Journal of Educational Psychology

Scaffolding Attention and Partial Word Learning Through Interactive Coviewing of Educational Media: An Eye-Tracking Study With Low-Income Preschoolers

Susan B. Neuman, Preeti Samudra, Kevin M. Wong, and Tanya Kaefer

Online First Publication, October 31, 2019. http://dx.doi.org/10.1037/edu0000428

* 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.

American Psychological Association

Journal of Educational Psychology

ISSN: 0022-0663

http://dx.doi.org/10.1037/edu0000428

Scaffolding Attention and Partial Word Learning Through Interactive Coviewing of Educational Media: An Eye-Tracking Study With Low-Income Preschoolers

Susan B. Neuman, Preeti Samudra, and Kevin M. Wong
New York University

Tanya Kaefer
Lakehead University

This study was designed to examine the effects of coviewing on low-income children's attention to and understanding of novel words in educational media. In addition, we sought to understand coviewing's contribution to children's receptive and expressive word learning when some target words were repeated more or less frequently. Using a within-subjects design, 83 preschoolers viewed 2 educational media stories, 1 with an adult coviewer, and the other without, in a counterbalanced approach. Eye-tracking technology recorded children's attention throughout viewing; pre- and posttests examined children's gains in receptive and expressive word identification. Results indicated that children's attention to target words was greater in the coviewing condition but appeared to contribute to expressive word learning only of lower repetition words. Attention mediated the relationship between coviewing and low-repetition word learning for expressive, but not receptive, vocabulary. Regardless of condition, children learned more words when they were repeated more frequently. This study provides further evidence that low-income children can pick up at least partial word knowledge on their own, particularly when words are repeated frequently.

Educational Impact and Implications Statement
Numerous policymakers have recommended adult coviewing of educational media to enhance young children's learning. This study focuses on its potential to enhance low-income preschoolers' word learning in programs, some of which were repeated more frequently than others. Results of our study suggests that coviewing's contribution was limited to situations when the word repetition was low; when words were repeated frequently, children seemed to pick up partial word knowledge on their own. Taken together, this research highlights both the features of educational media and the social supports that might contribute to low-income children's language learning.

Keywords: educational media, language, preschoolers, coviewing

Children learn words through educational media (Linebarger & Piotrowski, 2010). Viewed most frequently on mobile devices, educational media are programs deliberately and systematically designed to enhance children's school readiness and academic development (Rideout, 2017). Studies have shown that young preschoolers are able to engage in rapid, online processing of words while viewing such educational programs like *Martha Speaks*, picking up at least a partial understanding of these words in video contexts (Linebarger, Moses, Garrity Liebeskind, & McMenamin, 2013; Rice & Woodsmall, 1988). Furthermore, studies suggest that children can do so on their own, with minimal adult support. Takacs, Swart, and Bus (2014), for example, in a recent meta-analysis of 29 studies, found no significant differences in children's learning outcomes between viewing multimedia stories and sharing traditional print-like stories with an adult. According to these researchers (Bus, Takacs, & Kegel, 2015), such multimedia features like animated illustrations, background music, and sound effects may provide similar scaffolding for story comprehension and word learning as an adult.

Nevertheless, not all young children pick up words so effortlessly. For example, studies have shown that low-income children are likely to seriously lag behind their middle-class peers in vocabulary and oral language comprehension (Morgan, Farkas, Hillemeier, Hammer, & Maczuga, 2015). Research has documented a clear relation between socioeconomic status (SES), par-

Susan B. Neuman, Preeti Samudra, and Kevin M. Wong, Department of Teaching and Learning, New York University; Tanya Kaefer, Department of Educational Studies, Lakehead University.

Preeti Samudra is now at the Department of Psychology, State University of New York.

Correspondence concerning this article should be addressed to Susan B. Neuman, Department of Teaching and Learning, New York University, 239 Greene Street, New York, NY 10003. E-mail: sbneuman@nyu.edu

ticularly parent education and family income, and children's vocabulary development (Hart & Risley, 1995; Rowe, 2018). As early as 18 months, studies have documented striking differences in vocabulary and language processing efficiency for these economically disadvantaged children (Fernald, Marchman, & Weisleder, 2013; Halle et al., 2009); by 24 months, there is a 6-month gap compared with their more advantaged peers. Even more troubling, evidence from a number of longitudinal studies suggests that once behind (Cunningham & Stanovich, 1997; Juel, 1988), these children are likely to stay behind in vocabulary development, reading, and later academic achievement.

However, although many of these studies have shown stark differences in language input between middle- and lower income groups, few have reported on the potential variation within SES groups, particularly children who come from low-income groups. For example, a substantial portion of these studies have categorized low-income children as if they comprise one homogenous group, making it difficult to detect important within-group variability. Yet recent studies have documented large variation in the amount and lexical diversity of talk within low-income groups (Rowe, Pan, & Ayoub, 2005). In a recent study examining the ecocultural patterns of family engagement among low-income Latino families of preschool children (McWayne, Melzi, Limlingan, & Schick, 2016), for example, researchers found evidence of heterogeneity in patterns of family engagement within group, which related to practices associated with school readiness and children's language skills. This variability is often obscured in cross-group comparisons.

Therefore, applying a within-group lens could help to inform instructional practices and recommendations for promoting vocabulary for children who come from low-income circumstances. Certain formal features in educational media, such as animation, sounds, and music, for example, may hold some children's attention in learning new words. For example, Verhallen and Bus (2010) found that the second language learners (L2) children in their low-income sample seemed to especially benefit from digital storybooks compared with books read with static images. Yet in other research among low-income children, there is some initial evidence that the viewing of educational media might actually exacerbate the gap rather than close it (Neuman & Celano, 2006). Studies have shown that children with stronger vocabularies tend to learn more words than those with weaker vocabularies (Blewitt & Langan, 2016). In a recent study, for example, low-income Head Start children with slightly higher vocabulary scores used the pedagogical features in the educational media programs to their advantage, identifying more novel words in and out of context than their lower language peers. Unfortunately, neither ostensive nor attention-directing cues appeared to exert additional support for children with lower receptive language scores (Neuman, Wong, Flynn, & Kaefer, 2019). Subsequent studies (Samudra, Wong, & Neuman, in press), adjusting the pacing of educational programs, or providing definitional cues (Korat, Levin, Atishkin, & Turgeman, 2014) have shown only modest improvements in word learning.

