

Coviewing Educational Media: Does Coviewing Help Low-Income Preschoolers Learn Auditory and Audiovisual Vocabulary Associations?

Preeti G. Samudra

New York University

Rachel M. Flynn

Northwestern University

Kevin M. Wong

New York University

Coviewing media is a practice commonly recommended to parents of young children. However, little is known about how coviewing might scaffold the vocabulary learning of low-income preschoolers. The present study focused on how coviewing educational media influences children's learning of two different vocabulary associations—auditory-only and audiovisual vocabulary associations. We additionally studied whether children with weaker baseline vocabularies might particularly benefit from coviewing. One hundred twenty-eight low-income preschoolers viewed five educational media clips either with an adult coviewer or alone. Audiovisual and auditory vocabulary associations were then assessed. Results show that coviewing did not support vocabulary learning overall but did specifically support the development of auditory-only vocabulary associations for children with weaker baseline vocabularies. This suggests that coviewing may not provide a ubiquitous benefit but rather predicts learning in the mode of coviewer input (auditory) specifically for the children who need additional supports the most.

Keywords: *early childhood, child development, instructional technologies, learning environments, media, vocabulary*

A common perception of children's media use is that the “digital babysitter” is detrimental to child well-being. Nonetheless, children are heavily exposed to screen media such as television, DVDs, computers, and mobile phones from an early age. In fact, 2- to 4-year-old children view screen media for over 2.5 hours per day, with children in lower income households spending more time with screen media than those from higher income homes (Rideout, 2017). In light of the ubiquity of screen media, the American Academy of Pediatrics (AAP) recently updated its position from one that discourages media use in preschoolers to one that recommends that adults coview high-quality media with children (AAP, 2016).

High-quality programming for preschoolers often comes in the form of educational media—programs that are specifically intended to teach children valuable, school-focused information and skills such as vocabulary (Vandewater & Bickham, 2004). Vocabulary development during the preschool years is a critical predictor of later academic success,

making literacy-focused educational media a potentially strong genre to serve as a beneficial rather than detrimental “digital babysitter” (Storch & Whitehurst, 2002; Tamis-LeMonda, Kuchirko, Luo, Escobar, & Bornstein, 2017). A vast body of research aligns with this interpretation—preschoolers successfully learn vocabulary from educational media, even when viewing media alone (e.g., Mares & Pan, 2013; Rice, Conti-Ramsden, & Snow, 1990). Relatively little is known, however, about how and for whom the practice of coviewing might benefit vocabulary learning from literacy-focused educational media.

In the present study, we investigate the impact of coviewing on low-income preschoolers' vocabulary learning from educational media segments that provide explicit instruction of vocabulary words. We additionally investigate the learning of two forms of vocabulary associations—audiovisual and auditory only—to clarify how the added coviewer auditory input impacts the learning of vocabulary associations in different modalities. Finally, we investigate the role of



baseline vocabulary size in predicting how coviewing impacts vocabulary learning. In particular, we study whether coviewing might be particularly beneficial for children who may be more likely to need additional support—those with weaker initial vocabularies.

The Promise of Coviewing

It is clear why the AAP would view coviewing as a positive media practice for preschoolers. Prior research shows that interactive activities such as shared book reading and even child talk and conversational interactions more generally are beneficial for vocabulary and language growth (e.g., Gilkerson et al., 2017; Marulis & Neuman, 2010). As such, making the media consumption experience an interactive one could improve learning. Even in less interactive coviewing enactments, coviewing could promote child interest or attention to media (Saloman, 1977), possibly creating a more supportive environment for growth. Additionally, the educational component of media consumption could be enhanced through repetition and elaboration. The mere practice of increasing the amount of repetition and elaboration of content might make it more likely for children to encode and retain that information (Watkins, Calvert, Huston-Stein, & Wright, 1980). In this manner, coviewing may operate as a scaffold to the educational media program, which could create a more complete and robust learning experience for children.

Educational Media Affordances

In practice, however, educational media formats have many affordances that may render the additional scaffold of coviewing redundant for learning. A central theory that delineates the strengths of educational media as a learning format is dual coding theory (Clark & Paivio, 1991; Paivio, 1990). Dual coding theory proposes that two differing sensory modes of presentation of information (e.g., visual, auditory) promote learning of that information better than just one mode of presentation. This is because the two modalities are theorized to tap into different cognitive resources and therefore not compete for the same limited processing resources. Thus, combining multiple modalities to teach the same information is beneficial to learning. In terms of the efficacy of educational media in teaching vocabulary to young children, dual coding theory would suggest that receiving vocabulary-relevant content through both the auditory and visual channels results in better processing than one channel alone, allowing for educational media to support vocabulary growth (Clark & Paivio, 1991; Paivio, 1990).

More traditional formats, such as books, only offer static pictures. An adult is required to provide auditory input. Educational media combine both sensory modes seamlessly and uniformly in a manner that is optimized for its audience. Following from this, Takacs, Swart, and Bus (2015) conducted

a meta-analysis to determine if multimedia features, such as sound and animations, may promote the vocabulary growth of children from less stimulating family environments. They found that multimedia features produced an added benefit over static elements alone (e.g., traditional storybooks), particularly for children from disadvantaged backgrounds. Other research has similarly shown that educational media book formats, such as e-books and video storybooks, result in higher vocabulary gains than static, traditional book formats (Shamir, Korat, & Fellah, 2012; Verhallen & Bus, 2010; Verhallen, Bus, & de Jong, 2006). As such, the combination of visual and auditory input in educational media environments has the capacity to support learning even without adult interaction—sometimes even more so than interactive activities such as shared book reading. It is therefore unclear as to whether the addition of a coviewer makes the educational media environment significantly more supportive of learning.

