Journal of Applied Developmental Psychology 67 (2020) 101108

Contents lists available at ScienceDirect

Journal of Applied Developmental Psychology

journal homepage: www.elsevier.com/locate/jappdp

Is attention the missing link? Coviewing and preschoolers' comprehension of educational media

Preeti G. Samudra[a,*], Kevin M. Wong[b], Susan B. Neuman[b]

[a] State University of New York, Plattsburgh, USA
[b] New York University, USA

ARTICLE INFO

Keywords:
Coviewing
Attention
Comprehension
Educational media
Vocabulary

ABSTRACT

Coviewing is a commonly recommended practice, but little is known about how coviewing impacts children's educational media viewing experience. We investigated how coviewing impacts attention and comprehension of educational media, as well as the role of baseline vocabulary in understanding these associations. Eighty-three preschoolers viewed two videos on an eye-tracker – one with an adult coviewer and the other without. Children's baseline vocabulary, attention, and comprehension were assessed. Results indicated that coviewing benefited visual attention. Neither coviewing condition nor attention, however, predicted children's comprehension. Instead, comprehension was predicted by age, vocabulary, and an interaction between coviewing condition, vocabulary, and attention. The interaction revealed that comprehension was stronger in the coviewing condition than the noninteractive condition only when children also had stronger visual attention to the program and larger vocabularies. Results suggest that coviewing benefits attention, but that both attention and child language are integrally tied to whether coviewing predicts comprehension.

Introduction

Young children are avid consumers of media in today's society – with children aged two to four watching over 2 h of television per day (Rideout, 2017). Fortunately, media targeting preschoolers often have educationally relevant goals, and preschool-aged children are skilled at comprehending and learning from these educational media programs, even when viewing media alone (Mares & Pan, 2013; Takacs, Swart, & Bus, 2015). Nonetheless, recommendations regarding children's screen media use, such as those by the American Academy of Pediatrics, 2016, suggest that parents should coview media with preschoolers to help them better understand what they see. Coviewing may be beneficial in many ways, including allowing parents to discuss and potentially mitigate any harmful effects of exposure to violence or risk-taking behaviors in media programming. However, in the context of educational media – programs that have the explicit intent to teach children a school-related skill rather than be primarily entertaining (Vandewater & Bickham, 2004) – it is less clear if coviewing enhances the viewing and educational experience for young children. In our study, we focus on the role of coviewing in enhancing the learning environment of video programs viewed on screen-based educational media platforms, such as television, streamed videos, iPads, and smart phones.

In fact, past research on the learning benefits of coviewing educational media has been largely inconclusive – some studies have found learning benefits to coviewing over viewing media alone, while others have found no added learning benefit to coviewing (Reiser, Tessmer, & Phelps, 1984; Skouteris & Kelly, 2006; Strouse, O'doherty, & Troseth, 2013). Additionally, little work examines how coviewing might impact the processes involved with media consumption and comprehension, such as child attention. Attention is necessary though not sufficient for understanding the content displayed on screen (Smith, Colunga, & Yoshida, 2010), so increasing our understanding of how coviewing impacts visual attention might help us learn more about the proximal effects of coviewing on children's viewing experience. Much like the lack of process-level data, little work examines how coviewing interacts with child characteristics, such as baseline vocabulary size. In order to comprehend the narrative of an educational media program, children need not only attention to the media, but also the language necessary to understand the media narration along with the added coviewer speech. The present study therefore investigates how coviewing impacts attention and comprehension, as well as the role of child baseline vocabulary in understanding these associations.

The overall goal of the present study is to add to our understanding of the coviewing process by investigating how a clearly defined, educationally-relevant form of coviewing – one that reflects processes that are commonly seen in parent-child interactions – impacts low-income

* Corresponding author.
E-mail address: psamu003@plattsburgh.edu (P.G. Samudra).

https://doi.org/10.1016/j.appdev.2019.101108
Received 31 January 2019; Received in revised form 23 December 2019; Accepted 27 December 2019

preschoolers' attention to and comprehension of educational media. We chose to focus on a sample of preschoolers from lower-income households in order to better understand a potentially supportive context for the children who may be in greater need of additional scaffolds. Prior research has shown consistent differences in language processing, production, and comprehension based on socioeconomic differences from an early age (Ginsborg, 2006; Pace, Luo, Hirsh-Pasek, & Golinkoff, 2017). At the same time, children from lower socioeconomic status households tend to consume more media than their peers (Rideout, 2017). Considering the human resources needed for coviewing and the inconsistent findings of prior research on the benefits of coviewing, understanding how best to invest resources to support the comprehension of children from lower-income households is particularly important. Within our sample, we therefore also investigate the role of children's baseline vocabulary in predicting attention and comprehension. Prior research has found that extant vocabulary plays an important role in predicting learning from media (e.g. Blewitt, Rump, Shealy, & Cook, 2009). As such, in a sample at risk for weaker language skills, we recognized the need to understand if the added language input of a coviewer might differentially support children based on the language competence they initially bring to the experience. In other words, we investigated whether coviewing might be a stronger support for children with high enough language skills to process two sources of language input. We therefore investigated the potential interactions between coviewing, attention, and baseline vocabulary that could help illuminate the circumstances under which coviewing relates to children's comprehension of educational media.

Coviewing educational media

There are multiple pathways through which coviewing might benefit attention and comprehension of media. Salomon (1977), for example, discussed how parent-child coviewing promoted enjoyment of viewing for both parties, which in turn could support children's attention to the program. Coviewers could additionally provide discussion and help as needed, thereby supporting comprehension of media content. For example, providing children with repetition and elaboration of important plot information could enhance the amount of input they receive related to the plot content, supporting their level of rehearsal of the content and comprehension of it (Watkins, Huston-Stein, & Wright, 1981).

In spite of the comprehension support coviewing might provide, research on the impact of coviewing on preschoolers' comprehension has produced fairly inconsistent results. Reiser et al. (1984) found that three- and four- year-old 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 researcher. This study employed an intensive questioning approach to coviewing, however, that is unlikely to be found under more naturalistic circumstances. Similarly, in a study of three-year-olds' video-storybook comprehension, Strouse et al., 2013 investigated two forms of coviewing, and found that one form – an intensive coviewing intervention that trained parents to pause the video and engage in dialogic questioning with their child – resulted in greater comprehension than a control. However, Strouse et al. (2013) discussed that parents rarely employed these questioning techniques spontaneously. As such, intensive coviewing interventions that are characterized by relatively non-naturalistic methods of child questioning have often demonstrated benefits of coviewing for comprehension.