Consequently, recognizing the variability within a low-income sample, some children are likely to need additional supports to accelerate their vocabulary development. And here, there are two likely candidates to provide such targeted assistance. The first includes the contextual support of an adult who may directly influence how a child views and makes meaning from a program. Recommended by the American Academy of Pediatrics (2016), coviewing may support learning through adult–child interaction while watching a program together. The second likely support includes word repetition—the number of times the word is actually used throughout the program. In their classic study, Rice and Woodsmall (1988), for example, theorized that the repetition of novel words (e.g., five to six repetitions), coupled with a depiction of the word's meaning, largely accounted for gains in word knowledge. Providing additional repetitions and recasts (e.g., repetitions in similar but not identical grammatical contexts), therefore, might be a prime candidate for increasing children's vocabulary.

In this study, we examine these potential supports and how they might contribute to low-income children's word learning. Specifically, our first objective was to determine the extent to which each of these supports independently might enhance children's attention to, and understanding of, novel words. Our second objective was to examine how these supports may interact to potentially bolster children's vocabulary. Together, our goal was to better understand the contextual and instructional design features of educational media that might bootstrap young children's vocabulary development.

The Potential of Coviewing

Coviewing typically refers to members of a household watching TV or a video together (Takeuchi & Stevens, 2011). Yet the term itself can have many different guises. For example, in one of the earliest studies, Salomon (1977) found that parent–child coobserving of Sesame Street seemed to have an affective influence on the lower income children's viewing but not for those in the middle class, which generated greater skills and comprehension of the program. Simply being present, Salomon hypothesized, might have targeted children's attention to the screen, resulting in improved performance.

In contrast to simply being there, however, several other studies have examined a more active mediational approach. Using questioning techniques and contingent feedback, Reiser, Tessmer, and Phelps (1984) found that 3- and 4-year-old children were more likely to identify letters and numbers while coviewing than when viewing with a silent adult. Coviewers asked the child to name the letters and numbers while viewing Sesame Street and gave contingent feedback throughout the program. Presumably, the questioning and feedback drew children's attention more deliberately to the screen. Even more prescriptive, Strouse, O'Doherty, and Troseth (2013) reported on the effects of a coviewing intervention that trained parents to pause a video and engage in dialogic questioning (e.g., open-ended questions) with their child. Among other comparisons, they compared the dialogic approach with one in which parents also paused the video but merely directed children's attention to the screen. Their results indicated that children in the dialogic group significantly outscored those in the other groups in vocabulary and comprehension, indicating that *what* parents did during the active mediation mattered more than simply directing children's attention.

Consequently, coviewing might support children's learning by drawing their attention to the screen, helping them to focus on the most important aspects of the program, and by extending the lessons presented in the program. It might also serve to guide children in more active viewing through comments and questions, enhancing the comprehensibility of the words and their meanings (Hirsh-Pasek et al., 2015). Furthermore, the interactive features of dialogic coviewing—asking open-ended questions and providing feedback may serve both a pragmatic and didactic function that fosters language development. Children not only learn words from other people but also make efforts to determine their communicative intentions.

Word Repetition

Coviewing might also support word repetition. For example, the dialogic questions in Strouse et al.'s (2013) study often required children to use story-specific vocabulary, repeating what they had heard in the program. Word repetition, in similar but not identical contexts, is known to support vocabulary development in print (Stahl, 2003) and screen media (Verhallen & Bus, 2010). Although no ideal number of repetitions has been empirically derived, much of the research suggests that a greater number of encounters improves children's ability to recall and comprehend them (Stahl & Nagy, 2006). In fact, McKeown, Beck, Omanson, and Pople (1985) found that although four encounters with a word did not reliably improve comprehension, 12 encounters did. Exposed to words repeated in multiple contexts (Biemiller & Boote, 2006), children began to learn more about those words than in a single context (Stahl, 2003).

There are a number of studies that have used repetition to their advantage, particularly among low-income children who might need additional exposure to novel words in multiple contexts. Verhallen and Bus (2010), for example, found that repeated exposure to a digital storybook (four times), presented with either static or video images, significantly improved vocabulary learning for low-income children, and that the video condition resulted in greater gains for expressive language compared with the static condition. Similarly, Linebarger and colleagues (2013) found that repeated exposure of a program (e.g., 5 words × 5 times) significantly predicted gains in expressive vocabulary for low-income children compared with working-class children, who did not show additional gains from repeated exposure.

Word repetition, therefore, might also support vocabulary development. However, there is some evidence that word repetition might have differential effects on the outcome variables measured. For example, Linebarger and colleagues (2013) reported gains among low-income children as a result of repeated exposure for expressive language but not for receptive language. Similar to Whitehurst et al.'s (1994) classic studies on dialogic reading, Strouse et al.'s (2013) study of dialogic coviewing also found gains particular to expressive language. On the other hand, repeated digital reading in Verhallen and Bus's (2010) research bolstered both receptive and expressive vocabulary, although children learned more words expressively than receptively. Showing similar differential effects with printed texts, Sénéchal and Cornell (1993; Sénéchal, 1997) have argued that the processes of acquisition of these two types of vocabulary might be different. It might be, for example, that a single exposure of words is sufficient for receptive language but that multiple exposures, as noted in the previous studies, are most beneficial for expressive language. According to these and other researchers, therefore, words should be assessed both receptively and expressively in order to better estimate the effects of repetition on word learning.

Similarly, coviewing might also have differential effects for word learning. For example, coviewing might help scaffold children's attention to target words that are not often repeated in programs; in cases in which the target words are frequently repeated in multiple contexts, coviewing might have negligible effects on attention. Moreover, given that children seem able to identify at least a portion of words based on a single exposure, the added value of coviewing might only be evident in expressive language and not receptive language. Sénéchal (1997), for example, found that the interactive techniques between adults and children in repeated storybook readings were more helpful in acquiring expressive than receptive vocabulary.

Therefore, this study was designed to examine the potential of coviewing on low-income children's attention to, and understanding of, novel words in educational media. In these media stories, some of the target words include many repetitions and recasts (eight to nine times), whereas in others, much less so (three to four times). Using eye-tracking technology, our goal was to understand how coviewing might differentially affect children's attention to words that were repeated at different frequencies and its effects on gains in children's receptive and expressive vocabulary. Specifically, we addressed the following questions: (a) How does coviewing affect low-income children's attention to novel words? Are there differences in attention based the number of word repetitions?; (b) How does coviewing affect gains in receptive and expressive vocabulary? Are there differential effects based on word repetitions?; and (c) Might attention mediate the associations between coviewing and receptive and expressive vocabulary?