Coviewing in Educational Media Environments

The limited coviewing literature has found fairly mixed results as to whether it benefits learning from educational media. Some studies find that specific enactments of coviewing promote word learning from media. For example, Reiser, Tessmer, and Phelps (1984) found that children performed better on letter and number naming when adult coviewers asked the child to name the letters and numbers and gave contingent feedback during the educational program compared to when viewing with a silent adult. In this coviewing enactment, trained coviewers asked children questions about the vocabulary taught in the educational media program and ensured they knew the answers prior to the final assessment of learning.

In a more recent investigation of video storybook comprehension, Strouse, O'Doherty, and Troseth (2013) found that parents who were trained to pause the video and use dialogic questioning techniques to converse with their 3-year-old child about the story were able to better enhance story comprehension compared to a control. In this enactment, parents were trained to ask their child open-ended questions and build their child's understanding of both vocabulary and comprehension of the media. Unfortunately, researchers found that these dialogic questioning techniques were rarely used spontaneously by parents. Overall, this suggests that coviewing interactions involving instructive and dialogic questioning are likely to enhance vocabulary learning from educational media—but these interactions are unlikely to occur in naturalistic coviewing enactments.

Other research, however, has found coviewing to exert little influence on learning. In addition to the coviewing enactment incorporating dialogic questioning, Strouse et al. (2013) studied a less intensive enactment in which parents talked to their child and directed their child's attention to the program but did not ask questions. They found that this coviewing

enactment did not enhance learning from educational media. Similarly, coviewing studies that focused on promoting socio-emotional development (Rasmussen et al., 2016) and story comprehension (Skouteris & Kelly, 2006) yielded no learning benefits of coviewing over viewing alone.

A critical difference between the studies by Rasmussen et al. (2016) and Skouteris and Kelly (2006) compared to those that found positive results for coviewing is that the former studies investigated coviewing as it naturally occurred. For example, Rasmussen et al. manipulated the coviewing scenario by asking parents in that group to talk to their child as much as possible about the show but did not prescribe how that should be done. Research suggests that for social interactions to be beneficial to learning, they need to be of high quality (Hirsh-Pasek et al., 2015), but due to the more naturalistic coviewing scenarios, these researchers were unable to describe the specific nature of the coviewing enactments. We therefore have a limited understanding of the quality of these coviewing interactions, limiting our understanding of the features of these coviewing enactments that failed to increase learning. In sum, the mixed coviewing literature suggests that we need a greater understanding of how more naturalistic enactments of coviewing impact vocabulary learning.

Our Coviewing Enactment

In the present study, we therefore utilized a clearly defined coviewing enactment that incorporated common interaction elements and behaviors that have been observed in parent-child shared book reading interactions. Some of these behaviors included pointing, repeated word labeling, and attention-directing statements (Ninio & Bruner, 1978). Much like Ninio and Bruner (1978), Evans et al. (2011) more recently found that when reading a book with their kindergarten-aged child, parental vocabulary-focused strategies primarily involved reiterating the vocabulary word and providing some conceptual or definitional information about the word.

In line with prior research on naturalistic forms of shared book reading, coviewers in the present study primarily focused on providing additional repetition of the word labels, examples, and definitions of words being taught in the educational media clips. They additionally promoted attention to the appropriate portions of the screen when providing word labels or naming pictured examples. Prior research suggests that cues within media that direct attention to relevant parts of the screen and provide repetition of vocabulary are particularly beneficial to vocabulary learning from educational media (Neuman, Wong, Flynn, & Kaefer, 2019). Coviewers were able to enhance the learning environment by including repetition and visual attention-directing elements through pointing and vocalizations indicating interest. Thus, by characterizing our coviewing enactment as

focusing on label repetition, example repetition, word defining, and pointing, we investigate how a vocabulary-focused, robust use of these naturalistic coviewing strategies impact vocabulary learning from educational media.

In addition to studying a clearly defined and consistently enacted form of coviewing that incorporates naturalistic parent-child interaction elements, we extend the coviewing literature by (a) studying children's learning of different forms of vocabulary associations—audiovisual and auditory only—and (b) investigating the role of children's baseline vocabulary in predicting how coviewing impacts learning.

Auditory and Audiovisual Vocabulary Associations

Prior research has rarely investigated vocabulary learning in combined versus individual sensory modalities. When studying coviewing, this might be particularly important because coviewer input falls primarily in the auditory medium. By providing extensive repetition of vocabulary terms and examples shown in the educational media clip, coviewers enhance the auditory environment while the media remains the primary source of visual information. Working memory models (e.g., Baddeley, 1986, 1992) discuss how two processes occur when input is received—rehearsing phonological material in one's mind to maintain it in working memory and translating visual input into a phonological form. However, preschool-aged children are thought to struggle with these processes. During the preschool years, children do not strategically mentally rehearse auditory information and are dependent on visual strategies for memory (Palmer, 2000).

As such, children may be more likely to remember visual information provided by media than the auditory information. Since preschool-aged children struggle with mental rehearsal of auditory information, the added vocabulary labels and auditory labeling of examples might allow these terms to be active in memory for longer than without the active coviewer. This might in turn help children retain the auditory information more effectively. The heightened auditory focus of the coviewer may also differentially impact learning of different associations. For audiovisual question formats, children have the visual representations to direct their associative memory. This visual cue is absent in auditory-only question formats. Therefore, coviewing has the potential to benefit both types of vocabulary associations but may be more influential for associations exclusively within the auditory medium.

The Role of Baseline Vocabulary

In addition to coviewing potentially differentially impacting vocabulary associations across different modalities, coviewing might also differentially benefit children's learning based on their extant vocabularies. Prior research suggests that having a smaller initial vocabulary size makes it more

challenging to learn new words (e.g., Blewitt, Rump, Shealy, & Cook, 2009; Neuman et al., 2019; Senechal, Thomas, & Monker, 1995). Having a stronger vocabulary allows children to connect the new words they hear to a wider variety of pre-existing representations, making it more likely that they encode and can later access information regarding the new word. Consequently, children with stronger baseline vocabularies may learn more words from educational media overall than those with weaker initial vocabularies.

