Journal of Children and Media

ISSN: 1748-2798 (Print) 1748-2801 (Online) Journal homepage: https://www.tandfonline.com/loi/rchm20

Children's attention to screen-based pedagogical supports: an eye-tracking study with low-income preschool children in the United States

Rachel M. Flynn, Kevin M. Wong, Susan B. Neuman & Tanya Kaefer

To cite this article: Rachel M. Flynn, Kevin M. Wong, Susan B. Neuman & Tanya Kaefer (2019) Children's attention to screen-based pedagogical supports: an eye-tracking study with low-income preschool children in the United States, Journal of Children and Media, 13:2, 180-200, DOI: 10.1080/17482798.2019.1575887

To link to this article: https://doi.org/10.1080/17482798.2019.1575887

Published online: 07 Feb 2019.

JOURNAL OF CHILDREN AND MEDIA
2019, VOL. 13, NO. 2, 180–200
https://doi.org/10.1080/17482798.2019.1575887

Routledge
Taylor & Francis Group

Children's attention to screen-based pedagogical supports: an eye-tracking study with low-income preschool children in the United States

Rachel M. Flynn[a], Kevin M. Wong[a], Susan B. Neuman[a] and Tanya Kaefer[b]

[a]Department of Teaching and Learning, New York University, New York, USA; [b]Department of Education, Lakehead University, Thunder Bay, Ontario, Canada

ABSTRACT

Educational screen media is increasingly salient in the lives of young children. Research affirms preschool-aged children can learn content from media when they attend to it, however less is known about how specific screen-based pedagogical supports (SBPS) might draw children's attention. Using eye-tracking methodology, the current study examines specific SBPSs that engage children's attention. The sample consisted of 106 3- to 5-year-olds from a poverty-impacted neighborhood. Participants viewed 12 video clips of Sesame Street that used four different SBPSs to support vocabulary: visual effects, visual + sound effects, explicit definitions, and explicit definitions + repetitions. Results indicated that children attended significantly more to the SBPSs with definitions. Findings also revealed differences in screen composition. Children attended more to people than objects, and attended more to on-screen conversations than conversations cut between screens. This study demonstrates the importance for educational media to use appropriate SBPSs and on-screen compositions to engage children.

ARTICLE HISTORY
Received 23 June 2017
Revised 24 January 2019
Accepted 25 January 2019

KEYWORDS
Video; television; media; attention; children; eye-tracking

Media is ubiquitous in the lives of young children around the world. It has become increasingly mobile and convenient to access with demonstrated benefits for learning across nations (Livingstone et al., 2017; Rideout, 2017). In the United States, preschoolers are watching over two and a half hours of content on various media platforms per day (e.g., television, mobile devices, computers) (Rideout, 2017), despite recommendations set by the American Academy of Pediatrics (2016) for 2- to 5-year-olds to view only one hour of high-quality screen media each day. The alarming amounts of media consumed by preschoolers may be attributed, in part, to parents who believe the content of programs benefit children and facilitate learning (Rideout, 2017). Given these trends in media use in this digital age, it is important for research to examine *how* children watch media and what they might learn from educational media programs.

Learning from screen media

Extensive research confirms that preschool-aged children can learn educational content, such as letters and numbers, from screen media (Crawley, Anderson, Wilder, Williams, &

CONTACT Rachel M. Flynn rachel.flynn@northwestern.edu New York University

Santomero, 1999; Kirkorian, Wartella, & Anderson, 2008; Linebarger, 2015; Linebarger & Walker, 2005; Wright et al., 2001). Two complementary theories supporting the idea that educational screen media can serve as a learning context for children are dual coding theory (Paivio, 2008) and the theory of synergy (Neuman, 2005, 2009). These two theories suggest that educational screen media may support learning by offering multiple sources and types of information to viewers. According to dual coding theory, information is processed both verbally and non-verbally (i.e., visual images) in parallel channels of the brain. When information is encoded both verbally and non-verbally, the interconnections between the two systems allow information to be processed more robustly than if it is encoded in separate channels. In addition, the theory of synergy asserts that multimedia presentations with visual and auditory effects can lead to a stronger mental representation of content on screen. Research frequently draws from these theories to investigate how multimedia presentations might lead to vocabulary learning among young children (Verhallen, Bus, & de Jong, 2006).

Relatedly, when presented with verbal and nonverbal information on screen, children can learn a wide range of topics such as science, math, history, or language. Across all content areas, media has the potential to provide rich learning experiences that build the vocabulary knowledge of viewers, which is particularly relevant as vocabulary may serve as the basis for conceptual development across subjects and disciplines (Neuman, Newman, & Dwyer, 2011). Children may develop an extensive understanding of new words and their meanings when presented with information in multiple ways on educational programs.

However, for preschool children to learn content from educational media, they must first attend to and comprehend the content (Anderson, Lorch, Field, & Sanders, 1981; Calvert, Huston, Watkins, & Wright, 1982; Fisch, Kirkorian, & Anderson, 2005; Lorch & Castle, 1997). On screens, there are specific features of educational media that can increase or decrease children's attention to content, which in turn can influence learning (Kirkorian & Anderson, 2008). For example, screen media that uses information that is irrelevant may distract children's ability to acquire new words and understand essential content (de Jong & Bus, 2004). Likewise, certain formal features and production techniques have the potential to support children's learning on screen (Calvert et al., 1982). To investigate how production techniques influence learning, early childhood research has generally used the method of looking at or away from screens as a measure of attention. Newer methods, such as eye-tracking technology, may provide more precise information about how young children view educational media, which could illuminate how specific aspects of screens influence children's visual attention (Anderson & Hanson, 2009).

Taken together, educational screen media is a vehicle for encouraging learning in the early childhood years, yet not all educational screen media is structured appropriately for learning (Vaala et al., 2010). For these reasons, to extend the theories of learning from media, the current study aims to use eye-tracking methods to examine specific screen-based pedagogical supports (SBPSs) that provide both visual and verbal sources of input to young learners. We seek to understand how these supports might differentially affect preschoolers' attention to screens, which ultimately can impact how children learn from educational screen media.

Attention to educational screen media

Early research on learning from television focused on how children processed or comprehended content. For example, while screen media may focus on teaching vocabulary to children, research has found that children must attend to the screen before they can learn the content (Anderson et al., 1981; Crawley et al., 1999; Kirkorian & Anderson, 2008). Television captures young children's attention through its formal features, such as cuts and pans, and visual and sound effects (Anderson et al., 1981; Calvert et al., 1982; Kirkorian et al., 2008). These formal features are suitable for supporting the presentation of vocabulary words with visual images (i.e., pictures or objects) or repeating the word throughout the segment; supports that can lead to increased vocabulary learning (Rice & Woodsmall, 1988). In addition, formal features in media are able to help children know what information to attend to, while auditory features re-engage inattentive viewers (Calvert et al., 1982).