Method

Participants

We recruited 83 preschoolers (M_{age} 4.3 years, SD = .37) from two Head Start centers located in high-poverty areas in a large urban city. Educational directors, teachers, and parents provided consent for participation. Children provided verbal assent. The sample was diverse: 29% African American, 49% Hispanic, 18% West Indian, and 4% Asian or biracial; 55% were female. All children qualified for free-and-reduced lunch. Standardized receptive language skills, measured by the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007) averaged 79.64 (SD = 15.76), more than one standard deviation below the norm.

Research Design

We used a within-subjects design to examine the effects of coviewing educational media stories on children's word learning. In this within-subjects design, each child viewed two video stories, one with an adult coviewer and the other without, in a counterbalanced approach. Condition order (e.g., coviewing vs. no coviewing) and the specific video (e.g., plants, shapes) used in each condition were counterbalanced between participants to ensure results were not tied to order or the specific video. Word learning in each condition was compared for each individual participant. In both conditions, children viewed the video on a computer equipped with eye-tracking technology to examine their attention throughout the programs.

There were a number of benefits in using this design. First, because each child received both treatments, we were able to control for between-subjects variability, reducing error and increasing our power to detect differences. And, second, within-subject designs may control for threats to internal validity because the participants essentially act as their own controls.

Digital Stories

We selected two full-length (9.5 min) narrative stories from the educational media program *Peep and the Big Wide World* (produced by WGBH, 2004). Designed for preschoolers, the cartoon characters—Peep, a newly hatched chicken and his pals, Chirp and Quack—go on weekly adventures, learning science concepts throughout their travels.

One story episode focused on plants and another focused on shapes. To measure how word repetition might affect word learning, we replaced the audio track of both programs with an adapted script that incorporated eight vocabulary words per episode. Following the plot line of the original scripts, actors (e.g., graduate students from the educational theater program) performed the voiceovers of the characters and the narrator in the new scripts. Half of the words in each video were repeated at a lower rate (three to four times), and the other half, at a higher rate (eight to 11 times). Words were nouns, clearly depicted on the screen at least three times during the video.

We selected words regarded as Tier II (e.g., words that children are later likely to encounter across all topics; Beck, McKeown, & Kucan, 2002). To heighten the likelihood that children would not already be familiar with these words, however, we also examined target words on ChildFreq, a database that shows the frequency of word occurrences by children's age from transcripts in the CHILDES database (MacWhinney, 2000). As shown in Table 1, all words were likely to be unfamiliar to children at this age level.

Table 1
Target Word Characteristics

Episode	Word	Repetitions in video	ChildFreq occurrences/1,000,000 words
Shapes	1. Pyramid	8	0
	2. Cube	9	7
	3. Corner	9	48
	4. Acorn	8	0
	5. Cone	4	28
	6. Beaver	4	11
	7. Dam	3	0
	8. Raccoon	3	4
Seed	1. Stream	8	3
	2. Sunflower	9	7
	3. Seed	9	27
	4. Soil	8	0
	5. Stem	4	20
	6. Petal	4	10
	7. Bud	3	12
	8. Seedling	3	0

Note. ChildFreq = word frequency of child's language from a large corpus of words in the CHILDES database.

Coviewing Condition

Our coviewing approach was based on the social nature of language development (Tomasello & Farrar, 1986) and the role that joint attention plays in early word learning. Tracing back to Bruner (1983), joint attention refers to moments when an adult and child are focused on the same thing and are mutually engaged in the discourse context. Examining the antecedents of labeling, Ninio and Bruner (1978), in their classic study of book reading, for example, showed how the mother and young child appeared to engage in a kind of informal scaffolding dialogue, with the mother initiating and responding to the child's vocal and gestural exchanges and directing the child's attention. Given the many longitudinal studies reporting positive correlations between joint attention and children's subsequent vocabulary (Morales et al., 2000), we attempted to adapt a coviewing approach that would use social cues to highlight what to learn and when to augment word learning. To do so, we developed several coviewing strategies to help children attend to target words. These included such techniques as pointing to an object on the screen, laughter, brief comments, or reactions to a character throughout the viewing. For example, Quack, one of the main characters says, "That's not a box, it's a CUBE! It has 6 sides that are squares and each square has four corners," followed by the coviewer saying, "Oh wow, it's a cube!"

To ensure that these coviewing comments were consistent, we created a script for each video. The script included not only what to say to the child but also when to say it. Five types of coviewing prompts were included: (a) repeating the target word; (b) pointing to the object when the characters on the screen said the word; (c) making real-life connections to the target word (e.g., for the word *cone*, the coviewer would state, "That looks like an ice cream cone!"); (d) providing brief recaps of certain plot points (e.g., responding to Chirp, who finds a circle, the coviewer says, "Chirp is right! That shell is a circle. It doesn't have any corners like other shapes"); and (e) reacting to the program's content (e.g., laughing when something funny happens). Table 2 provides excerpts of scripts and coviewing examples.

We videotaped an actress and a young child engaged in coviewing these programs. These videos were used to train two graduate research assistants in educational psychology in following the scripts and to ensure consistency of the implementation. Research assistants were trained to respond to comments or questions initiated by the child but were not to provide information about the target words beyond the scripted dialogue.

Measures

Prior to the start of the study, two trained graduate research assistants administered two pretest measures.

Receptive and expressive vocabulary. We administered a pretest measure to assess children's prior knowledge of target words. This measure was designed to provide further assurance that target words were not already familiar to children. Two formats were developed: receptive and expressive. Similar in format to the PPVT (Dunn & Dunn, 2007), for the receptive items, the child was asked to point to a picture of a word among four options. There were 24 items, 16 representing the target words from both stories (e.g., eight per video), along with eight foils (Cronbach's

Table 2
Excerpts of Scripts and Coviewing Examples

Seeds script	Coviewer
Narrator: When Peer first saw the yellow flower, he liked it right away because it looked like just like HIM.	
	"That flower does look just like him—like a yellow circle! No wonder he likes it so much."
Peep: Woohoo!	
Narrator: But when he went to see it, he discovered that it wasn't a little flower close to home, but a great big yellow flower head on top.	
	"Wow! That **stem** (point) does a good job of holding up all the leaves and the flower head."
Peep: It's beautiful!	
Narrator: The sunflower really was beautiful. It's a big flower head. . . . Peep loved it so much that he went to see it every day. . . . But one day when Peep got there. . .	
Peep: Ohhhhh!	"Ohhh!" The **sunflower**!
Narrator: Peep and Quack were very excited. they were looking for a TREASURE. The only problem is that they didn't know what a treasure was.	
	"That looks like fun! I hope Peep and Quack find **treasure** in the sand."
Peep: "Over here"	
Quack: "What is it?"	
Peep: "Look what we found."	"Ohhhh"
Narrator: Peep and Quack have found a pyramid. It's a special shape that has 4 triangle sides, a square bottom and a pointy top.	
	"Wow! Look at that **pyramid**. It has triangle sides and a square bottom (point)."

