended for preschoolers, ages 3–5; (b) targeted (at least in part) on language and literacy skills (per their description); and (c) recommended by expert review sites or awarded for their downloadable apps (e.g., Common Sense Media; PBS) with streamed media for this age group.

We collected an initial sample of 4,565 online streamed episodes from these top 100 programs. We subsequently eliminated redundancies, and then randomly selected two titles from each program for inclusion in the final sample. In total, our sample included 200 episodes, representing 108.9 hr of programming.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Content analysis coding strategy. Following a procedure by Fenstermacher et al. (2010), we identified a *scene* as our unit of analysis. A scene was defined as a *sequence of continuous action* in which a vocabulary word was introduced. For example, in Sesame Street's Word on the Street, the scene began with Cookie Monster asking "What does *respect* mean?" eliciting a response from a young girl "Treating people the way you want to be treated." The scene would end when it moved on to another topic. In total, we identified 2,277 scenes, with 700 novel words across the 200 programs.

Working collaboratively as research team, we watched a sample of 20 scenes from 10 different programs to generate categories of pedagogical supports. We identified two major categories of pedagogical supports: ostensive and attention-directing cues. Ostensive cues conveyed the meaning of a word through definition, multiple exemplars, and repetition. For example, the narrator might say "A hurricane is a very big storm with lots of wind and rain" with a background showing what a hurricane might look like. Or the onscreen character might give an explicit definition: "You're an author. That means you wrote your own book (pointing to the book)." As a production technique (Lesser, 1972), such ostensive cues have been described as direct teaching. They make salient the learning goal by telling and showing, often followed by telling and showing again. In each case, the gesture, picture, or demonstration is deliberately intended to bind the meaning of the word to its defined term. Repetition is designed to provide opportunities to practice these words as they become increasing familiar.

On the other hand, attention-directing cues were those that helped direct children's attention toward the target word, taking into account the narrow focusing capabilities of video (Lesser, 1972). For example, special sound effects might accompany the introduction of a word, such as "Is this box shaped like a *square*?" followed by a digital clicking sound effect, and dialogue "Right! It's the big *square* box." Or a character might use humor to get the young viewer's attention to a word, such as "Do you see any *pigeons*?" The responder says "It's right there on your head!" "My head? Ahhh *pigeon*!" As developers of Sesame Street reported in their formative research (Fisch, 2014), slapstick comedy, silliness in the form of pratfalls, and nonsensical events can serve to direct and sustain attention for preschoolers (Palmer & Fisch, 2000). In contrast to ostensive cues, attention-directing cues appeared to signal the importance of a target word, and not explicitly its meaning.

Within these larger categories, we then identified subcategories, along with explanations and examples from videos. With each iteratively developed coding session, we attempted to refine and clarify our codes, providing multiple examples from different programs. The final codebook included these two broad categories, with four ostensive and four attention-directing supports (see Table 1).

Two graduate research assistants were trained using a sample of scenes. Following the training, coders independently coded 20 scenes along with the second author. Percent agreement with the second author was calculated at 82.1%. Disagreements and areas of uncertainty were then flagged and resolved through discussion. After this session, a second set of 20 scenes were independently coded by the two research assistants. Overall reliability was 87.3%.

Results

In these results, we first describe the extent of vocabulary scenes across programs and the characteristics of these programs. We then describe the frequency and type of screen-based pedagogical supports most common throughout these programs.

Table 1
Screen-Based Pedagogical Supports

Type of support	Example
Ostensive cues	
Definitions	"A *subway* is an underground train." (Bubble Guppies)
Repetition	"Planet nine is *Pluto*. But, I don't see Pluto anywhere. Girl: That's because it's the SMALLEST planet, Pluto is very hard to find. Boy: Uhoh! Pluto is hiding! Boy: Yes, there he is! Pluto, it's time to go back where you belong. (Little Einstein)
Features of target words	"*Otters* have webbed feet that help them swim (Otter demonstrates). "They use their webbed feet like paddles." (Go Diego Go)
Examples	"A *career* is a job that you train for that you expect to have for a long time. That could be an architect, a teacher, or a scientist." (Sesame Street)
Attention-directing cues	
Visual effect	"Wow! A *volcano* (picture appears) (Aquaphonics adventure)
Sound effect	"Is this box shaped like a *square*?" [digital clicking sound effect] "Right! The big square box!" (Dora the Explorer)
Humor	Princess: This must be the *jungle*. It's on the other side of this . . . this . . . Girl: Quicksand (Around the World Adventure)
Guess/pause	A: "How about a piece of fruit?" B: "Hmmm . . . a piece of fruit would be great! Which of these is a piece of fruit? [pause] (Blue's Clues)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Extent of Vocabulary Opportunities

Characteristic	Description
Number of programs sampled	200
Programs with vocabulary scenes	132
Programs without vocabulary scenes	68
Number of hours coded	108.9
Number of scenes identified in apps	2,277
Average amount of screen type devoted to target word	19s
Average number of target words per episode in programs with vocabulary scenes	2.74 (*SD* 4.30) (R = 1–7)
Number of target words in scenes	700
Type of words	
Nouns	96%
Verbs	2%
Adjectives	2%

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Extent of vocabulary scenes. Table 2 describes the prevalence of vocabulary scenes across the 200 programs. Approximately two thirds of the educational programs included targeted vocabulary scenes, suggesting a significantly higher percentage than previous studies have reported (Vaala et al., 2010). Nevertheless, 68 programs in our sample did not provide any targeted vocabulary opportunities, representing a sizable portion of programs for young children.

As shown in Table 2, words targeted for instruction included mostly nouns, considered by researchers to be generally more visually salient, and more memorable for young children than other parts of speech (Rice & Woodsmall, 1988). Of the 700 novel words, 96% were nouns, 2% verbs, and 2% adjectives (see examples in Table 3). To calibrate the level of difficulty of words in these scenes, we used the collection of recordings from the CHILDES data set (MacWhinney, 2000), which consists of transcriptions of adult–child spoken interactions in different home and laboratory settings around the world. Based on this dataset, we took a random sample of target words from the scenes and compared them to those in CHILDES known to be familiar to typically developing children 5-years and under. As shown in Table 3, approximately half of the words would likely to be known by 3- and 4-year-old children, while the other half, more challenging. We could find no rationale for the selection of particular words in any additional materials.