A potential strength of coviewing might be to lessen this learning divide based on prior vocabulary. The children in our sample come from low-income backgrounds, making such a benefit particularly critical. The early childhood years represent a particularly salient time period for vocabulary development and the consequences of vocabulary growth for later academic trajectories (Wagmiller, Lennon, Kuang, Alberti, & Lawrence Aber, 2006). Additionally, stable and enduring differences in vocabulary growth based on social class differences are formed specifically during the early childhood years (Farkas & Beron, 2004). It is thus critical to investigate the potential for coviewing to ameliorate these differences in learning in our population of low-income preschoolers.

Coviewing might prove beneficial to learning because our enactment of coviewing serves as an added auditory scaffold, supporting the content of the media itself. Since the educational media content we use (*Sesame Street: Word on the Street*) was specifically designed to teach preschoolers new words, it is possible that children with relatively strong baseline vocabularies do not need the additional scaffolds provided by the coviewer to learn from the media clips. In contrast, children with relatively weak baseline vocabularies may benefit from the coviewing scaffold.

Vygotsky (1978) discussed how children benefit the most from adult scaffolding when they are learning something that falls in their zone of proximal development—consisting of the types of tasks that they cannot yet learn independently but can learn with adult assistance. Since educational media is designed to be comprehensible for preschool-aged children watching independently, it is likely that children with larger initial vocabularies are better able to learn from media without additional support. The children with weaker initial language skills, however, may struggle with word learning from educational media independently, so the task of vocabulary learning for these children may fall into their zones of proximal development in which adult scaffolds are most effective. As such, by providing additional scaffolds, a coviewer might particularly benefit the learning of children with weaker initial vocabularies who are in greater need of those scaffolds.

The Present Study

The present study therefore aims to better understand how having an adult coviewer might support the existing

strengths of educational media to further promote the vocabulary development of preschoolers from low-income backgrounds. Children viewed five educational media clips either with an adult coviewer or independently. Adult coviewers used a coviewing enactment that was characterized by strategies (e.g., word labeling, pointing, repeating examples) commonly used in naturalistic parent-child interactions. Word learning was assessed through vocabulary assessment formats that played to the strengths of educational media (combined visual-auditory formats) as well as a format that may specifically be benefited by the additional auditory input from a coviewer (an auditory-only format). In this way, we investigated whether coviewing would support different types of vocabulary associations and knowledge.

We were additionally interested in seeing if coviewing would be particularly beneficial for children in need of additional supports—those with lower baseline language skills. Prior research suggests that having a smaller initial vocabulary size makes it more challenging to learn new words (e.g., Blewitt, et al., 2009; Neuman et al., 2019), so we sought to determine if a coviewer might help close this learning divide. Overall, we examine learning from educational media and the impact of a naturalistic form of coviewing, focusing on three primary research questions:

Research Question 1: Do low-income preschoolers understand vocabulary words they are exposed to in educational media clips overall?

Research Question 2: Does a vocabulary-focused coviewing intervention that incorporates naturalistic parent-child interaction strategies enhance vocabulary knowledge from educational media?

Research Question 3: Does coviewing differentially support children based on their extant vocabularies?

Method

Participants

The present sample comprised 128 3- and 4-year old children ($M_{\text{age}} = 4$ years, 5 months, $SD_{\text{age}} = 4.71$ months); 45% were female. The sample was diverse: 44% were African American, 45% were Hispanic, 9% were Caucasian, and 2% were biracial or other. Participating children were enrolled in three Head Start centers located in high-poverty areas in a large urban city. Institutional Review Board approval was obtained from the host university, and preschool educational directors, teachers, and parents provided consent for participation. All children qualified for free and reduced lunch. Standardized receptive language scores, as measured by the Peabody Picture Vocabulary Test (PPVT), averaged 87.19 ($SD = 18.36$), which is approximately one standard deviation below the population mean. Demographic information by coviewing group is shown in Table 1.

TABLE 1
Demographic Information by Coviewing Group

	Coviewing Group	Noninteractive Group
Gender	47% female	43% female
Age (months)	53.28 (4.24)	53.41 (5.17)
Peabody Picture Vocabulary Test standard score	84.42 (17.84)	89.83 (18.59)

Research Design

Coviewing was a between-subjects comparison. Children were randomly assigned to view five educational media clips in one of two conditions: the coviewing condition, during which a researcher discussed the program with the child, or the noninteractive condition, during which the researcher did not interact with the child during the program. In both conditions, children participated in the study individually in a session with a researcher.

There were no significant differences in PPVT scores, age, or gender distribution based on coviewing condition, confirming that the resultant groups were equivalent on demographic and baseline language variables. Vocabulary clips were presented in six different orders, based on a Latin square design, to ensure effects were not limited to a single order of presentation.

Educational Media Episode and Word Selection

We used five vocabulary-focused segments from the television show *Sesame Street: Word on the Street*. Prior research has suggested that children can learn not just nouns but also verbs and adjectives from educational media (Rice & Woodsmall, 1988; Roseberry, Hirsh-Pasel, Parish-Morris, & Golinkoff, 2009). We therefore included a sample of different word types in the present study. Vocabulary words were additionally selected based on qualifying as Tier 2 words (Beck, McKeown, & Kucan, 2002) that are useful to target for vocabulary instruction. To heighten the likelihood that children would not already be familiar with the vocabulary we selected to teach, we only used words that had a low frequency (<3 instance) on ChildFreq, a database that shows the frequency of word occurrences by child age from transcripts in the CHILDES database (MacWhinney, 2014). Our final criteria for vocabulary selection required that the educational media clip provide a clear, ostensive definition of the word; provide examples of the word; and repeat the word at least 10 times.