Note. Bolded words are the target words to be learned in the clip.

α = .52). Low reliability likely resulted from children randomly selecting answers because of a lack of vocabulary knowledge.

Following a format of the Expressive One-Word Picture Vocabulary Test (EOWPVT; Martin & Brownell, 2000), for the expressive items, the child was shown a picture of a word and asked, "What's this?" Once again, there were 24 items, including the target words and foils (Cronbach's α= .71). In total, the pretest measure included 48 items administered over a 10-min time period.

Children were eligible to participate in the study if they answered half or fewer of the target receptive and expressive vocabulary items correctly, assuring us that there would be sufficient room for growth. Eleven children identified more than half of the target words and were eliminated for the remainder of the study.

Peabody Picture Vocabulary Test (Dunn & Dunn, 2007). We administered the PPVT to examine children's overall receptive language skills. Reliability was 0.91. We used standardized scores as an indicator of baseline vocabulary.

While viewing. Children viewed the programs from a computer connected to an eye-tracking device.

We tracked children's eye movements using the Tobii Technology T120 eye-tracker integrated into a 17-in. thin-film-transistor 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 [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.

Preschoolers sat approximately 60 cm 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 was shown at five points on the screen prior to the beginning of the eye-track task. After calibration, children would then view the program, with the research assistant able to follow the child's eye movements and behaviors using the live view on the second monitor.

Postviewing assessments. Following the viewing of each educational media story, children were administered two assessments in word identification.

Receptive word identification. Similar in format to the PPVT, children were shown four images and asked to point to the target word. Two items per word were examined, one that used a specific screenshot from the video and another that used a nonscreenshot cartoon image. Distractor images were all thematically perceptually similar to the target word. For example, to assess the target word *cube*, children were shown a picture of the target word along with distractors of a pyramid, cone, and round shell. A total of correct responses was calculated for each assessment. There were 16 items per assessment, for a total of 32 items across the two videos (Cronbach's α= .61).

Expressive word identification. Similar in format to the EOW-PVT, children were shown a screenshot of each target word and asked, "What is this?" Correct responses of the exact word (e.g., no synonyms were accepted) were calculated for each video. There were eight items per assessment, for a total of 16 items across both videos (Cronbach's $\alpha = .69$).

Procedure

Trained graduate student assessors administered all assessments individually to children in a quiet location at the center. Research assistants were randomly assigned to subjects. Pretests were administered a week before the start of the study. Children were randomly assigned to a counterbalancing condition (video in each condition; coviewing or noninteractive) to watch a 9.5-min video on a laptop, either with a coviewer or on their own. Following the viewing, posttests for the relevant video were administered. After approximately an hour, children would watch a second video (with or without the coviewer), followed by assessments. Therefore, each child received both conditions (in counterbalanced order), serving as his or her control. Each session, including the time for posttests, totaled 20 to 25 min.

In the coviewing condition, the research assistant would sit next to the child, following the protocol described earlier to ensure consistency of implementation. The assistant was trained to provide short responses to any comments or questions that might arise while viewing. However, they provided no additional repetitions, clarifications, or information on target words beyond the scripted dialogue.

Two strategies were used to ensure fidelity to the coviewing condition. Throughout the experiment, two of the authors of the study conducted spot-checks to verify the accuracy of the implementation. In addition, all coviewing sessions were audiotaped, and a random selection of these recordings were also examined for accuracy of implementation to ensure consistency throughout the experiment (e.g., eliminating the possibility of drift) when the observers were not present. Through observational spot-checks and audio-reviewed cases, our analysis indicated that research assistants accurately implemented the protocol and followed the scripts with high fidelity.

In contrast, in the noninteractive condition, children viewed the video without any adult interaction. In this case, the child watched the video on his or her own. The assistant remained in the room to supervise the child but made her presence less available by sitting approximately 10 ft. away. The assistant did not make eye contact or interact with the child while the video was playing. After the viewing, the relevant assessments were then given.

Analysis

From our eye-tracking data, we investigated attention in two ways. The first was to assess the percentage of time a child looked anywhere on the screen during the entire video. We calculated this percentage by summing the total length of all fixations on the video divided by length of the video. This calculation served as an index of the focus by the child on the program in general.

The second method was designed to provide a more precise estimate of attention to the visuals associated with the target words. Here, our goal was to examine how coviewing might affect the amount of time the child spent looking at the visual representation of a target word when it was named by the character. This served as an index of selective attention. It recognizes that in order to learn, a child needs to associate a label with its referent; if a child looks at a different object than the one referred to on the screen (e.g., a pyramid instead of a cube) then a child is unlikely to develop an accurate link between the word and its visual representation.

In order to calculate the percent of fixation duration, we drew areas of interest (AOIs) around the visual depiction of the target vocabulary word for up to 3 s each time it was labeled. We then extracted the fixation duration for each AOI of each word. In the case of the word *cone*, for example, we drew the AOI around the visual image on the screen at the same time the character said the word, with a 2-mm margin around the border. Because some target words were repeated more often than others, we computed the percent fixation duration on all words individually. We calculated a proportion for each word by adding the fixation durations in the AOIs for each word and then dividing that number by the total length of all AOIs of that word. This calculation was converted into a percentage by multiplying the proportion by 100.

We then used repeated measures ANCOVA with viewing condition as the within subject factor, and the child's age in months as a covariate, to examine the effects of attention and word repetition. We used an additional covariate, time (pre- to posttest), for examining receptive and expressive word learning. We followed up this analysis with *t* tests to examine group differences between conditions and word learning. Although most of our measures were non-normal, ANCOVA models are generally robust to violations of the assumption of normality (Blanca, Alarcón, Arnau, Bono, & Bendayan, 2017). However, to ensure we did not overinterpret our results, we replicated each of our analyses that yielded significant results with nonparametric tests, which do not depend on the assumption of normality. Because nonparametric tests do not allow for covariates, the covariate was not included in any of these analyses. For omnibus tests of main effects, we conducted Friedman's two-way analysis of variance by ranks. Because nonparametric tests do not produce interaction effects, we did not replicate these findings; rather, we moved immediately to the pairwise comparisons for these analyses using the Wilcoxon signed-ranks test. We did not find any changes to significance using the nonparametric tests; thus, for the sake of including the covariate and interaction effects, we continue to report the original analyses based on the general linear model.

