Toddlers' Word and Category Learning from Video:

Links Between Learning and the Screen Mediated Environment

by

Clare Elizabeth Sims

B.A., B.S., University of California Davis, 2006

M.A., University of Colorado Boulder, 2009

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Psychology and Neuroscience

2013

This thesis entitled:

Toddlers' Word and Category Learning from Video:

Links Between Learning and the Screen Mediated Environment

written by Clare Elizabeth Sims

has been approved for the

Department of Psychology and Neuroscience

____

Eliana Colunga (chair)

____

Matt Jones

____

Pui Fong Kan

____

Matt Keller

____

Mike Mozer

____

Yuko Munakata

March 19, 2013

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above-mentioned discipline.

IRB protocol: 1108.20

Sims, Clare Elizabeth (Ph.D., Psychology and Neuroscience)

Toddlers' Word and Category Learning from Video: Links Between Learning and the Screen Mediated Environment

Thesis directed by Eliana Colunga, Associate Professor, Department of Psychology and Neuroscience, University of Colorado Boulder

Screen media, such as television, videos, and computers, are an increasingly common environment for children's learning, even among infants and toddlers. Prior research suggests, however, that very young children learn less effectively from a screen than they do from face-to-face interaction with a person, the so-called video deficit effect. In three studies involving 165 2½- to 3-year-old children I investigated the characteristics and scope of toddlers' word and category learning from video, as well as aspects of the screen mediated environment that are associated with this learning. In the first study, two experiments examined children's word learning, generalization, and retention of novel lexical categories learned from watching a video compared to interacting with a person. While toddlers learned and retained novel, one-to-one word-referent mappings just as well from a screen as they did from a person, they experienced a video deficit in generalizing those words to novel categories. Children trained by video retained less clear, coherent categories after a delay compared to those trained in person. The second study investigated how the addition of either perceptual or social information supports toddlers' learning from a video. The results suggest that being able to interact with the physical objects to be learned about while watching a video may ameliorate the deficit, but learning words directly from a person while watching a video does not help. The third study assessed toddlers' word and category learning in relation to a common, naturalistic context of screen mediated learning:

parent-child co-viewing. The results showed that parents' use of label elicitation questions and positive feedback was positively associated with learning, whereas parental speech that focused on the video and content more broadly showed negative associations with toddlers' learning. Together these studies add to knowledge about the scope and limits of young children's learning from screen media, and how aspects of the environment may support this learning. Implications for explanatory accounts of the video deficit effect, potential uses in practice, and future directions for research are discussed.

Acknowledgments

First and foremost I would like to thank my advisor, Dr. Eliana Colunga. Thank you for seeing potential in me and giving me a chance to pursue my goals. Your mentorship has helped me become a better researcher, writer, teacher, and mentor.

I would also like to thank the faculty, staff, and my fellow students in the Department of Psychology and Neuroscience and the Institute of Cognitive Science. I am proud to have been a part of the supportive, collaborative, and interdisciplinary community of research that we have here at CU. I would especially like to thank all of the faculty members who have helped me as teachers and mentors over the course of my training, including my committee members, Matt Jones, Pui Fong Kan, Matt Keller, Mike Mozer, and Yuko Munakata.

This research would not have been possible without the many children and families from the Boulder area who gave their time to participate in the studies. I would also like to thank the undergraduate research assistants who helped collect and code the data presented here: Tess Amer, Lauren Becker, Katie Double, Allison Langley, Kristina Muntean, Danielle Robinson, Lindsey Sanders, and Kelsey Young.

I would like to thank my friends and members of my Holt, Patrick, and Sims families for your patience and support. Especially my grandfather Dr. Maurice Holt, who first inspired me to get my Ph.D.; my mom, Helen Holt, grandmothers Eileen Holt and Virginia Patrick, and stepmother Rose Tooman for showing me that strong women can accomplish anything they set their minds to; my dad Tom Patrick for unflagging positivity and support; and my sisters Natasha Holtpatrick and Liz Constantz for keeping me sane.