Frequency and type of screen-based pedagogical supports. Table 4 describes the pedagogical cues used most frequently across programs. Here, we coded each scene according to the type of cue most heavily featured in that particular scene.

As shown in Table 4, show designers used attention-directing cues far more often than ostensive cues. In these attention-directing scenes, the focus was to signal a target word without necessarily describing its meaning. Most often this included a visual example, or a visual effect of some kind. For example, in one scene a narrator says the word *pumpkin*, followed by a pumpkin glowing brightly with sparks coming out of it. Another type of attention-directing cue included sound effects. For example, in *Telo and Tula*, Tula notes "Today we're going to make an *apple pie*!" This is followed by the sound of an organ which plays as a picture of an *apple pie* appears on screen. Other attention-directing cues used humor, and pauses/questions, such as "Which is a piece of *fruit*?" followed by a pause then answer, although far less frequently than others.

Ostensive cues, on the other hand, were used in less than a third of the scenes. Of these cues, definition and repetition were the most common techniques for conveying the meaning of target words. In some cases, the ostensive cue might include an explicit definition, such as "A *subway* is an underground train." At other times, the characters might act out the meaning of the word, such as when Leo says "*Adagio,* that means we're going to go

Table 3
Examples of Target Words in Apps Based on CHILDES Dataset

Likely to be known by 4-year-olds	Likely to be known by 6- to 8-year-olds
Ears	saliva
Nose	violet
Snake	dam
eyes	baguette
purple	flock
ramp	hoofs
soccer	mercury
insect	ukulele
balloon	nibble
bird	flush
camel	hopeless
corn	ecstatic
egg	basilisk
finger	limestone
girl	level
green	skyscraper
ham	substitute
heart	antelope
horn	baboon
ice cream	broccoli
kite	fog
mouse	menorah
pie	quail
ring	sprinkler
snow	barbeque
star	unicorn
taxi	waterfall
flashlight	submarine
pot	octagon
Percentage of words: 56%	Percentage of words: 44%

Table 4
Frequency and Type of Screen-Based Pedagogical Supports (N = 2,277)

Type of support	*N*	Frequency
Ostensive cues		
Definition	434	19.2%
Repetition	277	12.2%
Features	72	5.2%
Examples/categories	36	4.2%
Attention-directing cues		
Visual effects	1,094	47.0%
Sound effects	241	16.6%
Humor	45	5.0%
Guess/pause	78	3.4%

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

sloooowwww." Repeating the word frequently was another type of ostensive cue. For example, in one Muppet scene, the puppet says "The Kraken is holding a *boat* (verbal emphasis). He must like playing with *boats*." Thinking it's a good idea, puppet Abby replies "That's it! Maybe we should try a *boat*. One *boat* coming up. Wave your wands or your fingers and say *boat, boat, boat! Boat, boat, boat!* (boat appears)." Other types of ostensive cues such as an explanation of the generic properties of a category of words, or a discussion of its features were used less frequently. These types of cues, however, are often associated with helping children develop a deeper meaning of vocabulary words and comprehension than mere definition and labeling (Gelman & Kalish, 2006).

Findings

In short, our content analysis painted a somewhat more optimistic picture of vocabulary opportunities in educational media currently on the market than previous content analyses. For example, in a previous study of infant-directed media, Vaala et al. (2010) reported less than 1% of the scenes were targeted to vocabulary in their content analysis of 58 digital programs. Similarly, in a more recent content analysis of two episodes in 15 educational series, researchers (Nichols Linebarger et al., 2017) reported only 4.92 scenes included new vocabulary, some of which were mislabeled or mismatched with their visual referents.

In contrast, our analysis showed that 66% of the programs included vocabulary instructional opportunities. These different calculations could reflect differences in the sample size of our analysis, or in recent changes in the marketplace. It could also reflect differences in the media platforms we reviewed; previous studies have examined educational or cable TV (Vaala et al., 2010). However, although we found far more opportunities for vocabulary development across programs, there was great variability: The average number of words across programs varied dramatically, and the choice of words seemed to represent a curious mix of challenging, unique, and most easily pictured nouns throughout the programs.

Our analysis identified two categories of screen-based pedagogical cues: attention-directing cues, designed to focus and signal children's attention to a target word, and ostensive cues, to provide explicit definitions, repetitions, and examples to help explain a word. Perhaps using the medium to its advantage, attentional cues through visual and sound effects were found to be far more prevalent than the ostensive cues which relied on verbal descriptions and definitions of words that might be outside children's direct experiences.

Study 2

Having identified these types of pedagogical cues, our next question was to explore whether these cues might influence children's attention, and subsequently lead to their ability to identify words—in the context in which they were seen in the video, and outside it, in a new context. Our goal was to understand how these media supports might affect low-income children's vocabulary, particularly for those who might need additional supports for learning unfamiliar words. Smeets and Bus (2012), for example, found that the availability of additional dimensions in stories with rich images, music, and sounds enabled more vulnerable young children to construct a coherent representation of story events better than compared with a traditional storybook with static features, resulting in an additional 6% increase in word learning.

Nevertheless, not all features are equally facilitative for learning. For example, attention-directing cues that zoom in on the critical details of words and their meaning might attract children's immediate attention, and have positive effects on their vocabulary development (Hirsh-Pasek et al., 2015). Visuals contingent with oral text may enable simultaneous processing, enhancing vocabulary meanings. At the same time, the focus on individual words might come at a cost, distracting children from the overall meaning of the story (Mayer, 2001). Ostensive cues, on the other hand, which rely on verbal exchanges and repetition might not grab children's attention, especially for those who might experience problems with processing verbal information. Although verbal information might be helpful in promoting comprehension of story events, children may be slower to fixate on these verbal cues for vocabulary development.