In fact, certain on-screen attributes lead children to attend more to the screen than others (see Kirkorian & Anderson, 2008 for a review). For example, children attend more when characters have a conversation about something in the immediate context, and they attend less when the conversation is about something that happened in the past, in the future, or when there is no conversation on screen at all (Anderson et al., 1981). In contrast, research also demonstrates that non-verbal information can support learning on screens. Fisch, McCann Brown, and Cohen (2001) found that children can comprehend television stories in the absence of dialogue by relying on visual images and sound effects to interpret the meaning of the program. Therefore, as dual coding theory suggests, both verbal and non-verbal information are important to consider when understanding the relationship between attention and learning.

In addition, content that is interesting to children is more likely to capture their attention (Anderson & Kirkorian, 2015; Kirkorian & Anderson, 2008). For example, preschool children learn more from television shows like *Dora the Explorer* that actively engage and ask viewers to respond to prompts and questions than from shows that children view more passively (Anderson, Bryant, Wilder, Santomero, Williams, & Crawley, 2000; Crawley et al., 1999; Linebarger & Vaala, 2010; Linebarger & Walker, 2005). These interactive shows layer content through repetition and encourage children to participate with on-screen characters by asking direct questions and pausing for children to respond (Linebarger & Walker, 2005). In fact, preschool viewers of *Blue's Clues*, an interactive television show, performed significantly better than non-viewers on problem-solving and riddle tasks after repeatedly watching the show (Anderson et al., 2000).

Although general attention to television is associated with comprehension and learning, there is less known about which specific on-screen teaching supports might draw children's attention while watching educational episodes. There is research that highlights how information that is tangentially related to the topic, but irrelevant to the narrative or theme, distracts children and prevents comprehension and learning (Fisch et al., 2005). Therefore, examining children's attention to relevant versus irrelevant on-screen content has the potential to help researchers illuminate the process of learning and the type of screen media that best supports learning.

Screen-based pedagogical supports

Identifying the specific factors of educational screen media that effectively engage children's attention and, ultimately, facilitates learning has a number of important implications. For example, understanding these factors may be particularly relevant for children from poverty-impacted environments as educational screen media has the potential to boost learning outcomes, such as vocabulary and language skills, which are critical for later literacy development (Cunningham & Stanovich, 1997; Marulis & Neuman, 2013). Indeed, a key focus of many educational television programs is to teach vocabulary, language, and literacy skills to preschool children (Vaala et al., 2010). While a number of studies seek to understand how different pedagogical features of educational screen media might support early literacy skills (Larson & Rahn, 2015; Linebarger & Piotrowski, 2010; Piotrowski, 2014; Vaala et al., 2010), more recent studies are using innovative methods that employ eye-tracking technology to precisely document children's attention to pedagogical features on screen (Neuman, Wong, Flynn, & Kaefer, 2019). Focusing particularly on how these pedagogical features can build vocabulary knowledge and command the visual attention of low-income children has potential to reduce the disparity in vocabulary skills between children from different socioeconomic groups (Cunningham & Stanovich, 1997; Larson & Rahn, 2015; Linebarger & Piotrowski, 2010; Rescorla, Alley, & Christine, 2001).

In the current study, we sought to examine how specific screen-based pedagogical supports (SBPS) influenced children's visual attention during vocabulary teaching episodes. We drew from recent research that identified specific SBPSs used to provide vocabulary learning experiences to young viewers (Neuman et al., 2019; Larson & Rahn, 2015; Vaala et al., 2010; Wong & Neuman, 2019). In a recent content analysis of educational media programs (N = 4,500), Neuman et al. (2019) identified 11 different SBPSs that supported vocabulary learning for preschool children. Of these 11 supports, the four most prevalent SBPSs were: 1) visual images, 2) sound effects, 3) explicit definitions, and 4) repetition (See Table 1). Providing young viewers with intentional vocabulary learning experiences, the current study aimed to gauge how each of these pedagogical supports differentially impacted children's attention to screens.

Theory supports the four most prevalent SBPSs as suitable ways to promote vocabulary learning among young children. According to the theory of synergy (Neuman, 2009), the "visual images" SBPS facilitates vocabulary learning because visual images provide children with robust mental representations of objects that promote deeper word knowledge. Dual-coding theory (Paivio, 2008) also supports these SBPS because both non-verbal stimuli (i.e., visual images) and verbal stimuli (i.e., sound effects) together lead to stronger comprehension and information recall than when either support is used in isolation. Aligned with these theories, Vaala et al. (2010) found in a content analysis that videos often used these types of verbal and non-verbal strategies in infant-directed media.

Shifting to the third and fourth SBPS, studies document the importance of explicit definitions as they provide preschool children with clear and robust instruction that scaffolds vocabulary learning and reading comprehension (Beck, McKeown, & Kucan, 2013). Also, word repetition is an important contributor to high-quality vocabulary instruction because it maximizes children's exposure to a novel word (Coyne, Simmons, & Kame'enui, 2004;

Table 1. Screen-based pedagogical supports, vocabulary words and program episode.

SBPS	Vocabulary Word	Sesame Street DVD	DVD released	Clip Duration	Composition	Conversation
Visual Effects	Caterpillar	*Firefly Fun and Buggie Buddie*	2010	0:14	Object &Person	None
	Grater	*C is for Cooking*	2013	0:07	Object &People & Muppets	None
	Key	*Abby in Wonderland*	2008	0:10	Object &People & Muppets	None
Visual + Sound Effects	Airplane	*Letter of the Month Club*	2006	0:13	Object &Muppets	None
	Pumpkin	*P is for Princess*	2010	0:41	Object &People & Muppets	None
	Square	*Guess that Shape and Color*	2006	0:19	Object	None
Explicit Definition	Comfort	*Being Brave*	2013	0:26	People & Muppets	Cut Screen
	Dusk	*Firefly Fun and Buggie Buddie*	2010	0:16	Person	None
	Shelter	*Wild Words and Outdoor Adventures*	2011	0:17	People & Muppets	On Screen
Explicit Definition + Repetition	Athlete	*Be a Good Sport*	2014	0:40	Muppet	On Screen
	Hurricane	*Friends to the Rescue*	2005	0:38	Object &People & Muppets	Cut Screen
	Whisk	*C is for Cooking*	2013	0:27	Object &People & Muppets	Cut Screen

Penno, Wilkinson, & Moore, 2002). Although repeated exposure can support vocabulary learning among young children, research has yet to determine whether explicit definitions should be used in concert with this pedagogical support. Without a definition to support the repeated word, for example, children may notice a frequently used word, but not fixate their attention to it or understand its meaning. The current study examined differences in children's visual fixation when words were presented with these four SBPS – the four most commonly used instructional supports in children's media.