And last but not least, thank you to my husband Brad Sims (and our dog Cali Holtpatrick-Sims!) for your unconditional support, patience, and love. Brad, you’ve cheered me on and kept

me going these past six years, and I cannot imagine having taken on this journey without you by my side.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Children develop within rich and complex environments, learning from many diverse sources. One source of information that is becoming increasingly prevalent, even for young children, is screen media. Screen media includes modalities such as television, videos, computers, and handheld devices like smart phones and tablets. Screen media has been put to use as a source of entertainment and education for children for many years, but recently there has been a dramatic proliferation in content aimed specifically at infants and toddlers. The American Academy of Pediatrics (AAP) has recommended that children under the age of two years not watch any TV or DVDs (AAP, 2010). However, survey data indicate that many families do not heed these recommendations and that screen media are a significant part of the lives of many young children in this country (Rideout & Hamel, 2006). Moreover, much of the content aimed at these young age groups makes claims of offering educational value (Fenstermacher et al., 2010). However, production of infant- and toddler-directed screen media has outpaced empirical research on the educational and social value of that content, and the existing evidence is mixed.

The collection of studies presented here aims to address several of the gaps in the existing literature on young children's learning from screen media. One of these gaps is the dearth of domains that have been studied up to this point. As will be shown shortly, empirical work comparing children's learning from a screen to learning from a person is limited to just a handful of domains and experimental paradigms. The current studies branch out from the literature by investigating children's screen-mediated category learning. This work bridges the gap by including a task that has been used before in this area of research, namely word learning, and then extending to related tasks that capture the related domain of category learning. The current studies also contribute to and extend on the literature by systematically investigating the characteristics of screen-mediated learning. These studies have been designed not only to document whether screen-mediated category learning is better or worse than learning from a person, but to examine how these types of learning differ from each other qualitatively. Further, by manipulating the kinds of perceptual and social information available in the context of learning, these studies explore whether changes to the environment can improve screen-mediated learning. Finally, the studies presented here have implications for applied uses of this research. The final study specifically links characteristics of the context of screen media use (how parents interact with their children while watching) to learning outcomes from the content of that media. Establishing these kinds of links is an important first step in translating the results of this research to recommendations that can be applied in practice.

The next chapter will set the stage for these studies. First I will review the existing literature on young children's learning from screen media, particularly focusing on what is known about how learning from a screen compares to learning from a person. There is consensus that meaningful differences exist, but several explanations for these differences have been proposed and are still in question. Next I will review the research that motivated my approach to studying category learning through the lens of word learning. I will clarify some of the important links between language and cognition and the theoretical support for this position. I will end by reviewing work on parental influences on children's screen mediated learning. This section will cover the existing literature on ways in which parents may mediate their children's learning of screen media content.

Chapter 2: Background Literature

Learning from Screen Media

What Can Children Learn From Screen Media. As will be discussed in more detail shortly, there is strong evidence pointing to a video deficit in young children, in which infants and toddlers learn less well from a screen than they do from face-to-face interactions (Anderson & Pempek, 2005). This deficit has generally been reported to persist until about 2½ to 3 years of age. However, research also shows that children across a broad age range, even infants and toddlers, can learn some things from screens. In this section I will review work on what children actually can learn from screen media, starting with infants and then looking at older age groups, before turning to the literature that focuses on the deficit.

Research shows that even very young infants can learn from screen media to some extent. This has been demonstrated using an imitation paradigm, in which infants watch a video of an experimenter performing several simple actions on a toy and are later given an opportunity to interact with that toy and imitate the modeled actions. Studies have shown that infants ranging from 6 to 24 months of age can imitate actions modeled by video both immediately and following a 24-hour delay (Barr, Muentener, & Garcia, 2007; Meltzoff, 1988). Other results show that infants are able to imitate increasingly complex actions with age (Barr & Hayne, 1999), and by 24 months their imitation from video is robust across changes in context (Strouse & Troseth, 2008). These studies show that even very young infants can learn, remember, and reenact information learned from a screen.