Therefore, this study was designed to test these assumptions. Recognizing that children cannot learn from educational messages to which they do not pay attention (Anderson & Pempek, 2005), we examine how the most common cues identified in Study 1 affect children's sustained attention and vocabulary identification. Specifically, we address the following questions: (a) To what extent do different screen-based pedagogical cues influence children's attention? Are there differences between attentional-directing and ostensive cues?; (b) Is there a relationship between the use of screen-based cues and word identification, in context, and out-of-context?; and (c) How is this relationship influenced by child characteristics (i.e., general vocabulary knowledge)?

Method

Sample. After receiving permission from educational directors in two Head Start Centers and parent consent, 12 classrooms of 3- to 4-year children were selected to participate in the study. Centers were located in a poverty-impacted neighborhood in a large urban city. Given the word difficulty levels reported in Study 1, 4-year-olds were the target focus for our analysis. From these classrooms, 110 4-year-old children were randomly selected (M = 4.39; SD = 0.71); 44% were female. The sample was culturally diverse; 60% were African American, 38% Hispanic, and 2% Caucasian. All children in these centers qualified for free and reduced lunch. Average receptive language score as measured by the Peabody Picture Vocabulary Test (PPVT) was 87.13 (SD 15.21).

Design. Based on our content analysis, we selected the most frequently reported attention-directing and ostensive cues across programs for our analysis. For attention-directing cues, for example, we isolated three scenes in which a vocabulary word was introduced through visual effects and three in which a sound effect signaled a vocabulary word. Similarly, for ostensive cues, we isolated three scenes in which an explicit definition was given to identify a word, and three in which repetition was used to identify a word. To avoid a program effect, all scenes were selected from different episodes of Sesame Street, (e.g., from Sesame Street, 2006–2013 archives) with the average scene length of 21.42 s. In total, 12 scenes from these episodes, three words per cue were used for our initial analysis, for a total of 257 s.

Given that our selection used extant scenes not subject to experimental manipulation, there was variability in word difficulty across scenes. To reduce error, we used a within-subject design in which all participants received all 12 scenes in a counterbalanced approach, serving as his or her own control. Therefore, in this study, each participant was shown four scenes representing each of the pedagogical supports counterbalanced for order effects, in three separate rounds. Because each child received all four pedagogical supports, we were able to control for between-subjects variability, increasing our power to detect differences. In addition, it allowed us to control for threats to internal validity since individuals act as their controls. Table 5 describes the words and a brief description of the scene.

We used multiple methods to examine children's attention to these pedagogical cues and word identification, including a standardized assessment, researcher-developed vocabulary measures, as well as an analyses of eye tracking movement patterns. As a noninvasive method, eye tracking would allow for a more precise analysis of how young children distributed their attention to these cues. In the current study, we used measures of fixation, specifically, when the eyes focused on a particular area. Fixations are typically identified as the center of visual attention (e.g., Henderson & Ferreira, 1990; Henderson & Macquistan, 1993), and are guided by attentional processes (e.g., Rayner, Sereno, & Raney, 1996).

Measures.

Screening measure. Each child was administered a brief screening measure prior to the study. The measure included a picture of each word on individual card, as well as additional picture foils for a total of 20 items. Designed to be an expressive task, the assessor asked "What is it?" Six children who accurately identified one or more words were screened out of the study.

Peabody Picture Vocabulary Test-IV (PPVT; Dunn & Dunn, 2007). Used as a baseline measure, the PPVT is an individually administered, norm-referenced test designed to be a valid and reliable measure of receptive language skills. Reliability ranged from .91–.94. For this study, raw scores were converted to age-related standard scores.

Word identification. Following the viewing of a set of four scenes (described below), children were individually administered an eight-item word identification task: words in context and words in new context. Similar in format to a PPVT, children were asked to point to the correct word among three other options. For words in context (four items), distractors included pictures from a similar clip, thematically related to the key word (i.e., key word "hurri-

Table 5
Examples of Target Words and Scenes by Pedagogical Support

Target word	Scene
Shelter	Jimmy: But, look. The sun is getting lower in the sky. Night will fall soon. I must start to build myself a *shelter*. Elmo: But wait, wait, wait, what's a *shelter*? Jimmy: A place where I can sleep, where I can stay warm and dry and protected from the elements. I must act quickly. (ostensive-definition) (7.9 s)
Whisk	Murray: What other tools do you use? Girl: Well you could use a whisk. Murray: What is a whisk? Wait what's a whisk? Girl: A whisk is something that you stir with. Murray: Where's the whisk? Girl: Here's the whisk. (ostensive-repetition) (8.7 s)
Grater	Murray: What tools do you use in the kitchen? Woman: We're gonna use a grater. (holds it) Murray: A grater? That sounds great. (Attention-directing-visual effect) (6.6 s)
Square	Zoe: Oh, oh I see one, a square. Elmo: Square. Zoe: The front of the cookie box (sound effect) Cookie Monster: Hey what you know about that, circle in square. Oh well break over, time to eat cookie. (Attention-directing-sound effect) (8.7 s)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

cane;" distractors meteorologist; blizzard; rain). For words in new contexts (four items), children were asked to select the correct word in a new context (not from a video scene), along with three other different distractors. Items were randomized and then presented in a set order across children. Children received a score for words in context and a score for words in new contexts for each of three rounds, totaling 12 in-context and 12-in new context word identification scores. Reliability, calculated for the 24-item assessment, was .80.

Eye-tracking technology.

Apparatus. Eye movements were measured with a Tobii Technology T120 eye-tracker integrated into a 17 in. thin film transistor (TFT) monitor (Psychology Software Tools, Pittsburgh PA). This is a remote eye-tracking system that had no contact with the child. The typical spatial accuracy of this system is approximately 0.5 visual degrees, and the sampling rate is 120 Hz. During tracking, the eye-tracker uses infrared diodes to generate reflection patterns on the corneas of the child's eyes. These reflection patterns, together with other visual information about the child, are collected by image sensors and used to calculate the three-dimensional position of each eye and gaze-point on screen. This system uses a binocular tracking method, which allows for increased head movements. Head movements typically result in a temporary accuracy error of approximately 0.2 visual degrees. In the case of particularly fast head movements (i.e., over 25 cm/s), there is a 300-ms recovery period to full tracking ability. An embedded camera is also used to record the child's reactions.