After implementing the aforementioned criteria for word selection, we used clips teaching the following five vocabulary words in the present study: *absorb*, *amphibian*, *fragile*, *ingredient*, and *strenuous*. Vocabulary clips averaged 1 minute, 45 seconds per clip. In each clip, a special guest star (a famous actor) and a *Sesame Street* puppet (e.g., Elmo)

discussed the target word by defining the word and discussing examples of the word. Target word labels (e.g., *strenuous*) were repeated 10 to 14 times per clip, and the auditory label for an example of the target word (e.g., *running*) was provided two to four times per clip. Visual depictions of the examples were always on screen when the examples were named.

Measures

Screening measure. Participants received a screening measure prior to the study to ensure they did not already know the target words being taught in the study. During this assessment, children viewed 13 pictures one at a time while an assessor asked them what each one was. The 13 pictures included the five target vocabulary words along with eight foils. Six children correctly identified one or more of the target words on the screening and therefore did not participate in the remainder of the study.

Peabody Picture Vocabulary Test. The Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007) is a validated, norm-referenced instrument that was used as a baseline assessment of receptive vocabulary. Reliability of the instrument ranges from .91 to .94. For this study, age-standardized scores were used as an indicator of baseline vocabulary.

Posttest vocabulary assessment. Children's learning from the educational media clips was assessed through a 30-item assessment that included three question types: word labeling (10 items), concept understanding (10 items), and auditory word example accuracy (10 items). Reliability for the posttest vocabulary assessment was $\alpha = .72$. Details of each question type are outlined in the following.

Audiovisual word labeling. Word labeling was assessed in a format similar to the PPVT, where children were shown three images and asked to point to the image that represented the target word. For example, to determine whether children acquired an understanding of the visual representation of the word *fragile*, they were shown a picture of a fragile object (e.g., a glass) along with two distractors that were perceptually and/or thematically similar (e.g., a nonbreakable pot and a plastic water bottle). They were then asked, "Point to *fragile*."

Audiovisual concept understanding. Concept understanding was assessed in the same format as the word labeling assessment but tested whether children could accurately select the picture that represented the meaning of the word rather than the vocabulary word itself. For example, to assess the understanding of the concept of fragile, children were asked, "Point to the one that you need to be careful with." Similarly, to assess concept understanding of the word ingredient, children were asked "Point to the one that is part of a recipe."

Auditory word example accuracy. Word example accuracy included yes/no questions about whether a certain object was a valid example of a target vocabulary word. This assessment did not use any visual aids and instead assessed whether children could link the verbal word label with an auditory example without a visual representation. For example, to assess this connection for the word *fragile*, we asked “Is a glass *fragile*?” A sample question for the word *strenuous* was “Is sleeping *strenuous*?”

Procedure

Trained graduate student assessors administered all assessments individually to children in a quiet location at their preschool. Children in the eligible age range were initially administered the screening tool for target words. Those who did not answer any of the screening target word questions correctly were then administered the PPVT to obtain their baseline standardized receptive vocabulary.

Children were then randomly assigned to the coviewing condition or the noninteractive condition. They viewed the five video clips in their respective condition. To reduce the effects of fatigue, children were shown the five clips in two blocks. Three clips were shown in the first block, and children received the posttest assessment for those three clips immediately after viewing the third clip. After a short break, children viewed the remaining clips and completed the posttest assessment for those clips. Clips were presented in six different orders to ensure effects were not tied to the order of presentation. The coviewing and noninteractive conditions are described in detail in the following sections.

Coviewing Condition. In the coviewing condition, children viewed the video clips with a trained graduate student on a laptop computer. The graduate student assessor engaged the child in naturalistic dialogue during the video clips.

Training of graduate student coviewers. Graduate students were trained by the authors of this study and a professional actress to include the following interaction elements while coviewing with children: repeating the keyword four to five times throughout the video clip, providing attention-directing cues (“Wow, look at that ingredient!”), physically pointing to the screen, repeating the examples provided (“It’s a toad! The toad is an amphibian!”), laughing at funny moments in the video, and making eye contact with the children. The graduate student did not ask the child any questions that require a verbal response.

During training, graduate students watched exemplar videos of the actress coviewing with another person. They then rehearsed the coviewing experience with each video multiple times to make the experience as natural as possible. They demonstrated the coviewing condition to the professional actress who provided feedback and helpful tips for improvement.

Sample coviewing scenario. The following describes a sample portion of the adult-child naturalistic coviewing experience employed in this study:

[Educational media clip]: A famous actor and *Sesame Street* puppet (Abbie) appear on screen and introduce themselves.

[Coviewer]: “Oh! It’s Abbie! I love her!” (early interaction to elicit interest)

[Educational media clip]: Introduction of vocabulary word: *amphibian*.

[Coviewer]: “Ooh, amphibian! I wonder what that is . . .” (attention/term repetition)

[Educational media clip]: Characters define the word *amphibian*.

[Coviewer]: Wow! It breathes through its moist skin? (concept repetition)

[Educational media clip]: Characters provide an example of an *amphibian*: a toad.

[Coviewer]: “It’s a toad! A toad is an amphibian! (Points to toad)” (attention/repetition)

As shown in the excerpt, the coviewing interactions were highly focused on providing attention-directing statements and gestures as well as repetition of the vocabulary terms, examples, and concepts. The coviewing interaction was not completely scripted but rather involved coviewers incorporating the same elements into the interaction with enough flexibility to account for spontaneous child talk. A total of five graduate students served as coviewers for this protocol. The third author made random spot-checks during data collection to ensure fidelity to the coviewing protocol and found that all assessors maintained strong fidelity to the coviewing protocol. As an added measure to ensure consistency, we used a one-way analysis of variance to compare learning on the three child assessments based on coviewer and found no significant differences ($ps > .3$) in child learning based on coviewer.