Contrastingly, Strouse et al., 2013 found that their other studied coviewing enactment – one in which parents did not ask questions and instead directed their child's attention and discussed the program with them – did not promote comprehension. Similarly, other studies employing more naturally-occurring coviewing enactments with three- and four-year-old children have yielded no overall benefits to comprehension or learning for experimentally manipulated coviewing over viewing alone (Rasmussen et al., 2016; Skouteris & Kelly, 2006). Unfortunately, these naturalistic studies did not describe how parents enacted coviewing, and instructions given to parents were quite general (e.g. “talk to your child as much as possible about the show”). As such, even though these studies had the strength of employing naturalistic parent-child coviewing, the lack of information on enactment or process variables makes it difficult to understand the characteristics of coviewing that failed to produce comprehension gains for children or why that may have been.

This presents a potential problem for policies such as that by the American Academy of Pediatrics that recommend coviewing. Policymakers rarely suggest specific strategies to use while coviewing, and the literature suggests that some of the strategies that parents use spontaneously may not be effective in supporting children's learning and comprehension. In fact, in a correlational study, Rice, Conti-Ramsden, and Snow (1990) found that viewing “Sesame Street” alone over two years was related to improved vocabulary gains, whereas viewing the show with an adult was unassociated with vocabulary improvements. This highlights a critical need to better understand how coviewing enactments that employ techniques that are typical of parent-child interactions impact not only comprehension, but also process variables (e.g. attention).

Coviewing enactment of present study

In the present study, we systematically investigate how a form of adult-child coviewing that incorporates elements commonly used in parent-child interactions influences both attention to and comprehension of educational media. In order to develop our coviewing enactment, we therefore drew on research documenting naturally occurring, positive parent-child interaction elements during shared book reading. These elements included pointing, discussing important word meanings, making comments or connections to help children understand past story elements, and making comments related to opinions or reactions related to story content (Evans, Reynolds, Shaw, & Pursoo, 2011; Fisch, Shulman, Akerman, & Levin, 2002; Ninio & Bruner, 1978; Roser & Martinez, 1985). We incorporated the aforementioned elements into our enactment of coviewing.

Another practice that often occurred during shared book reading was parents asking their child question questions (e.g. Fisch et al., 2002). Questioning has been a central feature of prior coviewing research that has demonstrated learning benefits to coviewing (e.g. Reiser et al., 1984; Strouse et al., 2013). However, we did not utilize questioning in our enactment due to the contextual differences between an educational media environment such as video and the traditional storybook environment. The pacing of traditional storybooks is self-determined, and pauses for discussion or questioning are easy to spontaneously embed within the interaction. Extensive questioning and discussion is not as well suited to a video viewing environment because the discussion tends to result in the child missing the content of the video that follows. It is therefore not a common practice to naturally pause a video to discuss its content. For example, the dialogic questioning coviewing intervention studied by Strouse et al., 2013 involved training parents to pause the video in order to engage in extensive questioning.

In order to maintain a more natural media consumption experience, we only utilized the elements of shared book reading that were best suited to the contextual constraints of the media environment – pointing, important vocabulary discussions, providing brief plot recaps and comments, reacting to the program content, and briefly elaborating on content. We provided a clearly defined, scripted, and focused coviewing procedure that represented a strong, educationally-relevant enactment of frequent parent-child interaction elements in a coviewing context. We therefore studied whether a rich use of these strategies would support the attention and comprehension of preschoolers who

were coviewing educational media with an adult.

Within the context of our coviewing enactment, we additionally sought to investigate potential reasons for why educationally-relevant coviewing may not necessarily support children's comprehension of educational media. There are multiple reasons why coviewing might not benefit comprehension. In the present study, we focus on two possible interpretations: i) the visual interpretation – that children might become visually distracted by the coviewer when he/she speaks to the child. The detriments of lessened visual attention coupled with the benefits of increased audio input from a coviewer might ultimately level off, and result in few comprehension gains from coviewing. ii) The auditory interpretation – that the additional audio input provided by the coviewer may become overwhelming for children with weaker language skills, and therefore only potentially benefit comprehension for children with stronger baseline vocabularies. If child baseline language, such as vocabulary, plays a role in how effectively children can take advantage of the coviewing experience, failure to take this into consideration might limit the ability to detect potential comprehension benefits to coviewing for certain children. As such, both the visual and auditory interpretations suggest different potential pathways through which the link between coviewing and comprehension might be integrally tied to attention and child baseline vocabulary.

The visual interpretation – child attention to educational media

Though some work has investigated the arousal processes associated with parent-child coviewing such as heart rate (e.g. Keene et al., 2019; Rasmussen, Keene, Berke, Densley, & Loof, 2017), little research has investigated how coviewing impacts child visual attention to an educational media program. Two possible hypotheses emerge on how coviewing educational media might interact with attention. On the one hand, as Salomon, 1977 and Strouse et al. (2013) suggested, a coviewer might enhance child attention by providing a model of attention, increasing interest in the program, and/or directing child attention deliberately. Studies by Keene et al. (2019) and Rasmussen et al. (2017) similarly suggest arousal patterns that reflect stronger engagement when viewing media with an adult compared to alone.

Alternatively, children may be inclined to look away from the screen and towards the coviewer while the coviewer is talking. In a study of peer coviewing, Anderson et al., 1981 found that children viewing educational media in groups of three showed weaker visual attention to television than children viewing without peers. Looking away from the screen would inherently reduce visual attention at those points, and may potentially disrupt visual engagement with the content of the program more generally. Richards and Anderson (2004) discuss how attentional inertia - or sustained looking at the screen without looking away –consistently predicts learning from television. If a coviewer visually distracts the child from the screen by talking to the child, this might interrupt the flow of attentional inertia, which may in turn be detrimental to visual engagement with the program.

In the present study, we therefore use eye-tracking to investigate how coviewing impacts preschoolers' attention to the screen. If our enactment of coviewing reduces child attention to the screen, this might be a potential explanation as to why such forms of coviewing rarely show benefits to comprehension. Nonetheless, prior research has rarely found coviewing to be detrimental to comprehension, so it is also possible that coviewing might have little effect on, or a positive effect on attention. If coviewing facilitates attention, the next step would be to ascertain if attention in turn predicts comprehension.