General procedure. Preschoolers sat approximately 60 cm/s from the monitor. Video scenes were displayed on the Tobii monitor with a second monitor facing the experimenter. Tobii Studio Professional 3.0 software was used for stimuli presentation and data processing.

To calibrate gaze, an attention-grabber was shown at five points on the screen. A manual calibration procedure was used; accuracy was checked by Tobii Studio software and repeated as necessary. Following calibration, a 2-s attention-grabber appeared in the center of the screen prior the beginning of each eye-tracking task. After calibration, children would then view four scenes. During each scene, the research assistant was able to follow the child's eye movements and behaviors using the live view on the second monitor. Total duration of each eye-tracking round was approximately 3 min. Children returned either on the same day or the day after for two more rounds for a total of 9 min of eye-tracking activities.

General data processing. Eye movement data was extracted using Tobii Studio 3.0 software. Fixations were defined as any gaze coordinates lasting at least 60 ms, and were identified using the Tobii Studio fixation filter. Adjacent gazes (i.e., gazes within a 0.5° radius, lasting less than 75 ms) were merged into a single fixation. To help visualize data, fixations were overlaid onto a video recording of stimuli presented in each scene. We then extracted fixation data of each area of interest (AOI) for each child.

Procedure. Children were individually administered the PPVT prior to the start of the study. Following baseline assessment, the child would be escorted to the library to watch video scenes on the eye tracker. Two trained graduate assistants assisted at all times in the data collection.

Children were assigned to one of three sequences of video scenes. For example, after calibrating the gaze, a child would watch four brief scenes (each with a different cue). The researcher assistant would then administer the word identification tasks. The next child would watch a different set of scenes followed by the appropriate word identification tasks. In this manner, we counterbalanced the treatment throughout the data collection. Second and third rounds occurred sometime later in the day or the next day following the same administrative protocol. Data for all three rounds included 104 4-year-olds.

Analysis. Dynamic AOIs were drawn around the target items within the screens for the entire span of time the item was on screen. In the case of the word, "hurricane," for example, it began with an image of a hurricane at the same time a newscaster was explaining the term. The image appeared on screen for the entire length of the clip, which was 38 s. In this segment, therefore, the AOI was drawn around the image of the hurricane and maintained for the entire time it was on screen.

To examine children's attention, we used three measures of fixation. First, we used the total fixation duration to the screen in each clip (i.e., this could include any object, conversation, or scene, not just the target word) as a measure of overall attention. Second, we calculated the time it took for children to fixate their attention on the novel object after it was named (i.e., how long it took to look at the visual of *hurricane* once the word was said) as a measure of orientation. Third, we calculated the amount of time spent fixating on the target item after it was named as a measure of targeted attention. These measures have been used extensively to examine visual attention (e.g., Both-deVries & Bus, 2014; Neuman, Pinkham, Kaefer, & Strouse, 2014).

Given that the video scenes varied in length (e.g., see minor variations in Table 5), we created percentages for fixation duration to examine differences in attention across the four pedagogical cues. Percentages were created by dividing each child's fixation duration by the total time of the scene. Once the percentage data was calculated for each word, we then averaged that information across words to get the mean fixation time (or orientation time) for each of our pedagogical categories. We then used repeated measures ANOVA with the four cues as the within subject factor, and the child's age in months as a covariate, followed by paired sample *t* tests to examine differences between these types of cues, and word identification in context and in new context. Age in months was used as a covariate to account for potential developmental differences across the 4-year-old age span among our sample. Finally, we explored whether these differences might reflect a language factor, using a median split, to compare these results for children who had higher or lower PPVT scores.

Results

Attention and screen-based pedagogical supports. Table 6 describes the means and standard deviations of children's attention to these screen-based supports. For overall visual attention, repeated measures analysis revealed no significant effect of age, $F(1, 103) = .075$, $p = .784$, or a significant age by cue interaction, $F(3, 101) = .51$, $p = .677$. There was, however, a statistically significant main effect of cue, $F(3, 101) = 6.23$, $p < .001$. To examine these differences, we collapsed the pedagogical cues into our two categories. As shown in the table, scenes that used ostensive cues (definition, repetition) attracted more attention overall than scenes that used attention-directing cues (visual effects; sound effects),

Table 6
Means and Standard Deviations of Children's Attention to Vocabulary Scenes

Variable	Ostensive cues	Attention-directing cues
Proportion of total time spent fixated on screen***	.57 (.19)	.25 (.10)
Proportion of time attending to target (after target named)***	.31 (.12)	.49 (.15)
Time to fixation on target (after target named)***	7.71 (1.71)	9.68 (1.19)
In context word learning	.61 (.23)	.62 (.23)
New context word learning**	.55 (.24)	.62 (.25)

** $p < .01$. *** $p < .001$.

$t(104) = 21.08$, $p < .001$. That is, children looked longer in general at everything on the screen when ostensive cues were used.

We then examined the time it took for the child to orient to the new item once the vocabulary word was used. Similarly, a repeated measures analysis showed no significant effect of the covariate, $F(1, 103) = .67$, $p = .416$, or a significant age by cue interaction, $F(3, 101) = .49$, $p = .692$. There was, however, a statistically significant main effect of cue, $F(3, 101) = 2.69$, $p = .046$. Again, we followed up this analysis by collapsing the pedagogical cues into our two categories. As shown in the table, ostensive cues were also faster to orient children toward the specific target word than attention-directing cues, $t(103) = 9.86$, $p < .001$. It took a shorter time for children to orient to target words with ostensive cues than attention-directing cues (7.71 compared to 9.68).