Noninteractive Condition. Children assigned to the noninteractive condition viewed the clips on a laptop computer without any adult interaction. Graduate student assessors told participating children that they would be watching some videos and answering some questions afterward. Children then watched the videos on a laptop computer. The assessor remained in close proximity to the child but did not make eye contact or interact with the child while the videos were playing.

Analysis

We conducted the following analyses to determine whether low-income preschoolers learned from educational media, whether coviewing benefited learning of different

TABLE 2
One-Sample t Tests Against Chance Values for the Full Sample

Question Type	<i>M</i>	<i>SD</i>	Chance	95% CI for		<i>t</i>	<i>df</i>	Significance
			Level	Mean Difference				
Labeling*	.57	.23	.34	.19, .27		11.14*	125	<.001
Concept understanding*	.66	.21	.34	.28, .35		16.79*	124	<.001
Auditory example questions*	.62	.14	.50	.09, .14		9.14*	125	<.001

* $p < .05$.

vocabulary associations, and how prior vocabulary size impacted learning. For all analyses, we converted posttest raw accuracy scores to proportion of items correct for each item type (word labeling, concept understanding, auditory word example accuracy). To assess whether children developed an understanding of the words taught in the educational media clips, we computed one-sample *t* tests against chance values for each assessment type.

We analyzed the influence of coviewing using multivariate analysis of covariance (MANCOVA) with coviewing condition (2: coviewing, noninteractive) as a between-subjects factor, age as a covariate to account for possible developmental variation, and the three posttest accuracy proportions as the dependent variables. We additionally included a median split on PPVT standard scores as a between-subjects factor (2: lower PPVT, higher PPVT) to explore whether learning from educational media or the effects of coviewing differed on a baseline language factor. The mean PPVT standard score for the lower PPVT group was 72.80 ($SD = 12.67$), almost two standard deviations below the standardized population mean. The higher PPVT group averaged 101.81 ($SD = 9.51$), roughly equal to the standardized population mean.

Results

In the present study, we sought to answer three primary questions: (a) Do children learn vocabulary from educational media, (b) is the learning of auditory-only and/or audiovisual vocabulary associations facilitated by watching educational media with an adult coviewer, and (c) are there differences in learning or the impact of coviewing based on prior vocabulary size? In particular, does coviewing predominantly benefit children with lower initial language skills? Results pertaining to each question are reported in the following sections.

Vocabulary Learning From Educational Media

We first analyzed whether children developed an understanding of the vocabulary words taught in the educational media segments overall. Three levels of vocabulary knowledge were assessed: audiovisual labeling, audiovisual concept

understanding, and auditory examples. An understanding of vocabulary words was defined as performing statistically significantly above chance levels on the assessment. To determine if participant scores were significantly above chance levels, we conducted one-sample *t* tests with chance level as the comparison value for each of the three question types. Analyses revealed that children successfully understood vocabulary taught in educational media for all three question types: labeling, $t(125) = 11.14, p < .001$; concept understanding, $t(124) = 16.79, p < .001$; and auditory example questions, $t(125) = 9.14, p < .001$ (See Table 2).

In this manner, results indicate that low-income preschoolers can benefit from vocabulary exposure through educational media overall. They were able to accurately select a picture that represented the vocabulary word (labeling) as well as the concept represented by the vocabulary word (concept understanding). They were similarly able to correctly answer yes/no questions asking if an object (e.g., glass) was an example of the vocabulary word being tested (e.g., fragile), which was a question type that did not have any visual aids. As such, after a single viewing of educational media vocabulary clips, low-income preschoolers linked the pictorial representation of vocabulary terms to the terms themselves, understood their meanings, and remembered auditory examples of the vocabulary words.

We also sought to determine if the children with only the lower PPVT scores genuinely understood vocabulary by conducting one-sample *t* tests including only the lower PPVT group against chance levels. Results mirrored the overall findings for the three question types: Children with lower PPVT scores performed significantly above chance levels on labeling, $t(61) = 5.60, p < .001$; concept understanding, $t(61) = 8.97, p < .001$; and auditory example questions, $t(62) = 5.35, p < .001$. We thus verified that even children with lower PPVT scores understood the vocabulary from educational media (see Table 3 for results for lower PPVT group).

Impact of Coviewing Educational Media

We next assessed whether coviewing enhanced vocabulary understanding. A MANCOVA with coviewing condition (2: coview, noninteractive) and PPVT scores (2: lower,

TABLE 3

One-Sample *t* Tests Against Chance Values for the Lower Peabody Picture Vocabulary Test Group

Question Type	<i>M</i>	<i>SD</i>	Chance	95% CI for		<i>t</i>	<i>df</i>	Significance
			Level	Mean Difference				
Labeling*	.50	.23	.34	.10	.22	5.60*	61	<.001
Concept understanding*	.58	.21	.34	.18	.29	8.97*	61	<.001
Auditory example questions*	.58	.14	.50	.05	.12	4.60*	62	<.001

**p* < .05.

TABLE 4

Means and Standard Deviations of Vocabulary Assessments (Proportion of Questions Correct) by Coviewing Group

	Coviewing		Noninteractive	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Labeling	.56	.25	.57	.20
Concept understanding	.66	.21	.65	.21
Auditory example questions	.63	.13	.61	.15

higher) and the covariate of age was conducted to assess whether coviewing impacted vocabulary understanding for low-income preschoolers. Results indicated that coviewing was neither facilitative nor detrimental to vocabulary understandings overall. Participants who viewed the videos with an adult and children who viewed the videos without adult interaction performed equivalently on all three question types: audiovisual labeling, $F(1, 117) = .03, p = .858$; audiovisual concept understanding, $F(1, 117) = .87, p = .354$; and auditory example questions, $F(1, 117) = 1.68, p = .197$ (see Table 4).