Unfortunately, the connection between looking time and learning has not been reliably established (Kirkorian, Pempek, & Choi, 2017). In some studies, the total time spent looking at learning-related stimuli were associated with learning (e.g., Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009), while in other cases, looking times did not predict learning (Schmitt & Anderson, 2002). Visual attention is only one of many forms of attention that can predict learning, and, as such, enhanced visual attention alone might not directly translate to enhanced comprehension. We therefore turn to the other sensory source of input – auditory input – and investigate the role of the child characteristic of baseline vocabulary in predicting the conditions under which coviewing might benefit comprehension. Since a coviewer is primarily a source of auditory input, children's language proficiency may be integrally tied to how effectively they can process and use the added input to support their comprehension.

The auditory interpretation – child baseline vocabulary

Our enactment of adult-child coviewing incorporated many auditory elements that are common in parent-child interactions, such as comments on past story events (Evans et al., 2011; Fisch et al., 2002; Ninio & Bruner, 1978). It is therefore possible that children's baseline language skills, such as vocabulary, play an important role in whether or not they are able to take advantage of the added auditory input. An influential theory that delineates this possibility is dual coding theory (Clark & Paivio, 1991; Paivio, 1986, 1990).

Dual coding theory proposes that two different 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 content is beneficial to learning and comprehension. Educational screen media taps into both the visual and auditory channels, providing a more complete representation of the story content than one channel alone.

Central to this theory is the notion that we have limited cognitive processing resources within a single modality. Once our processing resources are being fully utilized, additional input in the same modality would no longer be beneficial. In the context of coviewing educational media, the media itself provides both visual and auditory input, and the coviewer provides an additional source of auditory input. It is therefore possible that when children have weaker initial vocabularies, they may not be able to take advantage of the additional auditory input provided by the coviewer as their cognitive resources are being fully utilized to process the audio content of the media itself.

Children with stronger vocabularies, however, may require fewer resources to process the audio from the media, and may be better able to take advantage of the additional auditory input provided by the coviewer. If this is the case, children might need sufficiently strong language skills to process coviewer sources of auditory input in order for coviewing to support comprehension over viewing alone. In the present study, we therefore investigate whether coviewing might interact with children's baseline vocabulary to predict comprehension of educational media.

We additionally examine whether coviewing interacts with both vocabulary and attention to predict comprehension. Prior research shows that the attention might not only predict comprehension, but the reverse might also be found. Prior vocabulary knowledge and background knowledge have been shown to support the visual attentional processes of children (Anderson, Lorch, Field, & Sanders, 1981; Kaefer, 2018; Kaefer, Pinkham, & Neuman, 2017; Kaefer, Neuman, & Pinkham, 2015). This suggests that baseline factors such as vocabulary may ultimately work alongside visual attention to predict comprehension. Additionally, aligned with dual coding theory that emphasizes the additive nature of visual and auditory input, it may be that children need a combination of a stronger baseline vocabulary to successfully process audio input, and sufficient visual attention for visual processing in order for comprehension to be supported. We investigate these possibilities in the present study.

The present study

The present study focuses on a sample of low-income preschoolers

to extend the literature on coviewing, attention, baseline vocabulary, and comprehension. We investigate how a coviewing enactment incorporating educational elements of parent-child interaction impacts visual attention and story comprehension. Children viewed one educational media episode with, and another without an adult coviewer while being eye-tracked. Children's baseline vocabulary was assessed prior to viewing the videos, and children completed a comprehension assessment after viewing each video. Our study focused on the following questions:

i) Does coviewing impact visual attention to educational media?
ii) Does attention predict comprehension of educational media?
iii) Does coviewing benefit comprehension of educational media?
iv) Does coviewing interact with attention and/or child baseline vocabulary to predict stronger comprehension?

Overall, the present study aims to move beyond exclusively investigating the direct influence of coviewing on comprehension to developing a more nuanced understanding of the conditions under which coviewing might be more likely to support comprehension in preschoolers.

Method

Participants

Participants were 83 three- and four-year old children (M_{age} = 4.3 years, SD_{age} = 0.37 years; range = 42–59 months); 64% were female. Sample size determinations were made based on recommendations by Morgan and Case (2013) who suggest that a conservative sample size estimate for a repeated measures analysis of covariance can be approximated as a 44% reduction of the power estimates for a two-sample *t*-test. Power analyses in G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) for a 2-tailed, two-sample t-test with an estimated power of 0.8 suggested a sample size of 128. The 44% reduction resulted in a desired sample of 72 children in our study. We also used linear mixed modeling in our sample, and therefore used G*Power to determine the number of data points needed when using a model fully controlling for subject (the number of predictors in the model plus the number of participants, each representing a dummy-coded variable). A 2-tailed linear multiple regression model with 90 predictors (83 subject plus 7 predictors) at a power of 0.8 required a minimum of 94 data points. Our sample comprised 83 participants, each with two data points, thereby meeting the power requirements of this analysis.

Participating children were enrolled in two Head Start centers located in high poverty areas in a large urban city. The sample was diverse: 29% were African American, 49% were Hispanic, 18% were West Indian, and 4% were Asian or biracial. Educational directors, teachers, and parents provided consent for participation. Children provided verbal assent. IRB approval was attained from New York University (IRB-FY2016–1251, Title: “Educational Media Support for Low-income Preschoolers' Vocabulary Development”). All children qualified for free and reduced lunch. Standardized receptive language scores, as measured by the Peabody Picture Vocabulary Test-IV (PPVT-IV; Dunn & Dunn, 2007), averaged 79.64 (*SD* = 15.76), which is more than one standard deviation below the population mean.

Research design

We employed a within-subjects design in which participating children each viewed two videos – one in the coviewing condition, and the other in the non-interactive condition. In both conditions, children participated in the study individually in a one-on-one session with a researcher. The two conditions were run on the same day spaced approximately one hour apart. Coviewing condition order and the video used in each condition were counterbalanced between participants. Children were eye-tracked while watching both videos in order to determine the influence of coviewing on attention to the screen. Children were administered the PPVT-IV prior to viewing the educational videos as a baseline language indicator.