Finally, we examined the amount of time children looked at the target word after it was named. Once again, we found no significant effect of age, $F(1, 103) = 1.46$, $p = .230$, or a significant age by cue interaction, $F(3, 101) = 1.07$, $p = .367$. There was, however, a statistically significant main effect of cue, $F(3, 101) = 2.99$, $p = .035$. Follow-up tests showed children looked longer at the target word with attention-directing cues than with ostensive cues, $t(103) = 16.24$, $p < .001$. Children focused more time specifically on the target word *after* it had been named with these cues. In brief, it indicates that the two cues may have served somewhat different purposes: Children paid more attention to the screen with ostensive cues, suggesting that they were more engaged in the scene. As a consequence, they were faster to orient to the target item when it was announced. But they did not stay there. Rather, the attention-directing cues kept them directed to the target word, suggesting that the time they spent actually looking at the item was more important than looking at the overall scene.

Examining the results of our word identification tasks partially bears out this thesis. As shown in the table, there were no differences between these cues for in-context word identification $t(102) = 0.35$, $p = .730$. This was a relatively straightforward task, asking children to simply recall the scene in which the target word was given. However, there was a significant difference in word identification in new contexts, in which the task required the child to label a word without such contextual support, $t(102) = 3.09$ $p = .003$. In this case, children identified more words in new contexts when watching the attention-directing cues than the others. These results indicate that the focused time on the vocabulary word essentially paid off in a greater ability to identify the word.

Differences by language proficiency. Our final analysis focused on whether there were differences in attention and word identification by children's language proficiency. To do so, we created a categorical variable for PPVT, with those children who were higher ($M = 98.77$, SD 10.56) and those who were lower ($M = 75.50$, SD 8.95) in receptive language. As shown in the table, there were no significant differences in attention to either set of cues (all $ps > .01$). Both groups showed similar patterns, with children spending greater attention on the ostensive cues of definition and repetition than the visual or sound effect cues.

No significant differences between groups were reported in identifying words in context. Both groups appeared to score statistically equivalent on word identification in context (all $ps > .1$). This was not the case, however, for words in new contexts. In this case, children with higher PPVT scores were more likely to benefit from the attention-directing cues than ostensive cues $t(46) = 3.52$, $p = .001$, whereas there were no significant differences in cue type for children with lower PPVT scores $t(53) = 1.25$, $p = .216$. These results suggest that children with higher receptive language skills were able to use these cues more effectively to identify words in new contexts than those with lower receptive language skills (see Table 7).

Discussion

These studies were designed to survey the current landscape of educational media on streaming platforms and to examine the kinds of pedagogical features that might predict vocabulary learning. Building on previous research that scanned the marketplace (Fenstermacher et al., 2010; Vaala et al., 2010), our focus in Study 1 was to conduct a content analysis of vocabulary learning opportunities, and to determine the ways in which developers designed their programs to teach vocabulary. Based on this analysis, we then

Table 7
Means and Standard Deviations of Screen-Based Pedagogical Supports by PPVT

Variable	Ostensive cues	Attention-directing cues
Attention		
Lower PPVT	.53 (*SD* .21)	.24 (*SD* .17)
Higher PPVT	.60 (*SD* .19)	.26 (*SD* .06)
Word identification-in context		
Lower PPVT	.56 (*SD* .31)	.57 (*SD* .29)
Higher PPVT	.64 (*SD* .28)	.65 (*SD* .31)
Word identification-new context		
Lower PPVT	.50 (*SD* .26)	.54 (*SD* .27)
Higher PPVT	.59 (SD.22)	.72 (*SD* .22)

Note. PPVT = Peabody Picture Vocabulary Test-IV.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

attempted to isolate the most prevalent pedagogical cues to examine which might predict children's attention and subsequent ability to identify words both in-context and in new contexts.

Our findings revealed that over 66% of the programs streamed on video platforms had at least one or more vocabulary scenes, a percentage substantially more than previous scans of the marketplace of educational TV (Linebarger & Piotrowski, 2010; Nichols Linebarger et al., 2017). Nevertheless, there was considerable variability, ranging from one to seven scenes, across programs in how many opportunities children had to learn words. Of the over 2,000 vocabulary scenes identified, more than half of the words were likely to be already known by 4-year-olds, while the other were unique to this age group, and more likely to be known by 6- to 8-year-old children.

These results were in contrast to Larson and Rahn (2015), content analysis of vocabulary episodes in Sesame Street's Word on the Street initiative. In their case, word difficulty was measured using Beck and McKeown's Tiers heuristic (e.g., Tier 1 representing familiar words; Tier 2 words worth teaching; and Tier 3 content-related words; Beck, McKeown, & Kucan, 2002). These researchers found that over 76% met the criteria of Tier 2 words, with only 12% Tier 1, considered familiar. Such differences in our findings most likely reflect the selection of programs reviewed. In our review, for example, vocabulary was one among other skills taught in these educational media, whereas Word on the Street represented a deliberate initiative designed to improve vocabulary development for young children at risk.

Still, the choice of words in both content analyses was somewhat perplexing. As Fisch (2014) has described, it could reflect the tension often faced by media developers between the wish to entertain and the desire to educate. In our case, for example, words like *broccoli*, *baboon*, and *ukulele*—all words unlikely to be known by preschoolers would also not easily fall into a Tier 2 category, that is, high utility words across a variety of contexts. Similarly, words like *humongous*, *prickly*, and *splatter* in Word on the Street, would hardly seem exceptional candidates for direct teaching, another criterion for Tier 2 words. Rather, in each case, we could find no rationale for the selection of words, leading one to question the ultimate educational utility of vocabulary teaching in these media. Given that so few words can be directly taught (Anderson & Nagy, 1992), subsequent work by media developers might consider a more efficient, effective, and systematic approach to word selection. For example, there is now a plethora of word lists and norm-referenced lists that could serve as a resource for future development (Biemiller, 2009; Hiebert & Pearson, 2010).

To teach words, producers of media must rely on techniques to engage children's attention. Building on the research by Anderson and his colleagues (Anderson & Pempek, 2005), studies have shown that visual attention to media is associated with learning, and that program content and production techniques can maximize children's attention to programs. For example, Huston and Wright (1983) identified a set of formal features, such as music, dialogue, sound effects, zooms, and cuts, and demonstrated how these features encouraged young children's thoughtful processing. Nevertheless, attention to media does not necessarily predict comprehension. Young children may simply respond automatically to the saliency and unfamiliarity of formal features. Rather, the content of the educational message must be understandable (Neuman et al., 2017). Fisch (2000) has argued that embedding educational content within a narrative structure capitalizes on children's cognitive resources, which have limited capacity at a young age, and therefore, may aid in comprehension.