This study therefore demonstrated that after viewing educational media clips teaching specific words, children demonstrated an understanding of those words in general. Additionally, viewing the media with an adult who provided concurrent instruction (e.g., by providing the vocabulary labels additional times, repeating the examples and definitions) failed to produce an added benefit overall for audiovisual or auditory-only questions.

The role of baseline vocabulary. In light of the finding that coviewing had no impact on word knowledge in the overall sample, we investigated how extant receptive vocabulary (PPVT scores) impacted vocabulary understanding and the influence of coviewing. The same MANCOVA with the independent variables of coviewing condition (2: coview, noninteractive) and PPVT scores (2: lower, higher) and the covariate of age revealed that extant receptive vocabulary, as measured by the PPVT, was related to vocabulary understanding. Children with higher initial PPVT scores

TABLE 5

Means and Standard Deviations of Vocabulary Assessments (Proportion of Questions Correct) by Peabody Picture Vocabulary Test (PPVT) Group

	Lower PPVT		Higher PPVT	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Labeling*	.50	.23	.62	.21
Concept understanding*	.58	.21	.73	.19
Auditory example questions*	.58	.14	.65	.13

**p* < .05.

demonstrated a stronger understanding of vocabulary words taught to them through educational media than those with lower initial PPVT scores on all three question types: labeling, $F(1, 117) = 8.90, p = .003$; concept understanding, $F(1, 117) = 19.62, p < .001$; and auditory example questions, $F(1, 117) = 9.36, p = .003$. As such, those with a larger initial vocabulary size had a stronger word understanding than those with weaker extant vocabularies (see Table 5).

A question of particular interest related to whether or not coviewing had a differential impact on children's performance based on their PPVT scores (see Table 6). We were specifically interested in seeing if the additional support afforded by coviewing would particularly benefit those with weaker initial language skills. We found no interaction between PPVT scores and coviewing condition for the two audiovisual question types: labeling, $F(1, 117) = .00, p = .994$, and concept understanding, $F(1, 117) = .29, p = .591$.

On the final question type, auditory example questions, we found a significant interaction, $F(1, 117) = 5.31, p = .023$, between PPVT scores and coviewing condition. Follow-up analyses revealed that children with higher PPVT scores performed equivalently in the coviewing ($M = .64, SD = .15$) and noninteractive ($M = .69, SD = .15$) conditions, $t(61) = 1.40, p = .166$. Children with lower initial PPVT scores, however, performed significantly better in the coviewing condition ($M = .56, SD = .15$) than the noninteractive condition ($M = .56, SD = .17$), $t(61) = 2.43, p = .018$. In fact, when coviewing with an adult, performance on the auditory

TABLE 6
MANCOVA Inferential Statistics for All Vocabulary Assessments

Dependent Variable	Contrast	Coview Versus Noninteractive Condition					
		Main Effects and Interactions					
		<i>F</i>	<i>df</i>	Significance	MS _{Effect}	SS _{Error}	MS _{Error}
Labeling	Coview condition	.03	1, 117	.858	.002	5.71	.05
	Age	2.21	1, 117	.139	.11		
	PPVT group*	8.90	1, 117	.003	.43		
	Coview × PPVT	<.001	1, 117	.994	<.001		
Concept understanding	Coview condition	.87	1, 117	.354	.03	4.40	.04
	Age*	6.19	1, 117	.014	.23		
	PPVT group*	19.62	1, 117	<.001	.74		
	Coview × PPVT	.29	1, 117	.591	.01		
Auditory example questions	Coview condition	1.68	1, 117	.197	.03	2.15	.02
	Age	.07	1, 117	.793	.001		
	PPVT group*	9.36	1, 117	.003	.17		
	Coview × PPVT*	5.31	1, 117	.023	.10		

Note. SS = sum of squares; MS = mean square; PPVT = Peabody Picture Vocabulary Test.
* $p < .05$.

example questions for the lower and higher PPVT groups was comparable, $t(59) = .48, p = .630$. Inferential statistics for all analyses can be found in Table 6.

In sum, coviewing seemed to promote vocabulary understanding only for those with lower PPVT scores on the auditory-only question type: auditory example questions. Children with weaker baseline vocabularies were better able to demonstrate auditory vocabulary associations with the help of a coviewer and close the gap in learning for this question type. Children with higher baseline vocabularies still outperformed children with weaker vocabularies on audiovisual question types, however.

Developmental Considerations

Finally, due to the variation in the ages of children in our sample, we added age as a covariate in our analyses. We found that only one question type significantly varied by age: concept understanding questions, $F(1, 118) = 6.05, p = .015$. Follow-up analyses revealed that children performed better on concept understanding with increasing age, $r(125) = .23, p = .009$. As such, age did not factor into performance on the two question types that assessed learning of the vocabulary term itself, but older children were more likely to understand the meaning of the word independent of the vocabulary term.

Discussion

The present study aimed to understand how an educational media coviewing intervention impacted different types

of vocabulary knowledge. Clearly defined enactments of coviewing that incorporate strategies commonly used in naturalistic parent-child interactions have rarely been studied, making these findings particularly valuable to our understanding of the effects of coviewing.

Encouragingly, we found that *all* children—those with both higher and lower baseline vocabularies—were able to demonstrate both audiovisual and auditory-only vocabulary associations after a single, short exposure to educational media clips. This finding is consistent with prior research showing that children can learn vocabulary from educational media (e.g., Shamir et al., 2010; Verhallen & Bus, 2010) as well as research that shows that word learning can take place from a brief exposure to the word (Oetting, Rice, & Swank, 1994).