Examining screen composition

Beyond pedagogical supports, there are certain aspects of the screen that children attend to more than others (Anderson et al., 1981; Fisch et al., 2001; Kirkorian & Anderson, 2016). These varying aspects of screens are also known as screen composition, defined as the specific elements on screen that intentionally guide children's attention and scaffold learning. In line with dual-coding theory, one aspect of screen composition includes presenting visual and auditory stimuli to viewers. These stimuli may differentially draw children's attention to specific learning experiences in media by strategically using production techniques to capture children's attention (Vaala et al., 2010).

A second aspect of screen composition is when characters on screen engage in conversation with one another. Children appear to attend more to the screen when conversations take place – particularly when they are relevant and comprehensible – rather than when no conversations take place at all (Anderson et al., 1981; Fisch et al., 2001). Therefore, the current study also investigated how children fixated on characters, people or Muppets having a conversation relative to the amount of attention fixated to objects on the screen.

One final aspect of screen composition includes the use of cut screens, defined as a scene that takes place across the span of two different screen environments. In other words, as two characters have a conversation with one another, the camera does not pan smoothly from one character to the next, but cuts abruptly from the first character in the kitchen (screen environment #1) to the second character in the living room (screen environment #2). Based on Kirkorian and Anderson's (2016) work, which found that preschool children were slower to track objects across cut screens than adults, we investigated whether there were differences in visual fixation for conversations held on screen. The current study examined differences in children's visual fixation when characters had conversations in the same screen compared to conversations that cut across different screens.

Measuring attention while viewing educational screen media

In the past, to understand children's visual behavior on screens, research measured visual attention by examining how long children looked at the television screen and what was on the screen while they were attending (see Anderson et al., 1981; Calvert et al., 1982; Lorch & Castle, 1997; Pempek et al., 2010). These methods, while highly reliable and essential in shaping the educational screen media literature, do not allow for a precise interpretation of how children visually fixate on specific areas of the screen

(Kirkorian & Anderson, 2008). In response, Anderson and Hanson (2009) point out that new and innovative methods, such as eye-tracking technology, allow researchers to gain additional information and nuanced answers to traditional media research questions of how children learn from media.

While a number of studies in reading research use eye-tracking methods, there is less research that uses eye-tracking methods to examine screen media with children (Anderson & Hanson, 2009; Anderson & Kirkorian, 2015). Eye-tracking is a non-invasive methodology that permits high-resolution analyses of eye movement patterns. Tracking moment-to-moment changes in children's viewing behaviors while watching educational media enables a fine-grained examination of how screen-based supports might guide visual attention and the extent to which visual attention is related to educational outcomes. Because eye movement patterns are often compatible with cognitive understanding and knowledge (Thomas & Lleras, 2007), analyzing children's viewing behaviors may reveal additional information about how well they comprehend content on screen. Therefore, eye-tracking is an especially useful technique for studying children's online processing of educational media, which has been adopted by a number of recent studies that use eye-tracking to examine how young children watch educational media (e.g., Kirkorian & Anderson, 2016; Kirkorian, Anderson, & Keen, 2012; Neuman, Kaefer, Pinkham, & Strouse, 2014).

In particular, one study by Kirkorian et al. (2012) used eye-tracking methods to examine screen cuts, a formal feature, in media. They found that 4-year-old children and adults look at the center of the screen after a cut, which is optimal because it allows viewers to reorient their focus to changing content on the entire screen. Infants, on the other hand, had more variation in looking patterns after a cut on the screen, which shows a developmentally different viewing pattern for infants versus young children. Kirkorian and Anderson (2016) also used eye-tracking methods to examine if children anticipate scene transitions when objects cut off the screen and reappear on the opposite side of the screen. In their study, 4-year-old children were slower to track transitions and continued to look at the center of the screen, while adults' eye movements anticipated the object's movement. These eye-tracking studies help elucidate how preschool children respond to formal features while viewing. However, little research has used eye-tracking methods to examine how children's visual attention is influenced by specific content on the screen.

The current study uses eye-tracking methods to examine the specific learning features (i.e., screen-based pedagogical supports) that increase children's visual fixation. For example, by examining whether children look at characters or objects for a longer period of time, the current study allows for a deeper understanding of how malleable factors on-screen might facilitate visual attention and moderate word learning in young children.

Current study

Children's visual fixation to specific screen-based pedagogical supports and compositional features on screen may influence what they can learn from educational media. The current study aimed to explore how children looked at specific SBPSs during educational screen media viewing. There were four SBPSs used: visual images, visual

images + sound effects, explicit definitions, and explicit definitions + repetition, which are detailed further in the methods section. We used eye-tracking research methods to examine children's visual attention to SBPSs or certain on-screen compositions, such as characters or objects. While this study did not have a specific learning outcome, visual fixation to certain aspects of screen media is an important precursor to learning from media. By understanding the areas of the screen that low-income children visually fixate on while viewing, future studies can examine how specific supports and screen compositions might directly influence learning.

Research aims & hypotheses

The first aim was to investigate low-income children's visual fixation to certain SBPSs while watching educational media. Based on dual coding theory and the theory of synergy, we hypothesized that the SBPS that combined two supports (i.e., visual images + sound effects, explicit definitions + repetition), would hold children's attention longer than the other two supports (i.e., visual images only, explicit definitions only). The second research aim examined how long children looked at various on-screen compositions, characters, objects, and conversations. Based on prior research findings that children attend more when there are conversations on the screen than when there are no conversations (Anderson et al., 1981; Kirkorian & Anderson, 2016), we hypothesized that children would have a longer fixation-duration to characters than objects and on-screen conversations than cut-screen conversations.

Method

Participants

The study was conducted in two Head Start programs that provide free year-round preschool education to low-income children. All students qualified for free and reduced lunch. The centers were located in a poverty-impacted neighborhood in the northeast region of the United States. In total, twelve classrooms with 3- to 5-year-old children were invited to participate in the study. Teachers and parents provided written consent and children gave verbal assent. From these classrooms, 108 children were randomly selected, however, two students could not complete the study leaving 106 participants (44% female). Participant age ranged from 3 years 10 months to 5 years 6 months (M = 4.39; SD = 0.71). The two Head Start programs were in culturally diverse neighborhoods: 56% of the children were African-American, 38% were Hispanic, 1% White, and 7% Other. The sample also consisted of 45 English Learners (EL) (43%). Using a power calculator (Faul, Erdfelder, Lang, & Buchner, 2007), we determined that for a moderate effect, the sample would yield a two-tailed power of .85. A human subjects review board approved all aspects of the study.

Research design

To examine how children attended to SBPSs, we used a within-subjects design. In this type of design, each participant received all four SBPS conditions, and therefore, served as his/her own control. In this study, the within-subjects variable was the pedagogical

support used to teach a vocabulary word. There were three different vocabulary words in each SBPS condition resulting in 12 different video clips with 12 different words. The SBPS conditions were randomly ordered in three sequences to account for order effects and fatigue. Children were systematically assigned to one of these three sequences.