Results

Coviewing and Attention

Our first series of analyses addressed whether coviewing influenced attention and whether this effect might be impacted by the number of word repetitions. To examine these questions, we conducted a repeated measures ANCOVA, with children's fixation duration on the target words as a dependent variable, coviewing condition and repetition as within-subject independent variables, and age in months as a covariate. Our analyses reported a significant main effect of condition, $F(1, 69) = 49.33$, $p < .001$, $\eta_p^2 = .417$. There was also a significant main effect of repetition, $F(1, 69) = 11.64$, $p = .001$, $\eta_p^2 = .144$. However, there was no significant effect of the covariate, $F(1, 69) = .85$, $p = .359$, $\eta_p^2 = .012$, or significant interaction between condition and repetition, $F(1, 69) = .76$, $p = .388$ $\eta_p^2 = .011$. These results indicate that repetition and coviewing each influenced children's attention to the target words. As shown in Table 3, children attended to words that were repeated more often and spent more time attending to target words in the coviewing condition than when viewing on their own. In short, coviewing appeared to have a facilitative effect on attention to these target words.

Coviewing and Word Learning

Our next steps were to examine whether participating in the coviewing condition influenced children's receptive and expressive word learning. For this analysis, we conducted two 2 × 2 × 2 repeated measures ANCOVA, with receptive and expressive word learning scores as dependent variables. Time (pretest or posttest), coviewing condition, and repetition were entered as within-subjects independent variables, with age in months as a covariate in the analyses. Table 4 provides the means and standard deviations of words learned according to condition and repetition.

Receptive word learning. For receptive word learning, we found a significant main effect of time, $F(1, 80) = 581.84$, $p < .001$, $\eta_p^2 = .879$. Children were able to identify more words after viewing than at pretest, which suggests word learning. There was also a significant main effect of repetition, $F(1, 80) = 62.07$, $p < .001$, $\eta_p^2 = .437$, and a significant Repetition × Pre/Post interaction, $F(1, 80) = 8.15$, $p = .005$, $\eta_p^2 = .092$. Following up on the significant interaction, we found that children learned more high-repetitive than low-repetitive words, $t(82) = 2.50$, $p = .014$, $d = .34$. This suggests that words repeated more often were more easily learned by children. There was no significant effect of coviewing condition, $F(1, 80) = .38$, $p = .541$, $\eta_p^2 = .005$, the covariate, $F(1, 80) = .86$, $p = .357$, $\eta_p^2 = .011$, the Condition × Repetition interaction, $F(1, 80) = .65$, $p = .424$, $\eta_p^2 = .008$, or the Time × Condition × Repetition interaction, $F(1, 80) = .28$, $p = .600$, $\eta_p^2 = .003$. These results indicated that children were able to identify new words from videos and were more likely to learn words that were repeated more often. At the same time, however, coviewing did not appear to impact their receptive word learning.

Expressive word learning. For expressive word learning, we found a significant main effect of time (pre- to posttest), $F(1, 81) = 41.59$, $p < .001$, $\eta_p^2 = .339$, which, once again, suggests word learning. We also found a significant main effect of repetition, $F(1, 81) = 31.65$, $p < .001$, $\eta_p^2 = .281$, and a significant repetition by pre–post interaction, $F(1, 81) = 10.03$, $p = .002$, $\eta_p^2 = .110$. There was no significant effect of the covariate, $F(1, 81) = .42$, $p = .517$, partial $\eta_p^2 = .005$, condition, $F(1, 81) = 2.85$, $p = .095$, $\eta_p^2 = .034$, Condition × Time interaction, $F(1, 81) = 1.02$, $p = .316$, $\eta_p^2 = .012$, or Condition × Repetition interaction, $F(1, 81) = 2.67$, $p = .106$, $\eta_p^2 = .032$. These analyses show a similar pattern as with receptive word learning—children learned the target words and were more likely to learn them if they were repeated more often. However, here, we also found a significant three-way interaction between time (pre- to posttest), repetition, and coviewing condition, $F(1, 81) = 8.85$, $p = .004$, $\eta_p^2 = .098$. In order to further explore this interaction, we conducted pairwise t tests. In this case, we found that the coviewing condition made a significant difference for lower repetition words $t(82) = 2.87$, $p = .005$, $d = .45$, but not for higher repetition words $t(82) = .91$, $p = .365$, $d = .08$. That is, coviewing appeared to contribute to expressive word learning when words were not repeated often. However, coviewing made no difference when words were often repeated in the video itself. Taken together, these results suggest that coviewing may support children's expressive word learning with fewer repetitions.

Table 3
Percent of Time Spent Fixating on Video Vocabulary Labels by Word Repetition and Coviewing Condition

Percent of time	Coview	Noninteractive
Percent fixation low repetition*	19.03 (12.71)	11.21 (9.19)
Percent fixation high repetition*	20.56 (8.98)	15.16 (8.23)

* $p < .05$.

Coviewing, Attention, and Word Learning

In our final analysis, we attempted to consolidate what we had learned in the two previous analyses to better understand how coviewing might affect children's attention and their subsequent learning of words repeated more or less frequently in these videos. Although in the previous analyses, we found no direct effect of coviewing on word learning, a direct effect is not always required to demonstrate a meaningful indirect effect (Hayes, 2009). Therefore, we examined whether coviewing may guide children's attention, having an indirect effect on word learning.

To do so, we conducted a mediation analysis to determine if the relationship between coviewing and learning might be mediated by attention. Because our conditions were manipulated within subject, a mixed analysis was used to examine potential mediation effects including a random intercept (Vuorre & Bolger, 2018). To test a mediation model, we first entered condition, the pretest variable, and age in months into each model. In Step 2, we added attention. To demonstrate a mediation, we would expect that the effect of coviewing would decrease between the two models.

Receptive word learning. For receptive word learning, as shown in Table 5, we did not find evidence for a mediation effect. For both high- and low-repetition words, as expected, there was no direct effect of coviewing in Model 1. In Model 2, for low-repetition words, there was a significant effect of attention, over and above the effect of condition. Although there was a slight reduction in the effects of coviewing, this difference was nonsignificant ($z = .74$, $p = .230$). For high-repetition words, we also found no evidence for a mediation effect. The results of the first model failed to show an overall significant effect of fixation duration on receptive word learning. In Model 2, the effect of attention was nonsignificant, and there was no decrease in the effect of condition. These results were consistent with our findings in the previous analyses and continue to suggest that although coviewing impacted attention, and attention impacted receptive word learning, these effects operated distinctly from one another.