Educational media episodes

Since we were interested in studying how coviewing impacts comprehension, we used two narrative videos that each focused on a different science-related plot. Our educational media stimuli were developed from two 9.5-min narrative videos from the television show *Peep and the Big Wide World*, a program teaching science concepts to 3–5 year-old children. In one episode, the main characters found a beautiful flower far from home, and decided to grow their own flower closer to home. The bulk of the episode focused on the steps it took to grow the flower from the seed to a fully-grown plant. In the second episode, the characters were looking for buried treasure when they spotted a square in the sand. Upon pulling it out of the sand, they discovered it was actually a three-dimensional block (cube). They found more shapes in the sand, leading one character to think he had a unique treasure-finding ability. The characters learned more about the representation of three-dimensional shapes throughout the episode. The episodes incorporated a strong narrative with clear visual representations of main plot points, which aided in the assessment of comprehension.

Eye-tracking apparatus and data processing

Researchers utilized the child-friendly Tobii T120 system, a remote eye-tracking system that has an infrared-based eye-tracker integrated into a computer with an LCD screen. The Tobii T120 samples at 120 Hz and has an accuracy of 0.5 visual degrees. Children's eye movements while viewing videos in the coviewing and non-interactive conditions were recorded on this eye-tracker. Children were calibrated prior to viewing each video using a 5-point manual calibration on screen. Data were processed and exported using Tobii Studio 3.0.

Measures

Peabody picture vocabulary test – fourth edition (PPVT-IV)

(Dunn & Dunn, 2007). The PPVT is a validated, norm-referenced instrument that was used as a baseline assessment of receptive vocabulary. In this assessment, children are asked to point to one of four image options that depicts a named word. The assessment provides both raw and age-standardized scores as an indicator of baseline vocabulary. Both scores are reported, though only the raw scores are used in the linear mixed model since age is added a separate predictor.

Narrative story comprehension

In order to assess children's narrative story comprehension, assessors showed children six screenshots from each video. Screenshots each depicted an important plot point in the video narrative, and were used to cue children's story recall. For each picture, children were asked, “What happened during this part?” Children provided their responses, and were given an additional prompt, “Anything else?” after their initial response was complete. Assessors wrote down children's responses verbatim for later coding.

All child responses were transcribed, and a trained primary coder coded all comprehension responses by noting the number of accurate statements children made about the story (see Table 1 for examples of child responses and assigned codes). The primary coder was blind to the condition within which each video was viewed. Coded scores were summed across all six pictures to provide a measure of overall narrative story comprehension. A second trained coder independently coded 10% of responses, and inter-rater reliability was established at

Table 1
Comprehension coding examples.

Video	Child comprehension response	Score
Shapes video		
Image: Characters find a shape in the sand.	"They saw a shape"	1
	"They saw something./ They take it out."	2
	"The purple one is sad."	0 (inaccurate)
Seed video		
Image: character planting the seed.	"They are just putting the seed."	1
	"He put the seed in./ He put water/ and it growed up."	3
	"He fell down."	0 (inaccurate)

Kappa = 0.96.

Eye-tracking fixation duration

We were interested in seeing how coviewing impacted visual attention to the screen while viewing the video. We extracted the total fixation duration children spent looking at the screen during the full video. Fixation durations were extracted using Tobii Studio 3.0 software. Fixations were defined as coordinates lasting 60 milliseconds or more, which were also identified by the fixation filter in the software program. Fixation durations, or the amount of time spent looking at a specific location, have frequently been used as an index of attention and processing of visual information (Just & Carpenter, 1980; Tsai, Hou, Lai, Liu, & Yang, 2012).

Procedure

Trained graduate student assessors administered all assessments individually to children in a quiet location at their preschool. Children first completed the PPVT-IV. They were then randomly assigned to a counterbalancing condition (condition order; video in each condition). On a later day, children completed both the coviewing and non-interactive conditions spaced approximately one hour apart. In each condition, children were calibrated on an eye-tracker and watched one 9.5-min video while their eye movements were recorded. The comprehension assessment for the relevant video was administered immediately following the video viewing in both conditions. The coviewing and non-interactive sessions each lasted 20 min. The coviewing and non-interactive conditions are described in detail below.

Coviewing condition

In the coviewing condition, children viewed the video clips with a trained graduate student on an eye-tracker monitor computer. In order to ensure the coviewing enactments were consistent, graduate student assessors were trained to follow a specific coviewing script for each video. The script was designed to engage the child in an educational manner while not being too disruptive.

Coviewing elements

The coviewing script incorporated interaction elements that required the coviewer to provide additional information about important concepts or words in the story, make real-life connections, reiterate certain plot points, and display engagement with the program by reacting to the program in a way that aligned with the content of the program (e.g., by laughing when something funny happens). An excerpt of the video and coviewing script can be found in Table 2. All of the aforementioned interaction elements were fully scripted to ensure consistency of implementation. Comments or questions from children in this study were extremely rare. Nonetheless, in the few cases where the child asked a question or made a comment, coviewers provided a short contingent response to acknowledge that they heard the comment, such as saying "yeah!", "uh-huh" or "mmmm" in an interested tone.

Non-interactive condition

In the non-interactive condition, children viewed the video on the eye-tracker monitor without any adult interaction. Graduate student assessors told participating children that they would be watching a video and answering some questions afterwards. Assessors remained in the room to supervise the child, but made their presence less salient by sitting 10 ft away from the child and pretending to read a book. They did not make eye contact or interact with the child while the video was playing.

Analysis

The present study investigated how coviewing educational media with an adult impacted preschoolers' visual attention and story comprehension. In order to analyze how coviewing impacted visual attention, we conduced a repeated measures analysis of covariance with the dependent variable of visual attention, the within-subjects factor of coviewing condition (2: coviewing, noninteractive) and the (mean-centered) covariates of PPVT raw scores and child age. To answer our remaining research questions investigating whether attention, coviewing, or an interaction between attention, coviewing and/or PPVT scores predict comprehension, we conducted a two-level HLM with participant as the level-2 factor and coviewing condition as the repeated measures factor. The model contained a random L2 intercept as well as the fixed predictors of child age, coviewing condition, PPVT raw score, fixation duration, and all two- and three-way interactions between coviewing condition, PPVT scores, and fixation duration. There were no significant correlations between attention, age, and PPVT scores, verifying low multicollinearity among predictors. All covariates and predictors used in these analyses were mean-centered. Data were analyzed using IBM SPSS Statistics version 25. Post-hoc simple slopes analyses were conducted to interpret interactions using Stata version 15.