A within-subjects design was selected for many reasons. First, because students received all SBPS conditions, we were able to control for between-subjects variability. This reduced error and increased our power to detect potential differences between conditions. Second, threats to a carry-over effect were minimal since twelve different video clips were examined. Third, because participants essentially served as their own controls, a within-subjects design accounted for significant threats to internal validity.

Materials

Video clip stimuli

The twelve videos clips were selected from the children's television program, *Sesame Street* (2005–2013). While preschool children are accessing mobile devices to play interactive games that are educational or for entertainment more than ever before, watching television and videos remains the most common form of media for children ages 3 to 5 (Kabali et al., 2015). For this reason, the current study examined television as the media.

We chose the educational television show *Sesame Street* for three reasons. First, decades of research have used *Sesame Street* to examine children's ability to learn content, such as vocabulary, from screens (Larson & Rahn, 2015), particularly among culturally diverse populations which are reflected in our study's sample (Fisch & Truglio, 2001). Second, although *Sesame Street* is often catered to children slightly younger than the participants in our sample (i.e., 4-year-olds), children from poverty-impacted neighborhoods often have lower vocabularies than their peers, which makes *Sesame Street* an appropriate program to use. Third, it was necessary to choose one program for all clips to avoid effects of program. After examining many educational media shows for preschoolers that included the four SBPSs, *Sesame Street* provided clear exemplars of the SBPSs with a variety of vocabulary words to include.

The average video length was 21.42 seconds (SD = 8.77). A total of twelve video clips were used, with three clips representing each SBPS, for a total viewing time of 257 seconds. Information about the video clips including vocabulary word, SBPS condition, Sesame Street episode and clip duration is included in Table 1.

Screen-based pedagogical supports

The current study focused on if children visually fixated on the four SBPSs that were found to be the most commonly used supports in commercially-available educational screen media (Neuman et al., 2019). The following section breaks down each SBPS found in the video clips. Each SBPS had to focus on teaching only one vocabulary word.

Visual images

Video clips with visual images as a pedagogical support explicitly highlighted the vocabulary words using images or objects to promote acquisition. In addition, vocabulary words were isolated from other characters and objects in the composition of the screen. In our video sample, to teach the word "caterpillar," a character on Sesame Street had an image of a caterpillar propped on an easel. He then said, "Look, Dorothy, a caterpillar."

Visual images + sound effects

Video clips with this pedagogical support used sound in conjunction with a visual image as a tool to draw children's attention to the vocabulary word. For example, besides using visual images that might show an object depicting the vocabulary word, there is also a distinct sound that may draw viewers' attention to the word. In our video sample, to teach the word "pumpkin," a character on *Sesame Street* waved her wand around an object, which magically became a pumpkin. The camera then zoomed in on the pumpkin so that it took over the majority of the screen, and the outline of the pumpkin was covered in sparkles. A shimmering sound occurred simultaneous to the visual sparkles of the pumpkin outline. This SBPS was distinct from visual images on their own as the multimedia presentation may lead children to look at the screen for longer than an image alone, and the auditory features may also elicit attention from inattentive viewers (Calvert et al., 1982).

Explicit definition

Media clips with this pedagogical support used explicit definitions to teach vocabulary words. In other words, they intentionally stated the definition of a word in a clear and straightforward manner. In our video sample, to teach the word "shelter," a character on *Sesame Street* said a shelter is "a place where I can sleep. Where I can stay warm and dry and protected from the elements!"

Explicit definition + repetition

Media clips with this support used explicit definitions to teach vocabulary words, which are then repeated at least three times after the definition is given. Vocabulary words can also be repeated by the same, or multiple characters. In our video sample, to teach the word "hurricane," a character on *Sesame Street* said, "a hurricane is a very, very big storm with lots and lots of wind and rain." The word "hurricane" was then repeated six times throughout the segment. Repetitions ranged from 3 to 6 times across the three segments.

Measures

Eye-tracker

Eye-tracking technology was used to investigate the visual fixation of preschoolers while watching educational media. This innovative eye-tracking methodology was used to systematically examine children's visual attention when exposed to SBPSs during each video clip and assessment. Recent research using eye-tracking methods (Kirkorian & Anderson, 2016; Kirkorian et al., 2012) highlight developmental differences in on-screen

looking patterns by infants, 4-year-olds, and adults, which means the formal features used in educational media for preschool children should be age-appropriate and based on their viewing patterns.

Apparatus

To operationalize attention, eye movements were recorded using a Tobii Technology T120 eye-tracker integrated into a 17″ thin film transistor (TFT) monitor (Psychology Software Tools, Pittsburgh, PA). The sampling rate is typically 120 Hz, with a spatial accuracy of about 0.5 visual degrees. Using infrared diodes, the eye tracker generates reflection patterns on the corneas of each participant's eyes. Image sensors collect these reflection patterns, and other visual information about the participant to calculate a three-dimensional position of each eye and gaze point on screen. The TFT monitor utilizes active matrix technology with transistors that control each pixel on screen. This improves image quality and contrast relative to passive-matrix technologies. The TFT monitor also has a display resolution of 96 pixels per inch so that images are discernible.

The T120 eye tracker is a particularly appropriate apparatus for collecting data with young children (Kirkorian & Anderson, 2016; Kirkorian et al., 2012; Neuman et al., 2014). Using a binocular tracking method, this system allows for increased head movements. Typically, head movements result in a temporary accuracy error of about 0.2 visual degrees. For head movements that are especially active (i.e., over 25 cm/s), there is a 300-ms recovery period to full tracking ability. In addition, the system includes an embedded camera that records children's behavior and reactions to video clips and assessments. Calibration and stimulus materials are presented on the TFT monitor with Tobii Studio Professional 3.0 software.

Eye-tracking procedure

Preschoolers sat in a chair approximately 60 cm from the TFT monitor. While they received stimuli on the Tobii monitor, the researcher sat beside the child and viewed a second monitor. Tobii Studio Professional 3.0 software was used to present stimuli and process data.

To calibrate the gaze of preschoolers, participants were asked to follow an attention grabber to five points on the screen. A manual calibration procedure was used, which was monitored by Tobii Studio software and repeated when necessary. After calibration, a 2-second attention grabber appeared in the center of the screen at the beginning of each eye-tracking task. During each video clip, the researcher was able to follow the participants' eye movements and behaviors using the live viewer on the second monitor. Each cycle took approximately 25 minutes without breaks, including both familiarization and testing. If participants were agitated or restless, the screen was made blank, and they were given a short break. If a child was entirely noncompliant, the session was terminated.