Studies of toddlers' learning from screen media have made use of a few different types of tasks. For example, Troseth and DeLoache (1998) used an object retrieval paradigm to show that 2- and 2½-year-old children can use information learned on a screen to guide their subsequent search for a hidden toy. Another task that has been used with toddlers is word learning, in which children must map a novel or unfamiliar word onto an item in the world and remember that mapping over time. A series of studies presented evidence that two-year-olds can successfully recognize newly learned words trained in a screen presentation, both when tested through a screen and with real objects (Allen & Scofield, 2010; Scofield, Williams, & Behrend, 2007). However, related studies begin to hint at the limitations to screen-mediated word learning. Scofield and Williams (2009) confirmed the results above, but also found that children were unable to infer that an untrained novel label should go with an untrained item presented on screen, an ability that has been demonstrated with in-person tasks among this age group (Liitschwager & Markman, 1994). A similar result was found in a study of verb learning: three-year-olds demonstrated learning of novel verbs trained and tested on screen, but could not map an untrained verb to an untrained action (Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009). These results demonstrate that toddlers can learn some information from screens, but also start to show the limitations in this learning that characterize the video deficit effect.

Children's ability to learn from screen media increases in the preschool years, as children become increasingly competent in the conventions of these media, especially television (Anderson & Hanson, 2010). As children gain a better understanding of the form of screen media, they are better able to process the content. For example, Rice and Woodsmall (1988) taught three- and five-year-olds unfamiliar vocabulary words through a commercial television program, a more naturalistic context for children's screen media learning. Both age groups subsequently recognized pictorial depictions of the vocabulary words, although five-year-olds showed higher accuracy and recognized more items than three-year-olds. Finally, a classic study of screen mediated learning in preschoolers shows how children can learn social information from screens.

Bandura, Ross, and Ross (1963) showed three- to six-year-olds videos depicting aggressive behaviors. In the video, which was shown while children played in a room, an experimenter hit, kicked, and threw a Bobo doll. Children were then put into a frustrating situation, in which they were denied access to some desirable toys, and then led to a room with other toys including a Bobo doll. Children who had watched the video carried out more aggressive behaviors in the testing phase than did children who had not seen any aggressive model. This also shows that children pick up on various sorts of information conveyed on a screen, even if adults may consider it negative or undesirable.

Altogether, research with children ranging from six months of age to preschool-aged show that children can effectively learn and use information from a screen. Children can retain screen learning over both short and somewhat longer delays, and can learn different content such as actions, words, and behaviors. Children can also extend information learned from a screen to new contexts and situations. As is generally also the case in face-to-face learning tasks, as children get older they are able to learn more complex information from screens. The fact that children can learn some information from screens is important for the studies that I will present here. However, although children can effectively learn some things from screens, the video deficit literature demonstrates that this learning differs from in person, face-to-face learning. In general, children are less effective at learning and using new information when it is learned from a screen rather than face-to-face. In the next section I will review the literature that demonstrates this video deficit effect.

The Video Deficit in Infants and Toddlers. Although it has been demonstrated that children can learn, retain, and extend information presented on a screen, studies that compare screen media to face-to-face learning typically find a discrepancy. This section will cover the emerging area of research that explores the video deficit effect in infants and toddlers. First I will give an overview of studies that have used survey and observational methodologies to establish correlations between aspects of screen media exposure and behavioral measures in children. Then I will turn to the literature of experimental studies of the video deficit effect, with a focus on research in the domain of language. This experimental work is better able to address the question of how screen media may actually cause differences in learning. Finally I will discuss two types of explanatory accounts of the video deficit in young children.

Survey and observational studies. One way that researchers study the effects of screen media, especially TV viewing, on children's learning is by collecting surveys and diaries from parents of young children. Children's screen media viewing habits are analyzed in the context of various behavioral measures, such as vocabulary learning or cognitive development. In this way researchers have started to explore how screen media use influences children's development, and what other factors play a role in such relationships. For example, Zimmerman and Christakis (2005) used data from a large-scale longitudinal study to link children's early screen media use to later cognitive outcomes. The authors found that television use before age three was a negative predictor for reading recognition, reading comprehension, and digit span memory tests at ages six and seven. Interestingly, television viewing between the ages of three and five actually had a positive relationship with the later reading recognition measure. Another longitudinal study found that watching *Sesame Street* had a positive correlation to vocabulary development in three- to five-year-olds, but had less benefits for slightly older children (Rice, Huston, Truglio, & Wright, 1990). A large-scale survey of parents found an association between viewing baby videos or DVDs and lower vocabulary scores in 8- to 16-month-olds, but no association for 17- to 24-month-olds (Zimmerman, Christakis, & Meltzoff, 2007). All of these results suggest that early screen media exposure can have a negative effect on subsequent cognitive development, but also that the effects of screen media differ depending on other factors such as age.