Results

The present study investigated the connections between adult-child coviewing, visual attention to screen, baseline language proficiency, and comprehension. We specifically looked at i) whether adult-child coviewing impacted child attention, and ii) the potential predictors of comprehension including coviewing, attention, PPVT scores, and interactions.

Preliminary analyses

In order to determine whether condition order or video used in each condition would need to be used in further analyses, we first determined whether these counterbalanced variables affected children's comprehension or attention. There were no statistically significant differences in comprehension based on video, $F(1, 81) = 3.72$, $p = .057$ or condition order, $F(1, 81) = 1.04$, $p = .312$. Similarly, there were no significant differences in attention between the two condition orders, $F(1, 67) = 2.88$, $p = .094$ or videos, $F(1, 67) = 0.37$, $p = .544$. As such, video and condition order were excluded from all further analyses.

Coviewing and visual attention

We sought to understand how a coviewing affects child attention to educational screen media. We aimed to ascertain which of two competing hypotheses – one predicting that children would be distracted from the video by the coviewer, and the other suggesting that children's interest and therefore attention might be enhanced by a coviewer – would be supported by the data. We conducted a repeated measures ANCOVA with the within-subjects factor of coviewing condition (2: coviewing, noninteractive) and the covariates of PPVT raw score and

Table 2
Excerpt from video and coviewing script.

Seeds video script	Coviewer script
Narrator: Peep did not give up. That's when he learned that waiting is hard to do.	
	It is hard, huh?
Narrator: Peep got water from the stream every day. He watered and waited…	
	Wow, he's working hard to get water from the stream! But that's what helps flowers grow!
Narrator: for days… and days… and days… and days… and days… and days. Quack and Chirp began to worry.	[laugh]
Chirp: Nothing is every going to grow. How can we help Peep?	
Quack: We have to dig up the seed and eat it, so he will give up. Quack: I can't find it!	Oh no!
Chirp: Did he bury it by this green thing?	
Peep: That's it! It grew into a baby plant!	
	Cool! It started growing! They can see it now!

age on children's attention to the educational videos.

Children visually attended to a significantly greater percentage of the program when viewing with a coviewer (M = 73.35, SD = 18.97) than when viewing alone (M = 66.87, SD = 21.60), $F(1, 68) = 9.29$, $p = .003$. There were no significant main effects or interactions with child age or PPVT score, suggesting that attentional processes were similar regardless of baseline language proficiency and age in the 18-month range investigated in this study. See Table 3 for inferential statistics. Overall, results suggest that an interactive adult coviewer did not visually distract children from the screen, but rather strengthened children's visual engagement with the educational media program. Additionally, this stronger visual attention was not related to the child characteristics of age or baseline vocabulary size. We next turned to whether attention, PPVT scores, and/or coviewing might predict comprehension.

Predictors of comprehension

We next analyzed whether coviewing, visual attention, or an interaction between coviewing, attention and/or PPVT raw scores might impact comprehension. We conducted a two-level linear mixed model on children's comprehension, with a L2 factor of participant and a repeated factor of coviewing condition. Fixed predictors in the model were child age, coviewing condition, PPVT raw score, fixation duration, and two- and three-way interactions between coviewing condition, PPVT scores, and fixation duration. This analysis revealed three significant predictors of children's comprehension of educational media: child age, $t(73.26) = 2.07$, $p = .042$, PPVT raw score, $t(74.35) = 4.60$, $p < .001$, and the three-way interaction between coviewing condition, fixation duration, and PPVT standard score, $t(75.63) = 2.38$, $p = .020$. Children received higher comprehension scores with increasing age, $r(166) = 0.30$, $p < .001$, and PPVT scores, $r(166) = 0.40$, $p < .001$. There were no significant differences in comprehension between the coviewing (M = 9.01, SD = 5.34) and noninteractive (M = 8.83, SD = 4.83) conditions, $t(71.31) = 0.003$, $p = .998$. Statistics related to predictors in this model are in Table 4, and correlations between predictors can be found in Table 5.

In order to interpret the 3-way interaction, data were graphed by plotting the comprehension scores of children based on attention to screen and PPVT scores. Fig. 1 shows comprehension scores for each coviewing condition based on a median split of coviewing visual attention and PPVT scores. When comparing the coviewing to the noninteractive condition, comprehension in the coviewing condition began to surpass the noninteractive condition only when children had both stronger vocabularies and higher attention to media, $t(20) = 2.01$, $p = .059$. Little difference was observed by coviewing condition when both vocabulary and attention were low, $t(19) = 0.41$, $p = .689$. When children had only one of the two characteristics (attention or vocabulary), Fig. 2 their comprehension scores were actually stronger in the

Table 4
Estimates and significance of HLM predictors.

Predictors	Estimate	Standard error	df	t	Sig.
Coview condition	< .001	.21	71.43	.003	.998
PPVT*	.12	.03	74.35	4.60	< .001*
Age*	2.75	1.33	73.26	2.07	.042*
Fixation duration	.03	.02	131.28	1.75	.083
Coview condition by fixation duration	.01	.01	72.16	.95	.348
Coview condition by PPVT	.01	.01	69.83	.76	.448
PPVT by fixation duration	.001	.001	115.42	1.17	.243
Coview condition by PPVT by fixation duration*	.002	.001	75.63	2.38	.020*

* $p < .05$.

Table 5
Correlations between comprehension and model predictors.

Variables	1	2	3	4	5	6
1. Comprehension	–					
2. Age	.30**	–				
3. PPVT raw score	.50**	.16*	–			
4. Fixation duration	.20*	.13	.08	–		
5. Coview condition	.02	–	–	.19*	–	

* < 0.05.
** < 0.01.

Table 3
Main effects and interactions for attention by coviewing condition.

		Coview vs. noninteractive condition main effects and interactions					
Dependent variable	Contrast	F	df	Sig.	MS_{Effect}	SS_{Error}	MS_{Error}
Percent fixation duration on overall video	Coview Condition	9.29	1/68	.003*	1537.42	11252.30	165.48
	Coview by Age	.26	1/68	.614	42.37		
	Coview by PPVT	1.06	1/68	.306	175.69		
	Age	.85	1/68	.359	571.90	45522.27	669.45
	PPVT	.17	1/68	.680	114.95		

* $p < .05$.