Eye-tracking data processing

Using Tobii Studio Professional 3.0 software (Tobii Technology, Falls Church, VA), eye movement data were extracted for analysis. To process the data, areas of interest (AOIs) were first drawn manually around relevant stimuli (e.g., objects, characters) presented on the TFT monitor. AOIs were drawn at a distance subtending approximately 1.1 visual

degrees from the most protruded point of each stimulus. By isolating these AOIs, Tobii software (Colombo, Mitchell, & Horowitz, 1988) was able to calculate the amount of time spent looking at AOIs (e.g., total fixation-duration).

During recording, the eye-tracker collects raw movement data every 8.3 msec. Each data point is automatically identified by the software with a timestamp and (x, y) coordinates corresponding to a child's gaze at the given sampling moment, calculated using Tobii's pupil-centered corneal reflection technique. This information is sent to the Tobii Studio analysis application, which was used to extract information about gazes to (x, y) coordinates within given AOIs. The Tobii Studio fixation filter then grouped the raw eye movement data into fixations. Fixations are defined as any gaze coordinates lasting at least 60ms and located within 0.5 visual degrees, according to the algorithm set by the Tobii fixation filter. To help visualize data, fixations were overlaid onto a video recording of stimuli presented in each video clip. After fixation data was processed, we used Tobii Professional software to extract fixation data for each AOI for each child. Data were extracted to .txt files and then formatted to be compatible with the statistical software package.

Areas of interest

In each video clip, AOIs were drawn according to various compositional elements on screen that focused on teaching vocabulary words. By understanding where children looked at the screen, we were able to see how children specifically engaged with each video clip. We created variables that were proportional by dividing the time the child looked at the AOI to the total time of each AOI. AOIs were drawn according to the following two constructs:

Attention to characters and objects. Children learn vocabulary in screen media through robust representations of objects (e.g., a picture of a vocabulary word), and through characters and people (e.g., characters talking about a vocabulary word). To capture this, an AOI ("character") was drawn when a character was on screen defining a vocabulary word. Another AOI ("object") was drawn when an image of the vocabulary word appeared on screen after it had been introduced and defined. Video clips could have both types of AOIs if they used both types of composition. For example, for the word *hurricane* in the video clip, two AOIs would be drawn: one around the character that defined the word and one around the object depicting the hurricane. Table 1 illustrates which videos had characters or objects or both characters and objects.

Attention to conversations. To capture conversations in the composition of screen media, AOIs were drawn on characters who engaged in conversation with each other around the definition of a specific vocabulary word. This included two types of back-and-forth conversations: first, "on-screen conversations," which occurred when two or more characters were represented on the same screen, and engaging in conversation about the vocabulary word. Second, "cut-screen conversations," which occurred when characters were on different screens, where the camera would *cut* from one screen to the next as characters engaged in conversation about the vocabulary word. To be classified as a cut screen the conversation had to take place across two different scenes (see Table 1 for clips with cut screen conversations).

Visual fixation variables

To examine our first research question, to what extent does the type of SBPS influence how low-income preschool children watch educational media, we created composite visual fixation variables for each SBPS. These variables were created by first adding the fixation-duration time of all AOIs for each video, then dividing it by the video's length and finally averaging those proportions for each type of SBPS. This resulted in the following four SBPS fixation-duration variables that represented proportions: *Visual Fixation to Visual Images, Visual Fixation to Visual Images + Sound Effects, Visual Fixation to Explicit Definitions, and Visual Fixation to Explicit Definitions + Repetition*. The fixations to the composition AOIs were used to examine our second research question, under what conditions the composition of characters, objects, and conversations on screen influenced preschooler attention. Proportion variables were created by dividing each participant's fixation-duration to each object/character by the total time of that AOI (i.e., the time that object/character appeared on the screen). Only objects representing the vocabulary word or characters defining the vocabulary word were included in the AOIs, and therefore served as dependent variables. In addition, proportion variables were created for the conversation AOIs by dividing participant's fixation-duration to on-screen or cut-screen conversations by the total time of that AOI (i.e., the time that the conversation took place). These four composition variables were: *Visual Fixation to Objects, Visual Fixation to Characters, Visual Fixation to On-Screen Conversations, and Visual Fixation to Cut-Screen Conversations.*

Procedure

Three research assistants with Masters' degrees in education were trained to conduct the research. A scripted protocol was developed for one-on-one data collection with participants. Children were randomly selected from twelve classrooms to participate in the study. Each child participated in the study in one session for 25 minutes. Each child was escorted to a library to watch video clips on the eye-tracker. After calibrating gaze, participants watched twelve video clips featuring four SBPSs. Children were assigned to one of three sequences of video clips. Children were praised at the end of the study and escorted back to their classrooms.

Statistical analysis overview

Preliminary analysis revealed there were no differences by gender or age on the visual fixation variables; therefore these variables were not included in analyses. In addition, we examined if there were order effects based on the three different sequences, and found that there were no significant effects by order. For our primary analysis, we approached the data in two ways. First, to examine if children visually fixated for a longer duration on certain SBPSs, we used Repeated Measures Analysis of Variance (ANOVA) with the four SBPS attention variables as the within-subjects factor. Second, to examine the different types of compositions, we used paired samples *t*-tests to analyze children's visual fixation. We explored if children looked longer at characters or objects. Then, we explored if children looked longer to on-screen or cut-screen conversations.

Results

In the following results, we discuss the overall patterns of visual fixation towards various SBPSs and highlight how these pedagogical supports are represented in screen media. Afterward, we move to screen compositions and describe some of the features on the screen that children looked at for longer. It is important to note that visual fixation is based on a proportion, which allows for comparisons across clips of different lengths.

Children's visual fixation to the screen-based pedagogical supports

To examine our research question investigating children's visual fixation to certain SBPSs while watching educational media, we compared the proportion of time children looked at the different AOIs within the different SBPSs. Table 2 describes the means and standard deviations for the visual fixation variables for each specific SBPS. The supports that children looked at the longest were explicit definitions and explicit definitions + repetitions. These occurred when children looked at characters presenting clear definitions of words on screen. In fact, on average, children visually fixated on these supports four times more than supports that used visual images and 1.5 times more than visual images + sound effects.

Investigating further, we used Repeated Measures ANOVA to find that there were differences in looking time between the various SBPSs, as children looked significantly longer at certain supports, $F(3, 102) = 258.72$, $p < .001$, $\eta_p^2 = .713$. Children looked at the two SBPSs with definitions longer than the two SBPSs with visual images supporting vocabulary words (see Table 2). To examine the specific differences between these two groups of pedagogical supports, we used follow-up paired samples *t*-tests to examine the differences in fixation-duration by each of the SBPSs. Between definitions and visual images, we found that children looked longer at the relevant AOIs with explicit definitions than AOIs with visual images ($t(104) = 23.74$; $p < 0.001$) or visual images + sound effects ($t(104) = 11.25$; $p < 0.001$). Children also looked longer at the AOIs with definitions + repetition than visual images ($t(104) = 22.67$; $p < 0.001$) or visual images + sound effects ($t(104) = 10.46$; $p < 0.001$). Noting the discrepancy between these two SBPSs with visual images, we investigated the influence of visual images on looking time when sound effects were also included to reorient attention. Using paired samples *t*-tests, we noted that SBPSs with visual images + sound effects did result in greater fixation-duration than visual images on their own ($t(104) = 15.28$; $p < 0.001$). Finally, we turned to the two supports that included explicit definitions (i.e., explicit definition and definition + repetition), and found that there were no significant differences in fixation-duration between these two supports. Overall, children had a longer fixation-duration on relevant teachable moments (i.e., SBPSs) when characters provided definitions for the vocabulary words rather than visual images, with or without sound effects, to support the vocabulary word.