Expressive word learning. For expressive word learning, however, we found a different pattern of results (Table 6). In the case of low-repetition words for Model 1, we found a significant effect of condition, suggesting that there may have been a direct effect of coviewing on outcomes when broken down by repetition.

Table 4
Means (and Standard Deviations) of Receptive and Expressive Words Identified by Word Repetition and Coviewing Condition

	Low repetition				High repetition			
	Coview		Noninteractive		Coview		Noninteractive	
Target words	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Receptive words	.86 (.86)	3.28 (1.59)	.78 (.82)	3.06 (1.66)	1.29 (.84)	4.08 (1.72)	1.26 (.80)	4.12 (1.69)
Expressive words	.18 (.56)	.49* (.80)	.13 (.38)	.20* (.44)	.36 (.62)	.71 (.86)	.32 (.54)	.79 (.88)

Note. Asterisk indicates significant difference between coview conditions. Pre = pretest; Post = posttest.
* $p < .05$.

When attention was entered into Model 2, the effect of fixation was significant, but the effect of coviewing had been reduced sufficiently to be nonsignificant. This may provide some evidence for a mediation effect in expressive word learning of low-repetition words. For high-repetition words, there was no effect of coviewing in the first model. In the second model, there was an effect of attention, but the effect of condition was not reduced. In fact, coviewing showed a stronger effect when attention was entered into the model, which was counter to our original hypotheses.

Taken together, this analysis suggests that for receptive language, although coviewing may impact attention and attention may impact word learning, these two processes appeared to be acting separately. For expressive language, although the pattern is similar for high-repetition words, attention may mediate the relationship between coviewing and word learning for low-repetition words.

Discussion

Language learning for young children occurs in social contexts (Bruner, 1983). Consequently, coviewing is thought to support a more optimal context for young children's language learning from educational media than when viewing on their own. In the coviewing context, adults may engage in brief interactions, model behaviors, and provide informal social cues for making meaning. Studies of coviewing (Reiser, Williamson, & Suzuki, 1988; Salomon, 1977), primarily of educational TV viewing, showed promise that an adult presence could enhance children's learning from the screen. However, in a more active mediational role, studies have shown that specific pedagogical techniques by parents, such as pausing an educational video at various time points to ask questions, encouraging children to retell parts of the story, could improve children's expressive vocabulary (Strouse et al., 2013) and knowledge of program content (Valkenburg, Krcmar, & de Roos, 1998).

Nevertheless, in today's media environment, children are likely to view educational media programs largely on mobile devices (e.g., smartphones, tablets, computers), not on DVDs or large-screen videos (Rideout, 2017), for which no such pausing may be possible. In these more typical settings, parents and children may view and talk synchronously. Consequently, our coviewing approach was designed to engage adults and children in joint activity, positioning the adult as a cocreator of meaning, similar to other word-learning situations. Reflecting the social-pragmatic dimension of language acquisition (Tomasello, 2000), our model assumes that children learn words not merely by having an adult label an object but by developing, through adult–child social interaction, a mutual understanding in a joint context.

This coviewing approach is designed to model a social context that is more typical of the intersubjective communication between parents and children in their day-to-day interactions. In contrast to instructive mediation (Valkenburg, Krcmar, Peeters, & Marseille, 1999), or the use of pedagogical techniques throughout the viewing process (e.g., asking questions; recalling events), our model attempted to blend the social coviewing process with brief attention-directing cues (e.g., laughter, pointing, repeating) that could indicate for the child the adult's intended referent. Our results indicated that this coviewing approach had its intended effect: Children spent longer times looking at the target word in the coviewing condition than when viewing on their own. Acting as a brief scaffolding device, coviewing seemed to call children's attention to words. Furthermore, children attended more to words in the coviewing condition that were repeated more frequently (e.g., eight to 10 times) than when repeated only three to four times. These results are further supported in recent research (Samudra, Flynn, & Wong, 2019), in which coviewing was found to enhance children's visual attention to target words.

Table 5
Mixed Models for Receptive Word Learning Examining the Relationship Between Word Learning, Attention, and Condition

Model	Estimate	*SE*	*t* value	*p* value
Low-repetition words				
Model 1				
Age in months	.02	.32	.05	.961
Pretest*	.62	.14	4.32	<.001
Condition	.18	.24	.755	.451
Model 2				
Age in months	.01	.35	.04	.967
Pretest*	.41	.16	2.60	.010
Condition	−.08	.27	−.32	.752
Percent fixated on vocabulary*	.02	.01	2.42	.017
High-repetition words				
Model 1				
Age in months	.41	.40	1.03	.307
Pretest	.11	.15	.745	.457
Condition	.03	.22	.153	.878
Model 2				
Age in months	.35	.45	.79	.432
Pretest	.08	.18	.48	.632
Condition	.14	.26	.55	.587
Percent fixated on vocabulary	.03	.02	1.69	.094

Note. Asterisk indicates significant predictor. *SE* = standard error.
* $p < .05$.

Table 6
Mixed Models for Expressive Word Learning Examining the Relationship Between Word Learning, Attention, and Condition

Model	Estimate	*SE*	*t* value	*p* value
Low-repetition words				
Model 1				
Age in months	−.05	.11	−.46	.647
Pretest*	.85	.08	10.53	<.001
Condition*	.25	.08	3.28	.002
Model 2				
Age in months	−.11	.12	−.89	.379
Pretest*	.82	.08	9.62	<.001
Condition	.14	.08	1.67	.098
Percent fixated on vocabulary*	.13	.004	3.31	.001
High-repetition words				
Model 1				
Age in months	.12	.19	.62	.537
Pretest*	.65	.10	6.24	<.001
Condition	.11	.09	1.21	.228
Model 2				
Age in months	.16	.21	.76	.452
Pretest*	.66	.11	5.90	<.001
Condition*	.22	.10	2.16	.034
Percent fixated on vocabulary*	.02	.01	2.26	.025

Note. Asterisk indicates significant predictor. *SE* = standard error.
* $p < .05$.

At the same time, there was considerable variability within this low-income sample. For example, our measure of attention, fixation duration, showed a fairly sizable range of seconds devoted to target words among the low-income children in our sample. Similarly, although word repetition seemed to have a facilitative effect, standard deviations on receptive and expressive word learning among this group were substantial. Rather than age, prior scores on receptive and expressive language seemed to best predict children's gains. These results highlight the heterogeneity of this low-income sample.