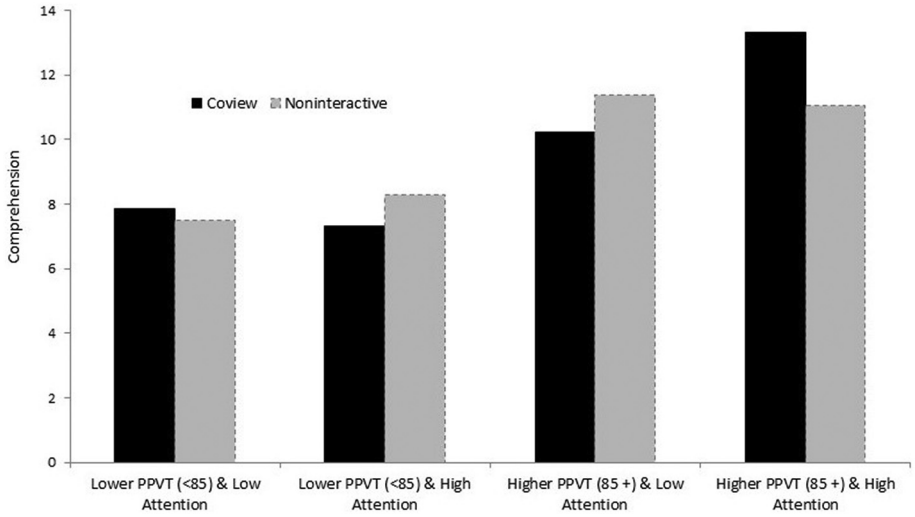

Fig. 1. Children's comprehension with lower and higher attention to coviewing video (median split) by coview condition in each of two PPVT groups (standard score below vs. within 1 SD of 100).

noninteractive condition.

We conducted an analysis of whether the simple slopes of the association between comprehension and PPVT scores varied in different combinations of attention and coviewing condition. Pairwise comparisons between slopes on the six comparisons revealed two significant differences between slopes. Specifically, the children in the coviewing condition with stronger attention to video had a steeper slope relating comprehension and PPVT compared to children with i) high attention in the noninteractive condition ($t = 3.04, p = .003$), or ii) low attention in the coviewing condition ($t = 2.02, p = .048$). The final group – children with low attention in the noninteractive condition did not significantly differ in slope from any other group. As such, we found that the combination of coviewing, high attention, and high vocabulary showed the clearest connection to comprehension. Overall, these results suggest that in order for coviewing to positively predict comprehension, children needed to have both a high enough baseline vocabulary and strong enough attention to video to take advantage of it. Having low attention to the video and/or low language proficiency corresponded with coviewing no longer benefiting comprehension.

Discussion

The present study examined how coviewing educational media impacted visual attention and story comprehension in a sample of low-income preschoolers. We found that coviewing heightened visual attention to the program, but neither attention nor coviewing directly predicted comprehension. Further explorations revealed that comprehension was stronger when coviewing versus viewing independently only when children had both strong enough vocabularies and high enough visual attention to take advantage of the coviewing experience.

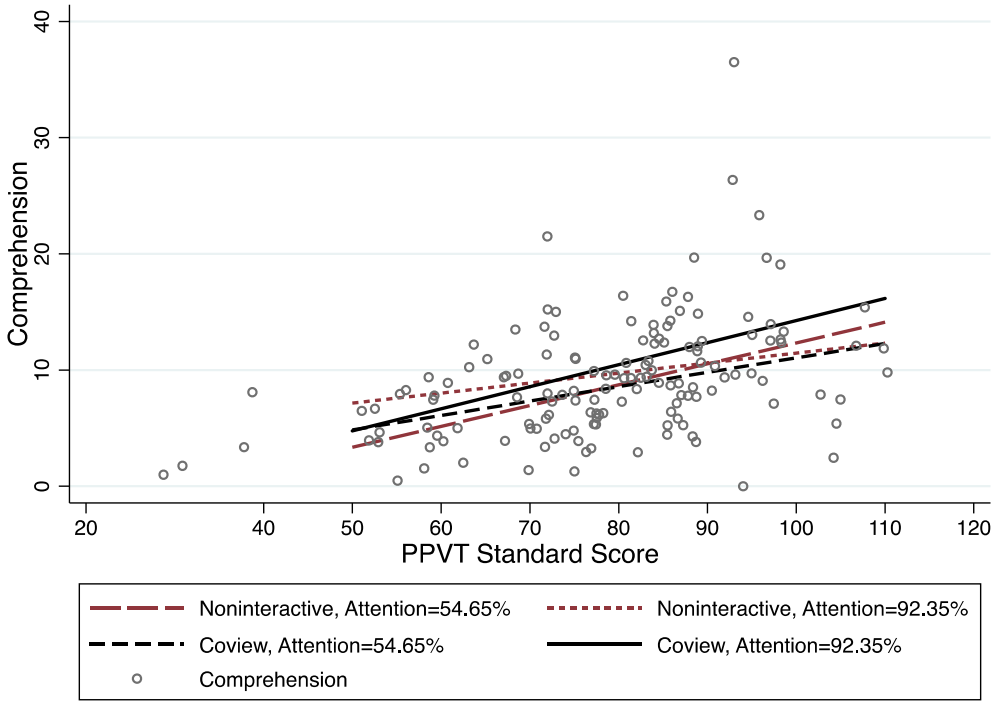

Fig. 2. Simple slopes of comprehension by PPVT scores for the coview and noninteractive conditions when children had very low and very high attention.

Our results align with prior research that has found inconsistent results regarding how coviewing impacts comprehension. Past work found that coviewing often did not benefit comprehension when naturalistic forms of coviewing were used (e.g. Rasmussen et al., 2016; Skouteris & Kelly, 2006). We offered two competing predictions related to attention as to why these forms of coviewing failed to provide added benefit over viewing alone. One prediction was that the coviewer distracts children from the video. Thus, visual attention to the video might be lower, leveling out any potential benefit of the added auditory input. Our alternative hypothesis was that, aligned with the interpretations of Salomon, 1977 and Strouse et al., 2013, coviewing actually enhances visual attention with the program. Our results supported the latter hypothesis- that attention was supported by coviewing. As such, lower attention did not seem like a valid reason for a lack of coviewing influences on comprehension.