Table 2. Means and (standard deviations) for fixation to screen-based pedagogical supports.

Screen-based pedagogical support	Fixation-Duration proportion
Attention to visual effects	0.14 (0.69)
Attention to visual + sound effects	0.36 (0.16)
Attention to explicit definition	0.57 (0.20)
Attention to explicit definitions + repetitions	0.56 (0.21)

Children's attention to different screen compositions

To examine our second research question of how long children looked at the various compositions on screen, we examined if children had longer fixation-duration times for the different types of compositions. We investigated how long children looked at people and Muppet characters versus objects. In addition, we examined fixation-duration for on-screen conversations versus cut-screen conversations. In our analyses, only clips that had these features were included in analyses. Table 3 reports the means and standard deviations for each type of composition.

The different types of compositions mattered. Children had significantly longer fixation-duration for certain aspects of the screen. First, we noted differences in the looking time between objects representing the vocabulary word and people/Muppet characters when they discussed the vocabulary word. It appeared that, on average, children fixated on people/Muppet characters on screen for twice as long as they did to the object representing the word being taught. Paired samples *t*-tests confirmed that children looked significantly longer at people and Muppet characters than at objects ($t(104) = 18.61$; $p < 0.001$). As shown in Table 3, the average fixation-duration time on people and Muppet characters when they were on the screen was higher than the average fixation-duration time to objects.

Finally, examining conversations about vocabulary words, we noticed a slight discrepancy in visual fixation when children viewed characters having conversations on one screen (on-screen) versus when they were having conversations with the camera panning between two different screens (cut-screen). Using paired samples *t*-tests, we found that children fixated for longer during on-screen conversations than during cut-screen conversations ($t(104) = 2.93$; $p < 0.01$). Table 3 also shows the average fixation-duration time for the two types of conversations, where children fixated for longer when the conversations were on-screen.

Taken together, these findings indicated that children looked for longer when viewing the pedagogical supports that used explicit definitions, people/Muppet characters, and on-screen conversations. These results suggest that, perhaps, preschool children prefer to look at characters that actively present knowledge in clear and explicit ways. Moreover, on-screen conversations may provide a learning context that is less cognitively demanding than cut-screen conversations, as children do not need to reorient their attention to changing content on the entire screen.

Discussion

Research demonstrates that children can learn from educational screen media when they attend to and comprehend the content (Kirkorian et al., 2008). While a large body

Table 3. Means and (standard deviations) for fixation on compositions.

Composition	Fixation-duration proportion
Objects	0.22 (0.09)
People and Muppet characters	0.54 (0.21)
On-screen conversations	0.52 (0.23)
Cut-screen conversations	0.46 (0.18)

of research investigates these relationships, very few studies have used eye-tracking methods to precisely measure visual fixation to on-screen compositions and supports (Anderson & Hanson, 2009; Anderson & Kirkorian, 2015). Eye-tracking methods can confirm findings from other research that has examined attention. These methods can contribute new findings to the research on attention as they are more precise in capturing eye-movements and gaze. Future research should examine the relationship between attention to certain on-screen information and learning using eye-tracking methods.

The current study introduces the concept of screen-based pedagogical supports (SBPSs), which are grounded in dual coding theory and the theory of synergy. The categories of these SBPSs draws from two bodies of research: how children learn vocabulary and the formal features of television that capture attention. We used a within-subjects design to examine if the proportion of time that children visually fixated on certain SBPSs was more than other SBPSs. We found low-income preschool children had different looking patterns while viewing educational media. Our first hypothesis that children would visually fixate for longer when the SBPSs had multiple supports (i.e., explicit definition + repetition; visual images + sound effects) was partially supported. Children visually fixated for longer on the relevant teaching information when watching the clips with definitions (i.e., explicit definitions; definitions + repetition). They did not look for as long when vocabulary words were taught using visual images or visual images + sound effects. However, children did look longer at clips with the combined visual images and sound effects compared to visual images on their own. This may be because the sound effect oriented the viewer to the image, which engaged their attention, resulting in a longer fixation overall. This finding could offer a better understanding of how visual and auditory information can work together, extending dual-coding theory. In educational screen media having both types of information present can be helpful as the sound effects can direct attention to the relevant information.

It is somewhat surprising, given the previous research on children's increased attention to formal features (e.g., visual and sound effects), that those SBPSs did not lead children to look longer at the screen. At the same time, this may be explained by compositional screen features including the presence of characters, objects and conversations on-screen and cut between screens. Specifically, supporting our second hypothesis, children looked for longer at the people and Muppet characters on the screen talking about a vocabulary word than to objects representing the vocabulary word. This validates previous research that relevant on-screen dialogue engages children's attention. Anderson et al. (1981) found that children attended more to conversations between characters about the present situation than when there was no dialogue on the screen. Our study extends this research, showing that different types of conversations in educational media (i.e., on-screen vs. cut-screen conversations) may also have differential effects on attention to screen. Moreover, Wass and Smith (2015) found that toddler-directed programs used low-level design techniques to direct attention to relevant information on the screen to increase comprehensibility for young children. The authors recommended examining if on-screen characters provide cues to direct attention to relevant objects (Wass & Smith, 2015). In our research, the SBPSs of visual images, and visual images + sound effects may have relied too heavily on those formal features

as a strategy to hold children's attention. It is possible that children may look for longer at visual images if characters draw attention to the image and provide a definition of the word. Future research should continue to examine the inter-related use of these SBPSs in educational screen media and the impact on attention.

Finally, also supporting our second hypothesis, our research revealed that children looked for longer when the characters were having a conversation on-screen than when the conversations cut between two different screens. In some ways, this questions the findings from previous research that show formal features, such as cuts, are helpful to getting children to attend as they signal changes in screen content and re-orient attention (see Anderson & Kirkorian, 2015 for a review). However, using eye-tracking methods to carefully examine where the child is looking on screen helps to explain the differences in this finding. Previous research has used eye-tracking methods to examine screen cuts as a formal feature (Kirkorian & Anderson, 2016; Kirkorian et al., 2012). These studies found that four-year-old children were slower to transition between scene changes and that after a cut they focused on the center of the screen. In other words, children's visual fixation on relevant teaching information may be disrupted if they are viewing characters on-screen and then a cut occurs. After the cut, if they focus on the center of the screen before reorienting to look at the relevant characters, they may be losing information in the process. The formal feature of a cut may help inattentive viewers to re-orient, however, our research illustrates that they may not be the best feature for already attentive viewers. In particular, if the cut occurs during the middle of a scene when content is taught or discussed, then it has the potential to interrupt the learning process. Future eye-tracking research should examine if the best practice would be to teach relevant information without any cuts or scene transitions. Perhaps after the information has is explicitly taught, formal features can be used to re-capture inattentive viewers' attention or orient attentive children to upcoming changes.