Dual coding theory (Clark & Paivio, 1991; Paivio, 1990) similarly suggests that media is particularly beneficial to learning because it combines multiple channels—auditory and visual—to teach the same information. Aligned with this, a meta-analysis conducted by Takacs, Swart, and Bus (2014) showed that multimedia elements scaffold independent vocabulary learning in educational media to the same degree that adult scaffolds provide learning support during shared book reading of traditional books. Our findings verify that children show strong knowledge of vocabulary after being exposed to this information effectively and independently from the audio and video information provided by educational media.

However, we also found that children with stronger initial vocabularies had stronger word knowledge than those with weaker extant vocabularies. This aligns with prior research

showing that those with stronger vocabularies are able to learn new words more easily (e.g., Blewitt et al., 2009; Neuman et al., 2019). As such, it seems that the same exposure to educational media for children with higher and lower baseline vocabularies is likely to exacerbate rather than ameliorate the gap in vocabularies between these groups.

The concerning differences in vocabulary learning based on prior language skills led us to investigate a potential contextual intervention that may help ameliorate this divide. The American Academy of Pediatrics (2016) recommends one such context that could be supportive to children—adult-child coviewing. However, they do not describe what coviewing interactions should actually look like, making it particularly important to understand how different enactments of coviewing impact children’s learning. We therefore studied the impact of a form of coviewing that incorporates naturalistic parent-child interaction elements that has rarely been systematically investigated in prior work.

We found that though coviewing had little overall impact on vocabulary knowledge, it specifically benefited children with lower baseline vocabularies on the auditory-only question-type example questions. This may have been because the assessment used an exclusively auditory format—so children with lower initial language skills may have particularly benefited from the additional auditory input by the coviewer for this question type. The visuals provided by the media clips may have been more critical for successful performance on the visual question formats. This suggests that the dual coding (Paivio, 1990) of auditory and visual information may be optimally supported by the educational media itself while auditory-only associations may benefit from a new mode of auditory input. Working memory models (Baddeley, 1986, 1992) similarly suggest that preschool-aged children struggle with memory strategies primarily in the phonological arena, with their visual strategies being better developed. The present study aligns with the interpretation that auditory vocabulary associations may be particularly scaffolded by the coviewer input.

We additionally found that coviewing enhanced auditory vocabulary associations specifically for children with weaker baseline vocabularies. This suggests that it was primarily children with weaker baseline skills who needed and benefited from the additional auditory scaffolds provided by the coviewer. Vygotsky (1978) suggested that adult scaffolding enhances learning most effectively when the task being learned falls in a child’s zone of proximal development—or the type of task that children cannot learn independently but can with adult support. Since the educational media clips we used were specifically designed to teach vocabulary to our studied age group, they provided a great deal of repetition within the media itself. As such, it is possible that children with stronger vocabularies were able to learn from the media itself independently while the same learning task fell into the zones of proximal development of the children with weaker vocabularies.

Prior research on coviewing showed fairly mixed findings, but little work specifically looked at our enactment of coviewing using both auditory-only and audiovisual assessments. Some prior work on a very different enactment of coviewing involving questioning techniques (e.g., Strouse et al., 2013) has found more general success for learning. These techniques involved far more training and discussion than our coviewing enactment, suggesting that a more intensive coviewing intervention may produce stronger and more general benefits to learning. Our intervention was designed to incorporate naturalistic parent-child interaction elements rather than more intensive techniques that are not used by parents spontaneously. This naturalistic form suggests a more targeted auditory benefit to children who have lower baseline vocabularies. Audiovisual associations, however, were not benefited by coviewing.

In lieu of questioning techniques, our intervention incorporated strategies frequently used in naturalistic parent-child shared book reading such as repeating the word label, repeating the visualized examples, and pointing. We therefore predicted that our intervention should still produce benefits to learning, potentially on even audiovisual formats. One potential reason we did not find such ubiquitous effects may be that the coviewer distracted the child from the audio in the educational media segments while they were talking. Strouse et al. (2013) trained parents to pause the video when conversing with their child, preventing this element of distraction. However, pausing the video is unlikely to occur in naturalistic situations, so in an effort to remain as naturalistic as possible, we did not pause the video in our study. Since children could not attend to the audio provided by both the media and the coviewer simultaneously, the coviewer may have distracted children from the educational content of the media while speaking. To investigate this possibility, we are currently studying children’s attention to screen media while coviewing using an eye-tracker.

Another possible consideration about why our coviewing enactment did not produce learning benefits across all children and assessment formats is the nature of the coviewer. We used trained assessors as coviewers rather than naturalistic interaction partners such as parents, teachers, or siblings. We utilized a more structured protocol with trained assessors primarily to ensure consistency in the strategies used while coviewing to avoid the limitations of prior naturalistic parent-child coviewing research that could not define the nature of the coviewing interactions. Nonetheless, our enactment was designed to use strategies common in parent-child interactions, and our findings mirrored those of prior work using parents as coviewers (e.g., Rasmussen et al., 2016).

The present study contributed to our understanding of the impact of a naturalistic form of coviewing educational media on vocabulary knowledge, but we had some limitations. The coviewers were not as familiar to children as parents or teachers. As such, children may have responded differently

to the interaction than they would have with a more familiar adult. Even so, coviewers used similar strategies to those observed in parent-child interactions, so the comfort and familiarity of the interaction style likely helped children respond naturally to the coviewing experience. Another limitation is that we only tested learning at one time point, immediately following the viewing of videos. Future research should investigate retention of learned information over a longer timeframe. We also used only one educational media program (*Sesame Street: Word on the Street*), so we suggest caution in generalizing the findings to very different educational media programs. Finally, though we screened children for an understanding of the target vocabulary, children were not pretested in all question formats, so there may have been some variation in extant partial understandings of words prior to the study. However, through random assignment, these possible variations were unlikely to systematically differ between groups. Nonetheless, our findings are well aligned with prior research and theory studying different forms of educational media, suggesting that the media we used are fairly representative of the types of educational media available to preschoolers.