Identifying the pedagogical features used by producers to teach vocabulary words was an effort to acknowledge both content and formal features. As one might predict, we found that these educational media relied on attention-directing, specifically visual effects, far more than ostensive cues. These results might be due to producers view of the developmental limitations of preschoolers' background knowledge. Ostensive definitions often rely on analogical reasoning or comparing one thing to another (Gelman & Kalish, 2006), assuming that the listener has sufficient background understanding to recognize the information being given. Attention-directing cues may also reflect the medium's ability to tell a story through sound, animated visual images, and music (Bus et al., 2015).

How these cues might function for vocabulary learning, in particular, was the focus of Study 2. The results of our analysis provided a more complicated pattern of attention than earlier studies have suggested (Bryant & Anderson, 1983). For example, recognizing that attention is a necessary prerequisite to understanding and retention, Anderson, Choi, and Lorch (1987) in their studies of Sesame Street observed a phenomenon described as *attentional inertia.* That is, the longer a look on a screen (in this case, TV) is maintained, the conditional probability that it will be further maintained increases substantially. In other words, the chance of losing attention is at its highest within that first look; from then on, the chances of looking away go down. The assumption is the longer the look, the greater the educational benefit.

Yet in our case the "longer the look" (at the scene) did not predict word learning. Rather, these active and engaged children appeared to make decisions about when and what to look for. Although children spent more time looking in scenes using ostensive cues, and were quicker to orient to the target word, they spent less time looking at it. On the other hand, with attention-directing cues, once the target word was named, children spent more time looking at it, and were more likely to identify the word in new contexts. For word learning, therefore, attention-directing cues seemed to be the most effective strategy. Children seemed to actively monitor the scene and to make ongoing decisions about what was most relevant within it. It reflected an active, "minds-on" pattern of viewing, in which children are more likely to quickly sample parts of a program most salient to them (Hirsh-Pasek et al., 2015). Whether this "sampling" draws children's attention to certain relevant content (e.g., in this case words) at the *expense* of the overall meaning or comprehension, however, is something that should be examined in future research.

Unfortunately, neither ostensive nor attention-directing cues by themselves appeared to exert additional support for children with lower PPVT. These cues did not help to level the playing field. Children with higher receptive language scores identified more words using both sets of cues than their lower PPVT peers. This finding replicates results from numerous studies on incidental word learning as well as explicit vocabulary instruction (Coyne et al., 2013), and further supports the existence of Matthew Effects (i.e., the rich get richer while the poor get poorer; Stanovich, 1986) in vocabulary development. It suggests that without intensifying vocabulary supports for children who are most at risk for language and/or reading difficulties, the current educational media might further exacerbate the gap rather than close it (Coyne et al., 2010; Penno, Wilkinson, & Moore, 2002).

We recognize that there are a number of limitations in the present study. The study was correlational in its design. We do not claim to draw causal inferences that these pedagogical cues foster vocabulary learning. In addition, our analysis of pedagogical cues was based on extant programming, representing the clearest examples of each set of cues and words targeted for instruction. To account for minor differences and time differentials within clips, we used a proportional score and a within-subject design to control for such variability. Further, although eye-tracking is a noninvasive strategy, we recognize that it does not represent a typical viewing context. In more natural settings, numerous distractions (e.g., dinner time, play activities, multitasking) may mediate children's attention from media. And finally, our analysis of visual attention was based on fixation variables. A more fine-grained analysis of eye movement patterns, for example, might better describe the dynamics of attention in relation to the pedagogical supports in educational media for word learning.

Recognizing these considerations, we believe that our findings represent an important first step in understanding the potential of streaming videos, their production techniques and how they may support children's word learning. Our analysis of the current landscape suggests that children may be exposed to more vocabulary-building experiences than previous reported. This is good news and may suggest that on the advice of experts, producers of media are beginning to address the more complex skills of vocabulary and knowledge-building experiences than in the past, when much educational programming and educational apps were largely focused on basic skills (Guernsey et al., 2012). Educational media has an enormous potential to enhance children's access to vocabulary through digital stories and informational programming when it is consistent with established theories of learning (Lesser, 1972). Studies have shown that when optimally designed, digital stories can facilitate the learning of new vocabulary and story comprehension (Takacs, Swart, & Bus, 2014).

Our analysis of pedagogical cues may offer several promising new directions for research and media production. It might extend the research on features that support learning, leading to a more nuanced model of attention that could be useful for media producers. For example, attention-directing cues might focus children's attention more deliberately, targeting particular skills for learning. Specific sound effects might prime children to pay attention to a new word. On the other hand, ostensive cues might be used at various points in a story to sustain attention and promote the narrative thread throughout a story, leading to greater overall comprehension. In both of these examples, different cues could be used more intentionally to engage children's thinking and to further bolster story and text comprehension.

Finally, studies have shown the potential advantage of engaging both sets of cues more intentionally (e.g., such as sound effects matched to verbal (definitional) cues) on children's learning, matching the nonverbal information sources with the oral dialogue or text. Consistently with our theoretical model of dual coding (Paivio, 2008) when there is close congruency and temporal proximity between channels, these cues can potentially support learning, and could address the more intensive support that children with language difficulties might need (Verhallen et al., 2006). Too often, however, studies have shown verbal/visual/sound effect mismatches (Linebarger & Piotrowski, 2010; Vaala et al., 2010), potentially diverting children's attention from the language and word meanings. Well-designed digital stories that use these pedagogical cues to intensify word learning in a synergistic manner might be especially helpful to children who have fewer background experiences or might need greater supports for deriving word meanings. These results could support a more intentional approach to media design to enhance children's opportunity to learn vocabulary. Given their enormous appeal to young audiences, maximizing the design capabilities of these digital assets may offer an important additional scaffold to facilitate low-income children's vocabulary development.