Future research and limitations

This study contributes to the literature on how children view educational screen media, though it should be considered with its limitations in mind. First, while this study indicates that visual fixation varies by child and different SBPSs, it does not speak to whether increased attention leads directly to learning. Still, rather than focusing only on the product of learning (i.e., vocabulary outcomes), this study makes significant contributions to research that investigate the processes of learning (i.e., attention as a mediating process that might facilitate learning). Future studies may contribute to this body of work by examining attention as a potential moderating influence on the connection between screen-based pedagogical supports and word learning (Calvert et al., 1982; Kirkorian et al., 2012). Follow up studies may consider first asking children if they know any of the vocabulary words in the video clips to better understand how prior knowledge might influence children's attention. Although we selected video clips with words that were comparable in difficulty, children's attention to the different SBPSs may have been influenced by their prior experience with the vocabulary words rather than the type of support. Besides, this research provides information about children's viewing on-screen compositions in a short teaching moment; future research is needed to explore the features that sustain children's attention over longer scenes, entire episodes and seasons. This research examined differences in looking time while viewing educational video (i.e., television or

DVD). However, future studies should consider using eye-tracking methods to examine moment-to-moment changes in attention while preschool children play interactive games and use applications on touch screen devices. The medium (i.e., video vs. interactive game) that the educational content is presented on may impact what children look at, and ultimately learning. Finally, the current research used commercially-available videos from a popular and evidence-based television show. Using these videos as stimuli has several limitations. First, children may have prior experience with the show and may or may not enjoy it. Second, using commercially-available videos means that we were limited in the words available and screen compositions. Future research should manipulate video to create an ideal support with different vocabulary words. Manipulating video would also allow for relatively precise identification of the specific features that children fixate on while viewing. For example, the same video could be layered where some clips have a visual image of the object alone and then the second condition would use the same clip, but add a sound effect. Finally, choosing from commercially-available videos means that there are perceptual differences across the clips (i.e., more movement or more colors) (Aslin, 2007). Future research should create or manipulate videos to carefully control for any perceptual differences in the videos. Finally, our sample was drawn from a poverty-impact neighborhood and from a center with children who predominantly receive free and reduced lunch. While this is a strength of the study, as this population is often under-researched, it also limits generalizability.

Conclusion

The current study is one of very few eye-tracking studies that examine what children look at while viewing educational screen media. Aligned with prior research, this study confirms that children look longer when there are on-screen conversations about immediate and relevant information. Our research sheds new light on the body of literature that examines how formal features increase children's looking time, as it reveals that in some situations, formal features may actually disrupt a learning context, which may have implications for researchers interested in children's educational media. Finally, this study highlights the importance of using high-quality SBPSs, such as providing explicit definitions and repeating vocabulary words, to help children attend to relevant information on the screen. Educational media often relies on its form through visual images and sound effects. This study suggests that screen media may rely too heavily on such form. Still, it is important to consider these findings in context of preschool children who view large amounts of screen-based media, especially because much of the content they view claims to be educational. Despite the body of research on children learning from educational screen media, there are still open questions about what defines high-quality educational screen media. Findings from this study reveal the importance of using eye-tracking methods to determine some of the mechanisms in screen media content and composition that effectively hold children's attention.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Institute of Education Sciences [R305A150143].

Notes on contributors

Rachel M. Flynn is a Research Assistant Professor in the Department of Medical Social Sciences at Northwestern University. She received her PhD in Developmental Psychology from the University of California, Riverside and completed her postdoctoral training at New York University. Her primary research examines media's impact on children's cognitive development. She is interested in studying individual differences factors, such as prior exposure, enjoyment, and attention, that differentially impact media effects.

Kevin M. Wong is a PhD candidate in the Department of Teaching and Learning at New York University, specializing in early literacy and multilingual education. His research examines pedagogical supports that promote L1 and L2 vocabulary development among dual-language learners, including the use of educational media. His work related to educational media has appeared in *Bilingual Research Journal, Reading and Writing*, and the *Journal of Educational Psychology*.

Susan B. Neuman is a Professor of Teaching and Learning at New York University specializing in teacher education and early literacy development. Her research and teaching interests include early childhood policy, curriculum, and early reading instruction, prek-grade 3 for children who live in poverty. Neuman has received two life-time achievement awards for research in literacy development and is a Fellow of the American Educational Research Association.

Tanya Kaefer is an Associate Professor in the Faculty of Education at Lakehead University in the field of Educational Psychology. She studies the development of children's knowledge and the cognitive processes involved in early learning. Specifically, she examines the sources of children's knowledge - experiences, people, books and media - and how this knowledge may influence future learning and development.

ORCID

Rachel M. Flynn http://orcid.org/0000-0002-2522-2932
Kevin M. Wong http://orcid.org/0000-0001-6237-0427

References

American Academy of Pediatrics. (2016). Council on communications and media: Media and young minds. *Pediatrics*, *138*, 5.

Anderson, D. R., & Kirkorian, H. L. (2015). Media and cognitive development. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science, 7th edition, Vol. 2: Cognitive processes* (pp. 949–994). Hoboken, NJ: Wiley.

Anderson, D. R., Bryant, J., Wilder, A., Santomero, A., Williams, M., & Crawley, A. M. (2000). Researching Blue's Clues: Viewing behavior and impact. *Media Psychology*, *2*(2), 179–194.

Anderson, D. R., & Hanson, K. G. (2009). Children, media, and methodology. *American Behavioral Scientist*, *52*(8), 1204–1219.

Anderson, D. R., Lorch, E. P., Field, D. E., & Sanders, J. (1981). The effects of TV program comprehensibility on preschool children's visual attention to television. *Child Development*, *52*, 151–157.

Aslin, R. N. (2007). What's in a look?. *Developmental Science*, *10*, 48–53.

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford Press.

Calvert, S. L., Huston, A. C., Watkins, B. A., & Wright, J. C. (1982). The relation between selective attention to television forms and children's comprehension of content. *Child Development*, *53*, 601–610.

Colombo, J., Mitchell, D. W., & Horowitz, F. D. (1988). Infant visual attention in the paired-comparison paradigm: Test-retest and attention-performance relations. *Child Development*, *59*(5), 1198–1210.