In summary, this study demonstrates both the facilitative effects of viewing educational media and the potential strengths and limitations of naturalistic enactments of coviewing for enhancing vocabulary learning in low-income preschoolers. On a positive note, we found that children with both stronger and weaker initial language skills demonstrated an understanding of vocabulary words taught in educational media after viewing the media independently. We also found that coviewing supported the growth of auditory vocabulary connections specifically for children with lower baseline vocabularies, closing the extant vocabulary gap in learning for the auditory assessment type. The bad news, however, was that baseline vocabulary still exerted a strong influence on the audiovisual question formats that could not be overcome by coviewing. Overall, though we have further to go in helping children maximize their learning from educational media, we suggest that educational media holds promise for exposing children to new vocabulary and that coviewing can, in some cases, support the children who need it the most.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by an IES Exploration grant R305a150143.

References

American Academy of Pediatrics. (2016). Media and young minds. *Pediatrics*, *138*(5). doi:10.1542/peds.2016-2591

Baddeley, A. (1986). *Oxford psychology series, No. 11. Working memory*. New York, NY: Clarendon Press/Oxford University Press.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559. doi:10.1126/science.1736359

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

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, *101*(2), 294–304.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210.

Dunn, D. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary test (PPVT)*. New York, NY: Pearson.

Evans, M. A., Reynolds, K., Shaw, D., & Pursoo, T. (2011). Parental explanations of vocabulary during shared book reading: A missed opportunity. *First Language*, *31*, 195–213.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, *33*(3), 464–497. doi:10.1016/j.ssresearch.2003.08.001

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, D. K., . . . Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *Journal of Speech-Language Pathology*, *26*, 248–265. doi:10.1044/2016_AJSLP-15-0169

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.

MacWhinney, B. (2014). *The CHILDES project: Tools for analyzing talk, Volume II: The database*. New York, NY: Psychology Press.

Mares, M.-L., & Pan, Z. (2013). Effects of *Sesame Street*: A meta-analysis of children's learning in 15 countries. *Journal of Applied Developmental Psychology*, *34*(3), 140–151.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, *80*, 300–335.

Neuman, S. B., Wong, K. M., 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. doi:10.1037/edu0000278

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

Oetting, J. B., Rice, M. L., & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech, Language, and Hearing Research*, *38*(2), 434–445.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.

Palmer, S. (2000). Working memory: A developmental study of phonological recoding. *Memory*, *8*, 179–193. doi:10.1080/096582100387597

Rasmussen, E. E., Shafer, A., Colwell, M. J., White, S., Punyanunt-Carter, N., Densley, R. L., & Wright, H. (2016). Relation between active mediation, exposure to *Daniel Tiger's Neighborhood*, and US preschoolers' social and emotional development. *Journal of Children and Media*, *10*(4), 443–461.

- Reiser, R. A., Tessmer, M. A., & Phelps, P. C. (1984). Adult-child interaction in children's learning from *Sesame Street*. *ECTJ*, 32(4), 217–223.
- Rice, M. L., Conti-Ramsden, G., & Snow, C. (1990). Preschoolers' QUIL: Quick incidental learning of words. *Children's Language*, 7, 171–195.
- 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 2017*. San Francisco, CA: Common Sense Media.
- Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development*, 80(5), 1360–1375.
- Salomon, G. (1977). Effects of encouraging Israeli mothers to co-observe *Sesame Street* with their five-year-olds. *Child Development*, 48, 1146–1151.
- Senechal, M., Thomas, E., & Monker, J. A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, 87, 218–229.
- Shamir, A., Korat, O., & Fellah, R. (2012). Promoting vocabulary, phonological awareness and concept about print among children at risk for learning disability: Can e-books help? *Reading and Writing*, 25(1), 45–69.
- Skouteris, H., & Kelly, L. (2006). Repeated-viewing and co-viewing of an animated video: An examination of factors that impact on young children's comprehension of video content. *Australian Journal of Early Childhood*, 31(3), 22–30.
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, 38(6), 934–947.
- Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology*, 49(12), 2368–2382.
- 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. doi:10.3389/fpsyg.2014.01366
- Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A meta-analysis. *Review of Educational Research*, 85(4), 698–739.
- Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science*, 20(6).
- Vandewater, E. A., & Bickham, D. S. (2004). The impact of educational television on young children's reading in the context of family stress. *Journal of Applied Developmental Psychology*, 25, 717–728. doi:10.1016/j.appdev.2004.09.006
- Verhallen, M. J., & Bus, A. G. (2010). Low-income immigrant pupils learning vocabulary through digital picture storybooks. *Journal of Educational Psychology*, 102(1), 54–61.
- Verhallen, M. 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.
- Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children*, 23(3), 34–41.
- Wagmiller, R. L., Lennon, M. C., Kuang, L., Alberti, P. M., & Lawrence Aber, J. (2006). The dynamics of economic disadvantage and children's life chances. *American Sociological Review*, 71(5), 847–866. doi:10.1177/000312240607100507
- Watkins, B. A., Calvert, S., Huston-Stein, A., & Wright, J. C. (1980). Children's recall of television material: Effects of presentation mode and adult labeling. *Developmental Psychology*, 16, 672–674. doi:10.1037/0012-1649.16.6.672

Authors

PREETI G. SAMUDRA is a postdoctoral associate in the Department of Teaching and Learning at New York University-Steinhardt. She is a developmental psychologist whose research focuses on supporting the language and literacy development of diverse learners.

RACHEL M. FLYNN is a research assistant professor in the Department of Medical Social Sciences and the associate director for the Institute for Innovations in Developmental Sciences at Northwestern University. She is a developmental psychologist whose research focuses on the influences of media and technology on children's cognition and learning.

KEVIN M. WONG is a PhD candidate at New York University and clinical assistant professor at Monroe College. His research interests include educational media and vocabulary development for young dual-language learners.