References

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195130324.003.005

American Academy of Pediatrics. (2016). Council on communications and media: Media and young minds. *Pediatrics, 138,* 2016–2591.

Anderson, D. R., Choi, H. P., & Lorch, E. P. (1987). Attentional inertia reduces distractibility during young children's TV viewing. *Child Development, 58,* 798–806. http://dx.doi.org/10.2307/1130217

Anderson, D., & Evans, M. (2001). Peril and potential of media for infants and toddlers. *Zero to Three, 22,* 10–16.

Anderson, D., & Pempek, T. (2005). Television and very young children. *American Behavioral Scientist, 48,* 505–522. http://dx.doi.org/10.1177/0002764204271506

Anderson, R. C., & Nagy, W. E. (1992). The vocabulary conundrum. *American Educator, 14,* 44–46.

Beck, I., McKeown, M., & Kucan, L. (2002). *Bringing words to life*. New York, NY: Guilford Press.

Biemiller, A. (2009). *Words worth teaching: Closing the vocabulary gap*. Columbus, OH: SRA/McGraw-Hill.

Both-de Vries, A. C., & Bus, A. G. (2014). Visual processing of pictures and letters in alphabet books and the implications for letter learning. *Contemporary Educational Psychology, 39,* 156–163. http://dx.doi.org/10.1016/j.cedpsych.2014.03.005

Bryant, J., & Anderson, D. (Eds.). (1983). *Children's understanding of television: Research on attention and comprehension*. New York, NY: Academic Press.

Bus, A., Takacs, Z., & Kegel, C. (2015). Affordances and limitations of electronic storybooks for young children's emergent literacy. *Developmental Review, 35,* 79–97. http://dx.doi.org/10.1016/j.dr.2014.12.004

Coyne, M., McCoach, D., Loftus, S., Ruby, M., Creveceour, Y., & Kapp, S. (2010). Direct and extended vocabulary instruction in kindergarten: Investigating transfer effects. *Journal of Educational Effectiveness, 3,* 93–120. http://dx.doi.org/10.1080/19345741003592410

Coyne, M., Little, M., Rawlinson, D., Simmons, D., Kwok, O., Kim, M., . . . Civetelli, C. (2013). Replicating the impact of a supplemental reading intervention: The role of instructional context. *Journal of Research in Educational Effectiveness, 6,* 1–23.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. Johnson (Eds.), *Processes of change in brain and cognitive development: Attention and performance XXI* (pp. 249–274). Oxford, UK: Oxford University Press.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13,* 148–153. http://dx.doi.org/10.1016/j.tics.2009.01.005

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934–945. http://dx.doi.org/10.1037/0012-1649.33.6.934

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

DeJong, M., & Bus, A. (2004). The efficacy of electronic books in fostering kindergarten children's emergent story understanding. *Reading Research Quarterly, 39,* 378–393. http://dx.doi.org/10.1598/RRQ.39.4.2

Dunn, L., & Dunn, D. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Bloomington, MN: Pearson Education, Inc.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research, 33,* 464–497. http://dx.doi.org/10.1016/j.ssresearch.2003.08.001

Fenstermacher, S., Barr, R., Salerno, K., Garcia, A., Shwery, C., Calvert, S., & Linebarger, D. (2010). Infant-directed media: An analysis of product information and claims. *Infant and Child Development, 19,* 557–576. http://dx.doi.org/10.1002/icd.718

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science, 16,* 234–248. http://dx.doi.org/10.1111/desc.12019

Fisch, S. (2000). A capacity model of children's comprehension of educational content on television. *Media Psychology, 2,* 63–91.

Fisch, S. (2014). *Children's learning from educational television: Sesame Street and beyond*. New York, NY: Routledge.

Gelman, S., & Kalish, C. (2006). Conceptual development. In W. Damon, R. Lerner, D. Kuhn, & R. Siegler (Eds.), *Handbook of child psychology* (pp. 687–733). New York, NY: Wiley.

Guernsey, L., Levine, M., Chiong, C., & Severns, M. (2012). *Pioneering literacy in the digital wild west: Empowering parents and educators*. Washington, DC: New America Institute.

Halle, T., Forry, N., Hair, E., Perper, K., Wandner, L., Wessel, J., & Vick, J. (2009). *Disparities in early learning and development: Lessons from the Early Childhood Longitudinal Study—Birth Cohort (ECLS-B)*. Washington, DC: Child Trends.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429. http://dx.doi.org/10.1037/0278-7393.16.3.417

Henderson, J. M., & Macquistan, A. D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics, 53,* 221–230. http://dx.doi.org/10.3758/BF03211732

Hiebert, E., & Pearson, P. D. (2010). *An examination of current text difficulty indices with early reading texts*. Santa Cruz, CA: TextProject, Inc.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., . . . Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science, 26,* 1071–1083. http://dx.doi.org/10.1177/0956797615581493

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science, 16,* 3–34.

Huston, A., & Wright, J. (1983). Children's processing of television: The informative functions of formal features. In J. Bryant & D. R. Anderson (Eds.), *Children's understanding of television: Research on attention and comprehension* (pp. 35–68). New York, NY: Academic Press.

Kabali, H. K., Irigoyen, M. M., Nunez-Davis, R., Budacki, J. G., Mohanty, S. H., Leister, K. P., & Bonner, R. L., Jr. (2015). Exposure and use of mobile media devices by young children. *Pediatrics, 136,* 1044–1050. http://dx.doi.org/10.1542/peds.2015-2151

Kamil, M., Inrator, S., & Kim, H. (2000). The effects of other technologies on literacy and literacy learning. In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 771–788). New York, NY: Routledge.