This variability often goes unnoticed in cross-SES comparisons and could have important instructional implications. For example, word repetition has been used as a primary catalyst for vocabulary learning in many studies of vocabulary, and often serves as an indicator of readability in text (Biemiller, 2006; Stahl & Nagy, 2006). However, some children from low-income circumstances might benefit from a substantially greater number of repetitions than others. In one study, for example, Pinkham (2011) reported that 28 repetitions were needed before a threshold of 80% proficiency (e.g., ability to label the word) was reached. These results suggest the importance of repetition in learning words from educational media (Linebarger et al., 2013) and may suggest differentiated exposure for those to take advantage of it. Through repeated exposures, children began to learn some of the statistical regularities of how the word may be used in multiple contexts. In our case, the repetitions were similar to recasts, often seen as a predictor of young children's syntactic growth. These results are consistent with Rice and Woodsmall's (1988) research, which found that word learning was associated with the repetition of words in similar, but not identical, contexts.

Yet there were differential patterns in gains for receptive and expressive vocabulary. Because word identification precedes the production of semantic context, it was not surprising that children identified more words receptively than expressively. Children identified about three words when repeated less frequently, and four words when repeated more frequently. For receptive language, coviewing did not contribute to greater word gains. This suggests that these educational media programs on their own may potentially contribute to children's vocabulary development. These results add to the accumulating evidence that preschoolers can learn words through rapid online processing of educational media without adult support (Takacs et al., 2014).

At the same time, gains were not as impressive for expressive language. Children made only modest improvements compared with receptive vocabulary. But here, the contribution of the coviewer seemed to add a helping hand, supporting children to use low-repetition words. Our mediational analysis further showed that attention may be one determining factor in the benefit of a coviewer, as attention partially mediated the relationship between coviewing and expressive word learning for low-repetition words. Whether other factors, like additional repetitions provided by the coviewer or the more informal cues or responses from them, also contributed to the effect of coviewing cannot be determined at this point. However, it does suggest that the coviewer scaffolded learning in the absence of sufficient input from the video itself.

The mediational role of attention and coviewing also differed for receptive and expressive language, once again emphasizing the importance of assessing word learning both ways. For receptive language, coviewing did not directly or indirectly affect word learning. But this was not the case for expressive language. Here, attention mediated the relationship between coviewing and low-repetition word learning, although it may have had a suppressive effect on the relationship between coviewing and high-repetition word learning. This finding was contrary to our hypotheses and, to our knowledge, has not been reported in previous studies. More research is needed to determine the theoretical or practical implications of such a finding.

Although these findings are difficult to disentangle, Verhallen and Bus (2010) speculated that unknown words are rarely learned expressively before receptively. Although children might have been able to identify words regardless of whether they or the coviewer had spoken them, expressive word learning may require children to have at least a partial knowledge of words and have spoken them while viewing, supporting the role of retrieval practice in acquiring expressive language.

Our findings for expressive language stand in contrast to research by Strouse and colleagues (2013), who reported improvements in expressive language resulting from their coviewing approach. Such differences in findings could be due to the differences in our approaches to coviewing. For example, in Strouse et al.'s study, the parent engaged in an active mediational role, stopping the program to ask questions and encouraging the child's retelling of the story. In contrast, our approach focused on the attentional dynamics in which adult–child dyads might engage in day-to-day mutual activity. Therefore, it could be that our approach did not sufficiently engage in talking about the words and their meaning in the program. Sénéchal (1997), for example, found that having a child answer questions during multiple readings of a storybook was more helpful to the acquisition of expressive than receptive vocabulary.

But it could also reflect an important limitation in our study. For example, in several of his studies, Tomasello (1999) found that the child had to first understand the communicative intentions of the adult in a novel communicative situation before the child could infer what a smile, frown, or gesture might mean. In other words, a smile or a frown was not sufficient by itself to indicate to the child the adult's intended referent. However, once a mutually understood joint attentional scene occurred, these behaviors could be better understood. Therefore, it could be that children in our study might not have understood the communicative intentions of these unfamiliar adults while viewing these educational media. Further research might wish to explore the effects of our coviewing approach among parent–child dyads to determine whether this might be the case.

Our conclusions must be qualified in several additional ways. First, we limited our analysis of receptive and expressive word learning to nouns rather than other parts of speech. Based on previous research (Harris, Golinkoff, & Hirsh-Pasek, 2011), however, we know that concrete nouns have an advantage in children's acquisition of words in digital media; therefore, these results might potentially inflate the number of words children receptively identified. It remains to be seen whether or not our findings are confirmed with other word types. Second, we also recognize that our measures examined immediate recognition of words. In future studies, we plan to examine whether words are later recalled or incorporated into children's language repertoires. Third, our analysis of the effects of coviewing was confined to word learning. It is entirely possible that coviewing has many other benefits, including the sheer enjoyment of engaging in joint activity with others. And lastly, and perhaps most importantly, our study was conducted in children's early care and education settings using a laptop, with trained research assistants as coviewers. Therefore, we cannot assume that children's attention or learning in a more naturalistic home setting on a smartphone or a tablet with their parents would yield similar results.

Given these limitations, however, this study provides further evidence that low-income children can pick up at least partial word knowledge as a result of viewing educational media. Furthermore, they can do so on their own, particularly when words are repeated frequently. Media producers should therefore consider word repetition in designing educational programming. In less ideal situations, when word repetition is low, coviewing seems to provide a temporary language scaffold—essentially, a brief bootstrap for expressive word learning. Taken together, this research begins to highlight the features of educational media and the social supports that might contribute to children's language learning.

References

American Academy of Pediatrics. (2016). Media and young minds: Council on Communications and Media. *Pediatrics, 138,* 1–6. http://dx.doi.org/10.1542/peds.2016-2591

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

Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In D. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 41–51). New York, NY: Guilford Press.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98,* 44–62. http://dx.doi.org/10.1037/0022-0663.98.1.44

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29,* 552–557.

Blewitt, P., & Langan, R. (2016). Learning words during shared book reading: The role of extratextual talk designed to increase child engagement. *Journal of Experimental Child Psychology, 150,* 404–410. http://dx.doi.org/10.1016/j.jecp.2016.06.009

Bruner, J. (1983). *Child's talk: Learning to use language*. New York, NY: W. W. Norton.

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

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

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

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

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.