Coyne, M. D., Simmons, D. C., & Kame'enui, E. J. (2004). Vocabulary instruction for young children at risk of experiencing reading difficulties: Teaching word meanings during shared storybook readings. In J. F. Bauman & E. J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 41–58). New York, NY: Guilford Press.

Crawley, A. M., Anderson, D. R., Wilder, A., Williams, M., & Santomero, A. (1999). Effects of repeated exposures to a single episode of the television program blues clues on the viewing behaviors and comprehension of preschool children. *Journal of Educational Psychology*, *91*(4), 630–637.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–935.

DeJong, M., & Bus, A. (2004). The efficacy of electronic books in fostering kindergarten children's emergent story understanding. *Reading Research Quarterly*, *39*, 378–393.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fisch, S., Kirkorian, H. L., & Anderson, D. R. (2005). Transfer of learning in informal education: The case for television. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 371–393). Charlotte, North Carolina: Information Age Publishing.

Fisch, S. M., McCann Brown, S. K., & Cohen, D. I. (2001). Young children's comprehension of educational television: The role of visual information and intonation. *Media Psychology*, *3*(4), 365–378.

Fisch, S. M., & Truglio, R. T. (Eds.). (2001). *G is for growing: thirty years of research on children and sesame street*. New York, NY: Routledge.

Kabali, H. K., Irigoyen, M. M., Nunez-Davis, R., Budacki, J. G., Mohanty, S. H., Leister, K. P., & Bonner, R. L. (2015). Exposure and use of mobile media devices by young children. *Pediatrics*, *136*(6), 1044–1050.

Kirkorian, H. L., & Anderson, D. R. (2008). Television and learning. In S. Neuman (Ed.), *Educating the Other America* (pp. 301–318). Baltimore, MD: Paul H. Brookes Publishing.

Kirkorian, H. L., & Anderson, D. R. (2016). Anticipatory eye movements while watching movement across shots in video sequences: A developmental study. *Child Development*. doi:10.1111/cdev.12651

Kirkorian, H. L., Anderson, D. R., & Keen, R. (2012). Age differences in online processing of video: An eye movement study. *Child Development*, *83*(2), 487–507.

Kirkorian, H. L., Wartella, E. A., & Anderson, D. R. (2008). Media and young children's learning. *The Future of Children*, *18*(1), 39–61. Retrieved from https://www.princeton.edu/futureofchildren/publications/docs/18_01_03.pdf

Larson, A., & Rahn, N. (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.

Linebarger, D. L. (2015). Super Why! to the rescue: Can preschoolers learn early literacy skills from educational television? *International Journal for Cross-Disciplinary Studies in Education*, *6*(1), 2060–2068. Retrieved from http://infonomics-society.org/wp-content/uploads/ijcdse/published-papers/volume-6-2015/Super-Why-to-the-Rescue.pdf

Linebarger, D. L., & Piotrowski, J. T. (2010). Structure and strategies in children's educational television: The role of program type and learning strategies in children's learning. *Child Development*, *81*(5), 1582–1597.

Linebarger, D. L., & Vaala, S. E. (2010). Screen media and language development in infants and toddlers: An ecological perspective. *Developmental Review*, *30*, 176–202.

Linebarger, D. L., & Walker, D. (2005). Infants' and toddlers' television viewing and language outcomes. *American Behavioral Scientist*, *48*(5), 624–645.

Livingstone, S., Lemish, D., Lim, S. S., Bulger, M., Cabello, P., Claro, M., … Nayar, P. (2017). Global perspectives on children's digital opportunities: An emerging research and policy agenda. *Pediatrics*, *140*(Supplement 2), S137–S141.

Lorch, E. P., & Castle, V. J. (1997). Preschool children's attention to television: Visual attention and probe response times. *Journal of Experimental Child Psychology*, *66*, 111–127.

Marulis, L. M., & Neuman, S. B. (2013). How vocabulary interventions affect young children at risk: A meta-analytic review. *Journal of Research on Educational Effectiveness*, *6*(3), 223–262.

Neuman, S. B. (2005). *Handbook of research on teaching literacy through the communicative and visual arts*. Eds. NY: Routledge.

Neuman, S. B. (2009). The case for multimedia presentations in learning. 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., Kaefer, T., Pinkham, A., & Strouse, G. (2014). Can babies learn to read? A randomized trial of baby media. *Journal of Educational Psychology*, *106*(3), 815–830.

Neuman, S. B., Newman, E. H., & Dwyer, J. (2011). Educational effects of a vocabulary intervention on preschoolers' word knowledge and conceptual development: A cluster-randomized trial. *Reading Research Quarterly*, *46*(3), 249–272.

Neuman, S. B., Wong, K. M., Flynn, R. M., & Kaefer, T. (2019). Learning vocabulary from educational media: The role of pedagogical supports for low-income preschoolers. *Journal of Educational Psychology*, *111*(1), 32–44.

Paivio, A. (2008). The dual coding theory. In S. B. Neuman (Ed.), *(2008). Educating the other America: Top experts tackle poverty, literacy, and achievement in our schools* (pp. 227–242). Baltimore, MD: Paul H. Brookes Publishing Company.

Pempek, T. A., Kirkorian, H. L., Richards, J. E., Anderson, D. R., Lund, A. F., & Stevens, M. (2010). Video comprehensibility and attention in very young children. *Developmental Psychology*, *46*(5), 1283–1293.

Penno, J. F., Wilkinson, I. A., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *94*(1), 23.

Piotrowski, J. T. (2014). The relationship between narrative processing demands and young American children's comprehension of educational television. *Journal of Children and Media*, *8*(3), 267–285.

Rescorla, L., Alley, A., & Christine, J. B. (2001). Word frequencies in toddlers' lexicons. *Journal of Speech, Language, and Hearing Research*, *44*(3), 598-609. doi:10.1044/1092-4388(2001/049)

Rice, M. L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development*, *59*(2), 420–429.

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

Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review*, *14*(4), 663–668.

Vaala, S. E., Linebarger, D. L., Fenstermacher, S., Tedone, A., Brey, E., Barr, R., … Calvert, S. (2010). Content analysis of language-promoting teaching strategies: Do strategies match the claims in infant directed media? *Infant and Child Development*, *19*(6), 628–648.

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

Wass, S. V., & Smith, T. J. (2015). Visual motherese? Signal-to-noise ratios in toddler-directed television. *Developmental Science*, *18*(1), 24–37.

Wong, K. M., & Neuman, S. B. (2019). Learning vocabulary on screen: A content analysis of pedagogical supports in educational media programs for dual-language learners. *Bilingual Research Journal*, *42*, 1.

Wright, J. C., Huston, A. C., Murphy, K. C., Peters, M. S., Pinon, M., & Scantlin, R. M. (2001). The relations of early television viewing to school readiness and vocabulary of children from low income families: The early window project. *Child Development*, *72*, 1347–1366.