Kirkorian, H. L., Wartella, E. A., & Anderson, D. R. (2008). Media and young children's learning. *The Future of Children, 18,* 39–61. http://dx.doi.org/10.1353/foc.0.0002

Larson, A. L., & Rahn, N. L. (2015). Vocabulary instruction on Sesame Street: A content analysis of the Word on the Street initiative. *Language, Speech, and Hearing Services in Schools, 46,* 207–221. http://dx.doi.org/10.1044/2015_LSHSS-14-0079

Lesser, G. (1972). Learning, teaching, and television production for children: The experience of Sesame Street. *Harvard Educational Review, 42,* 232–271. http://dx.doi.org/10.17763/haer.42.2.h13nq18915q61rw6

Linebarger, D., & Piotrowski, J. (2010). Structure and strategies in children's educational television: The roles of program type and learning strategies in children's learning. *Child Development, 81,* 1582–1597.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary training on word learning: A meta-analysis. *Review of Educational Research, 80,* 300–335. http://dx.doi.org/10.3102/0034654310377087

Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139164603

Meringoff, L. (1980). A story a story: The influence of the medium on children's comprehension of stories. *Journal of Educational Psychology, 72,* 240–244. http://dx.doi.org/10.1037/0022-0663.72.2.240

Meringoff, L., Vibbert, M., Char, C., Fernie, D., Banker, G., & Gardner, H. (1983). How is children's learning from television distinctive? Exploiting the medium methodologically. In J. Bryant & D. Anderson (Eds.), *Children's understanding of television: Research on attention and comprehension* (pp. 151–180). New York, NY: Academic Press.

Miller, E., & Warschauer, M. (2014). Young children and e-reading research to date and questions for the future. *Learning, Media and Technology, 39,* 283–305. http://dx.doi.org/10.1080/17439884.2013.867868

Neuman, S. B. (1992). Is learning from media distinctive? Examining children's inferencing strategies. *American Educational Research Journal, 29,* 119–140. http://dx.doi.org/10.3102/00028312029001119

Neuman, S. B. (2009). The case for multimedia presentations in learning: A theory of synergy. In A. Bus & S. B. Neuman (Eds.), *Multimedia and literacy development: Improving achievement for young learners* (pp. 44–56). New York, NY: Taylor & Francis.

Neuman, S. B., Pinkham, A., Kaefer, T., & Strouse, G. (2014). Can babies learn to read? A randomized trial of baby media. *Journal of Educational Psychology, 106,* 815–830. http://dx.doi.org/10.1037/a0035937

Neuman, S. B., Wong, K., & Kaefer, T. (2017). Content not form predicts vocabulary and comprehension for low-income preschoolers. *Reading and Writing, 30,* 1753–1771. http://dx.doi.org/10.1007/s11145-017-9750-4

Nichols Linebarger, D., Brey, E., Fenstermacher, S., & Barr, R. (2017). What makes preschool educational television educational? A content analysis of literacy, language-promoting, and prosocial preschool programming *Media exposure during infancy and early childhood* (pp. 97–133). Cham, Switzerland: Springer.

Paivio, A. (2008). The dual coding theory. In S. B. Neuman (Ed.), *Educating the other America* (pp. 227–242). Baltimore, MD: Brookes.

Palmer, E., & Fisch, S. (2000). The beginnings of Sesame Street research. In S. Fisch & R. Truglio (Eds.), *"G" is for growing: 30 years of research on children and Sesame Street* (pp. 5–24). New York, NY: Routledge.

Paris, S. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40,* 184–202. http://dx.doi.org/10.1598/RRQ.40.2.3

Penno, J., Wilkinson, I., & Moore, D. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 2002,* 23–33. http://dx.doi.org/10.1037/0022-0663.94.1.23

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188–1200. http://dx.doi.org/10.1037/0096-1523.22.5.1188

Rice, M. L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development, 59,* 420–429. http://dx.doi.org/10.2307/1130321

Rideout, V. (2017). *The Common Sense Census: Media use by kids age zero to eight*. San Francisco, CA: Common Sense Media.

Rideout, V., & Katz, V. (2016). *Opportunity for all? Technology and learning in lower-class families. A report of the Families and Media Project*. New York, NY: Joan Ganz Cooney Center.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York, NY: Guilford Press.

Sharp, D., Bransford, J., Goldman, S., Risko, V., Kinzer, C., & Vye, N. (1995). Dynamic visual support for story comprehension and mental model building by young, at-risk children. *Educational Technology Research and Development, 43,* 25–42. http://dx.doi.org/10.1007/BF02300489

Smeets, D. J., & Bus, A. G. (2012). Interactive electronic storybooks for kindergartners to promote vocabulary growth. *Journal of Experimental Child Psychology, 112,* 36–55. http://dx.doi.org/10.1016/j.jecp.2011.12.003

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–406. http://dx.doi.org/10.1598/RRQ.21.4.1

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934–947. http://dx.doi.org/10.1037/0012-1649.38.6.934

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2014). Can the computer replace the adult for storybook reading? A meta-analysis on the effects of multimedia stories as compared to sharing print stories with an adult. *Frontiers in Psychology, 5,* 1366.

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A meta-analysis. *Review of Educational Research, 85,* 698–739. http://dx.doi.org/10.3102/0034654314566989

Vaala, S., Linebarger, D., Fenstermacher, S., Tedone, A., Brey, E., Barr, R., . . . Calvert, S. (2010). Content analysis of language-promoting teaching strategies used in infant-directed media. *Infant and Child Development, 19,* 628–648. http://dx.doi.org/10.1002/icd.715

Vaala, S., Ly, A., & Levine, M. (2015). *Getting a read on the app stores*. New York, NY: Joan Ganz Cooney Center.

Verhallen, M., & Bus, A. (2010). Low-income immigrant pupils learning vocabulary through digital picture storybooks. *Journal of Educational Psychology, 102,* 54–61. http://dx.doi.org/10.1037/a0017133

Verhallen, M., Bus, A., & deJong, M. (2006). The promise of multimedia stories for kindergarten children at risk. *Journal of Educational Psychology, 98,* 410–429. http://dx.doi.org/10.1037/0022-0663.98.2.410

Wasik, B., Bond, M. A., & Hindman, A. (2006). The effect of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology, 98,* 63–74.

Wright, T., & Cervetti, G. (2016). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly, 52,* 203–226. http://dx.doi.org/10.1002/rrq.163

Received July 10, 2017
Revision received February 19, 2018
Accepted February 22, 2018 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.