However, subsequent analyses revealed that higher attention did not necessarily result in higher comprehension. Prior research has suggested a tenuous association between visual attention and learning (Kirkorian et al., 2017), and our study similarly confirmed that attention did not directly predict comprehension. Similarly, comprehension did not differ based on coviewing condition in our study. As such, even though visual attention was heightened as a result of coviewing, this did not necessarily translate into stronger comprehension.

Ultimately, we found that coviewing predicted stronger comprehension specifically when children had both higher attention to the screen and relatively high baseline vocabularies. Aligned with dual coding theory (Clark & Paivio, 1991; Paivio, 1990), these results suggest that having two sources of simultaneous auditory input (the media and the coviewer) may have been taxing on the auditory processing resources of children, particularly those with weaker vocabularies. When children had stronger baseline vocabularies, they may have been more likely to process and take advantage of both channels of auditory input. This alone was insufficient for coviewing to predict comprehension in our study, however. This visual channel also needed to have high input (reflected in higher visual attention to the video) in order for children with stronger vocabularies to demonstrate stronger comprehension in the coviewing condition.

As such, attention alone was not a missing link between coviewing and comprehension, but rather one of multiple factors that contributed to whether coviewing was more likely to predict comprehension. Children in our sample were from low-income backgrounds and had relatively weak vocabularies overall. Even children in our higher language group had mean vocabulary standard scores that were two-thirds of a standard deviation below the population mean. For these children with average to below-average language skills, neither language nor attention alone was not enough to support comprehension – both were necessary for coviewing to positively predict comprehension.

Situated within the coviewing literature, the present study suggests that some of the inconsistencies in prior work might relate to the role of attention, the style of coviewing, and child vocabulary. Prior work on naturalistic coviewing (e.g. Rasmussen et al., 2016; Skouteris & Kelly, 2006), much like our study, showed no overall comprehension benefit to coviewing. The present study suggests that a possible reason this style of coviewing may not demonstrate comprehension benefits is that the efficacy of coviewing interacts with child and process variables. Only under certain circumstances does coviewing predict stronger comprehension.

Our enactment of coviewing did not incorporate two elements that have been studied in past research – personalized, continent interactions, and questioning techniques. Prior studies on personalized, questioning-intensive coviewing interventions (Reiser et al., 1984; Strouse et al., 2013) have often shown benefits for comprehension. This instruction-heavy approach that incorporates extensive questioning and feedback may consistently improve learning, but is unlikely to be used spontaneously by parents. The present study found that an enactment of coviewing that used non-questioning educational interactions is insufficient to promote comprehension by itself.

In combination with the prior coviewing literature, our study suggests that children with varied background characteristics may be most likely to benefit from a more intensive coviewing approach than the one we studied – an enactment that is both personalized and questioning-focused, such as the dialogic questioning enactment studied by Strouse et al. (2013). Additionally, in order to allow for auditory processing time, pausing the video to have these discussions surrounding content is likely to benefit children. Unfortunately, these conditions are unlikely to reflect the spontaneous coviewing landscape. The enactment in our study was also more likely to reflect interactions used more spontaneously by parents when educationally coviewing – and it seemed potentially beneficial for only a subset of children.

As such, our study suggests that general recommendations to actively coview educational media with children may not always enhance comprehension without additional guidance on exactly how to coview. Even the clearly educational enactment of coviewing in our study did not produce an overall benefit. A less focused enactment that might occur in natural contexts where parents are unfamiliar with techniques to enhance learning may not produce the intended learning benefits to coviewing educational media. Rice et al. (1990) found that viewing educational media with an adult failed to predict learning, though viewing alone did. As such, the present research highlights the importance of providing parents and teachers guidance on the strategies to use while coviewing, as an unfocused enactment of coviewing may not generate the intended return-on-investment.

Overall, the present study extended our understanding of coviewing by investigating how coviewing impacts attention, as well as how attention and baseline vocabulary interact with coviewing to predict comprehension. Nonetheless, our study had some limitations. Our findings may not be generalizable to all coviewing partners (e.g. parents and teachers), since our study utilized researchers unfamiliar to the child. However, using researchers as coviewers allowed for an investigation of a clearly defined intervention with fewer distractions. Additionally, the strategies used in our enactment were quite common in typical parent-child shared book reading (e.g. Evans et al., 2011; Fisch et al., 2002; Ninio & Bruner, 1978). Finally, our results followed similar patterns to prior work studying similar enactments with parent-child dyads, suggesting that our study likely resembled a natural coviewing situation.

A second limitation is the restricted nature of our sample as well as our sample size. We studied only 83 children from low-income backgrounds with relatively weak baseline vocabularies approximately one standard deviation below the norm. This may limit generalizability populations with stronger language skills, and caution should be taken when generalizing findings from one study alone due to the sample size and variations between samples. For our sample, both visual attention and stronger PPVT scores interacted with coviewing to predict comprehension. Fewer or different components of the interaction may be needed for children from more advantaged backgrounds or for children with language skills that are very strong. Nonetheless, some prior work on coviewing (e.g. Strouse et al., 2013) has focused on relatively high-income, educated samples and has found that comprehension was still not directly benefited by coviewing enactments similar to ours. As such, it is likely that processes such as individual differences in attention and language proficiency are relevant to varied samples, though they may not show identical associations as our study.

Another limitation of our study was that we did not assess comprehension of concepts prior to viewing the video. It is therefore possible that some children had greater understanding of one video compared to another. We did randomly assign children to the specific video viewed in each condition. As such, possible variations in prior knowledge is likely to add to random rather than systematic noise in the data.

Finally, our coviewing enactment did not incorporate any questioning techniques, and it was not personalized to match individual children's experiences. Parents may link programs to their child's

personal lives more than our study enactment. However, the primary limitation of parent-child research on coviewing is a lack of clarity on enactment – which is something the present study was able to provide. We were also able to investigate a clearly educationally focused interaction that did not include questioning techniques.