Harris, J., Golinkoff, R., & Hirsh-Pasek, K. (2011). Lessons from the crib for the classroom: How children really learn vocabulary. In S. B. Neuman & D. Dickinson (Eds.), *Handbook of early literacy research* (Vol. 3, pp. 49–65). New York, NY: Guilford Press.

Hart, B., & Risley, T. (1995). *Meaningful differences*. Baltimore, MD: Brookes.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76,* 408–420. http://dx.doi.org/10.1080/03637750903310360

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 in the Public Interest, 16,* 3–34.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437–447. http://dx.doi.org/10.1037/0022-0663.80.4.437

Korat, O., Levin, I., Atishkin, S., & Turgeman, M. (2014). E-book as facilitator of vocabulary acquisition: Support of adults, dynamic dictionary and static dictionary. *Reading and Writing, 27,* 613–629. http://dx.doi.org/10.1007/s11145-013-9474-z

Linebarger, D., Moses, A., Garrity Liebeskind, K., & McMenamin, K. (2013). Learning vocabulary from television: Does onscreen print have a role? *Journal of Educational Psychology, 105,* 609–621. http://dx.doi.org/10.1037/a0032582

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. http://dx.doi.org/10.1111/j.1467-8624.2010.01493.x

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

Martin, N., & Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test* (4th ed.). Austin, TX: Pro-Ed.

McKeown, M., Beck, I., Omanson, R., & Pople, M. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20,* 522–535. http://dx.doi.org/10.2307/747940

McWayne, C. M., Melzi, G., Limlingan, M. C., & Schick, A. (2016). Ecocultural patterns of family engagement among low-income Latino families of preschool children. *Developmental Psychology, 52,* 1088–1102. http://dx.doi.org/10.1037/a0040343

Morales, M., Mundy, P., Delgado, C., Yale, M., Messinger, D., Neal, R., & Schwartz, H. (2000). Responding to joint attention across the 6- to 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology, 21,* 283–298. http://dx.doi.org/10.1016/S0193-3973(99)00040-4

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development, 86,* 1351–1370. http://dx.doi.org/10.1111/cdev.12398

Neuman, S. B., & Celano, D. (2006). The knowledge gap: Implications of leveling the playing field for low-income and middle-income children. *Reading Research Quarterly, 41,* 176–201. http://dx.doi.org/10.1598/rrq.41.2.2

Neuman, S. B., Wong, K., Flynn, R., & Kaefer, T. (2019). Learning vocabulary from educational media: The role of pedagogical supports for low-income preschoolers. *Journal of Educational Psychology, 111,* 32–44. http://dx.doi.org/10.1037/edu0000278

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language, 5,* 1–15. http://dx.doi.org/10.1017/S0305000900001896

Pinkham, A. (2011, December). *You can say that again: Preschoolers need repeated exposures to gain expressive vocabulary*. Paper presented at the Literacy Research Association, Jacksonville, FL.

Reiser, R., Tessmer, M., & Phelps, P. (1984). Adult-child interaction in children's learning from Sesame Street. *Educational Communication & Technology Journal, 32,* 217–223.

Reiser, R., Williamson, N., & Suzuki, K. (1988). Using "Sesame Street" to facilitate children's recognition of letters and numbers. *Educational Communication and Technology, 36,* 15–21.

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.

Rowe, M. (2018). Understanding socioeconomic differences in parents' speech to children. *Child Development Perspectives, 12,* 122–127. http://dx.doi.org/10.1111/cdep.12271

Rowe, M., Pan, B., & Ayoub, C. (2005). Predictors of variation in maternal talk to children: A longitudinal study of low-income families. *Parenting: Science and Practice, 5,* 259–283. http://dx.doi.org/10.1207/s15327922par0503_3

Salomon, G. (1977). Effects of encouraging Israeli mothers to c-observe Sesame Street with their five-year-olds. *Child Development, 48,* 1146–1151. http://dx.doi.org/10.2307/1128378

Samudra, P., Flynn, R., & Wong, K. (2019). Co-viewing educational media: Does co-viewing help low-income preschoolers learn auditory and auditory-visual vocabulary associations? *AERA Open*. Advance online publication. http://dx.doi.org/10.1177/2332858419853238

Samudra, P., Wong, K., & Neuman, S. B. (in press). Promoting low-income preschoolers' vocabulary from educational media: Does repetition and pacing support memory for learned word knowledge? *Journal of Cognitive Education and Psychology*.

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language, 24,* 123–138. http://dx.doi.org/10.1017/S0305000996003005

Sénéchal, M., & Cornell, E. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly, 28,* 360–374. http://dx.doi.org/10.2307/747933

Stahl, S. (2003). How words are learned incrementally over multiple exposures. *American Educator, 27,* 18–19.

Stahl, S., & Nagy, W. (2006). *Teaching word meanings*. Mahwah, NJ: Erlbaum.

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology, 49,* 2368–2382. http://dx.doi.org/10.1037/a0032463

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. http://dx.doi.org/10.3389/fpsyg.2014.01366

Takeuchi, L., & Stevens, R. (2011). *The new coviewing: Designing for learning through joint media engagement*. Retrieved from http://www.joanganzcooneycenter.org/wp-content/uploads/2011/12/jge_coviewing_desktop.pdf

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics, 10,* 401–413. http://dx.doi.org/10.1075/prag.10.4.01tom

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development, 57,* 1454–1463. http://dx.doi.org/10.2307/1130423

Valkenburg, P., Krcmar, M., & de Roos, S. (1998). The impact of a cultural children's program and adult mediation on children's knowledge of and attitudes toward opera. *Journal of Broadcasting & Electronic Media, 42,* 315–326. http://dx.doi.org/10.1080/08838159809364452

Valkenburg, P., Krcmar, M., Peeters, A., & Marseille, N. (1999). Developing a scale to assess three styles of television mediation: "Instructive mediation," "restrictive mediation," and "social coviewing." *Journal of Broadcasting & Electronic Media, 43,* 52–66. http://dx.doi.org/10.1080/08838159909364474

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

Vuorre, M., & Bolger, N. (2018). Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience. *Behavior Research Methods, 50,* 2125–2143. http://dx.doi.org/10.3758/s13428-017-0980-9

WGBH. (2004). *Peep and the big wide world*. Boston, MA: Author. Retrieved from https://blog.apastyle.org/apastyle/movies-and-tv-shows/

Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30,* 679–689. http://dx.doi.org/10.1037/0012-1649.30.5.679

Received November 3, 2018
Revision received September 18, 2019
Accepted September 23, 2019 ■