Parents are often recommended to coview media with their children (e.g. American Academy of Pediatrics, 2016). The present study suggests that the benefits of an educational, but non-intensive enactment of coviewing may not be ubiquitous. Within an educational media coviewing context, we suggest that practices such as coviewing need to be considered in combination with the child's language and attention. Coviewing in a less intensive manner may predict comprehension for children with adequate language skills and attention. However, for children with less developed skills, our enactment of coviewing may have produced an overwhelming rather than supportive language environment. As such, an educational but non-intensive form of coviewing may not be a high-yield practice to boost comprehension of educational media for children who are still developing the necessary language or attentional skills to take advantage of it.

Acknowledgements

This work was funded by IES grant R305a150143 awarded to Dr. Susan Neuman.

References

American Academy of Pediatrics (2016). Media and young minds: council on communications and media. *Pediatrics, 138*, 1–8. https://doi.org/10.1542/peds.2016-2591.

Anderson, D. R., Lorch, E. P., Field, D. E., & Sanders, J. (1981). The effects of TV program comprehensibility on preschool children's visual attention to television. *Child Development, 52*, 151–157. https://doi.org/10.2307/1129224.

Anderson, D. R., Lorch, E. P., Smith, R., Bradford, R., & Levin, S. R. (1981). Effects of peer presence on preschool children's television viewing. *Developmental Psychology, 17*, 446–453. https://doi.org/10.1037/0012-1649.17.4.446.

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*, 294–304. https://doi.org/10.1037/a0013844.

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

Dunn, D. M., & Dunn, L. M. (2007). *Peabody picture vocabulary test (PPVT)*. New York: 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. https://doi.org/10.1177/0142723710393795.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. https://doi.org/10.3758/BF03193146.

Fisch, S. M., Shulman, J. S., Akerman, A., & Levin, G. A. (2002). Reading between the pixels: parent-child interaction while reading online storybooks. *Early Education and Development, 13*, 435–451. https://doi.org/10.1207/s15566935eed1304_7.

Ginsborg, J. (2006). The effects of socio-economic status on children's language acquisition and use. In J. Clegg, & J. Ginsborg (Eds.). *Language and social disadvantage: Theory into practice* (pp. 9–27). West Sussex, England: John Wiley & Sons.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review, 87*, 329–354. https://doi.org/10.1037/0033-295X.87.4.329.

Kaefer, T. (2018). The role of topic-related background knowledge in visual attention to illustration and children's word learning during shared book reading. *Journal of Research in Reading, 41*, 582–596. https://doi.org/10.1111/1467-9817.12127.

Kaefer, T., Neuman, S. B., & Pinkham, A. M. (2015). Pre-existing background knowledge influences socioeconomic differences in preschoolers' word learning and comprehension. *Reading Psychology, 36*, 203–231. https://doi.org/10.1080/02702711.2013.843064.

Kaefer, T., Pinkham, A. M., & Neuman, S. B. (2017). Seeking and knowing: attention to illustrations during storybook reading and narrative comprehension in 2-year-olds. *Infant and Child Development,* (5), 1–10. https://doi.org/10.1002/icd.2018.

Keene, J. R., Rasmussen, E. E., Berke, C. K., Densley, R. L., Loof, T., Adams, R. B., ... Marshall, A. (2019). The effect of plot explicit, educational explicit, and implicit inference information and coviewing on children's internal and external cognitive processing. *Journal of Applied Communication Research, 47*, 153–174. https://doi.org/10.1080/00909882.2019.1581367.

Kirkorian, H., Pempek, T., & Choi, K. (2017). The role of online processing in young children's learning from interactive and noninteractive digital media. In R. Barr, & D. Linebarger (Eds.). *Media exposure during infancy and early childhood*Switzerland: Springer. https://doi.org/10.1007/978-3-319-45102-2_5.

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*, 140–151. https://doi.org/10.1016/j.appdev.2013.01.001.

Morgan, T. M., & Case, L. D. (2013). Conservative sample size determination for repeated measures analysis of covariance. *Annals of biometrics & biostatistics, 1*, 1002–1015.

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

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics, 3*, 285–308. https://doi.org/10.1146/annurev-linguistics-011516-034226.

Paivio, A. (1986). Psychological processes in the comprehension of metaphor. *Metophor and Thought, 163*. Retrieved from https://ci.nii.ac.jp/naid/10029347183/.

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

Rasmussen, E. E., Keene, J. R., Berke, C. K., Densley, R. L., & Loof, T. (2017). Explaining parental coviewing: the role of social facilitation and arousal. *Communication Monogrpahs, 84*, 365–384. https://doi.org/10.1080/03637751.2016.1259532.

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*, 443–461. https://doi.org/10.1080/17482798.2016.1203806.

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. https://doi.org/10.1007/BF02768893.

Rice, M. L., Conti-Ramsden, G., & Snow, C. (1990). Preschoolers' QUIL: quick incidental learning of words. Children's. *Language, 7*, 171–195.

Richards, J., & Anderson, D. R. (2004). Attentional inertia in children's extended looking at television. *Advances in Child Development and Behavior, 32*, 163–212. https://doi.org/10.1016/S0065-2407(04)80007-7.

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

Roser, N., & Martinez, M. (1985). Roles adults play in preschoolers' response to literature. *Language Arts, 62*, 85–90.

Salomon, G. (1977). Effects of encouraging Israeli mothers to co-observe sesame street with their five-year-olds. *Child Development, 48*, 1146–1151.

Schmitt, K., & Anderson, D. R. (2002). Television and reality: toddlers' use of visual information from video to guide behavior. *Media Psychology, 4*, 51–76. https://doi.org/10.1207/S1532785XMEP0401_03.

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*, 22–30.

Smith, B., Colunga, E., & Yoshida, H. (2010). Knowledge as process: contextually-cued attention and early word learning. *Cognitive Science, 34*, 1287–1314. https://doi.org/10.1111/j.1551-6709.2010.01130.x.

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. https://doi.org/10.1037/a0032463.

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

Tsai, M., Hou, H., Lai, M., Liu, W., & Yang, F. (2012). Visual attention for solving multiple-choice science problem: an eye-tracking analysis. *Computers & Education, 58*, 375–385. https://doi.org/10.1016/j.compedu.2011.07.012.

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. https://doi.org/10.1016/j.appdev.2004.09.006.

Watkins, B. A., Huston-Stein, A., & Wright, J. C. (1981). Effects of planned television programming. In E. L. Palmer, & A. Dorr (Eds.). *Children and the faces of television, teaching, violence, selling*. New York: Academic Press.