The content of what children watch is also important, and studies show that it is correlated with vocabulary size and expressive language development (Linebarger & Walker, 2005; Okuma & Tanimura, 2009). For example, Linebarger and Walker (2005) found that while watching some shows, including *Dora the Explorer*, correlated with larger vocabularies in 30-month-olds, watching other shows such as *Teletubbies* was negatively correlated to the same measure. Okuma and Tanimura (2009) further broke down a sample of children's favorite videos into types and characteristics, and found that children with delayed language abilities favored certain characteristics in videos, such as continually changing images. Altogether, survey and observational studies of screen mediated learning have revealed intriguing patterns. Child characteristics, such as age, come together with characteristics of the media itself, such as how content is presented, and associate with learning. Further, these studies show that screen media interacts with learning at both short- and long-term timescales. These studies have advantages in capturing such patterns across large samples of children and exploring naturalistic screen media use. However this work falls short in its ability to explore causal links between young children's screen media use and learning outcomes. Experimental research is necessary to help complete this picture.

Experimental studies. Experimental studies that manipulate whether information to be learned is presented through screen media are apt to investigate how and when screen media might cause deficits in learning, as well as the underlying causes of the video deficit effect. In this section I will review the body of research that uses this approach. Current research in this area has been conducted using a variety of paradigms, but mostly falls into the domains of imitation, object retrieval, and language. I will review the current state of the research with an emphasis on studies in the domain of language.

Several studies have demonstrated that certain basic behaviors exhibited by infants and toddlers differ in the context of screen media when compared to typical face-to-face interactions. For example, in the first year of life infants tend to show weaker affective reactions to screen mediated stimuli and prefer to look at live presentations when given a choice (Hains & Muir, 1996; Diener, Pierroutsakos, Troseth, & Roberts, 2008). Young children also show a video deficit in self-recognition tasks. Although 24-month-olds can easily succeed at the traditional self-recognition task involving a mirror, toddlers up to 36 months of age struggle when the same type of task is administered using video instead (Suddendorf, Simcock, & Nielsen, 2007; Zelazo, Sommerville, & Nichols, 1999). These studies suggest basic differences between how children perceive and use screen mediated and live presentations. Other work has focused more specifically on differences in learning across these contexts.

A common way to study learning in infants and toddlers is through an imitation paradigm in which an adult models a specific action that a child will be given an opportunity to imitate at some later time. Studies of 6- to 12-month-old infants show that these very young children can imitate simple actions both immediately and after a 24-hour delay no matter if they saw those actions demonstrated on a screen or by an adult in person (Barr et al., 2007; Krcmar, 2010). The video deficit effect in imitation seems to develop over the first year of life, and has been shown in several studies of 12-month-olds (Barr & Hayne, 1999; Klein, Hauf, & Aschersleben, 2006). However, there is some evidence that doubling exposure to video demonstrations can alleviate the deficit in this age group (Barr et al., 2007; Barr, Muentener, Garcia, Fujimoto, & Chavez, 2007). As children approach their second birthday, the video deficit persists under some conditions (Krcmar, 2010; McCall, Parke, & Kavanaugh, 1977), but children seem better able to take advantage of support such as repetition (Barr & Hayne, 1999; Barr et al., 2007). At 24 months of age children can successfully imitate simple actions as well from a screen as from a person (Krcmar, 2010), but still show a deficit in imitation of more complex, multi-step sequences of actions, both immediately and after a delay (Barr & Wyss, 2008; Hayne, Herbert, & Simcock, 2003; Strouse & Troseth, 2008). However by this age, children can use more complex information such as intentionality or end goals of action sequences to support imitation from screen media (Deocampo & Hudson, 2005; McGuigan, Whiten, Flynn, & Horner, 2007; Nielsen, Simcock, & Jenkins, 2008). By 36 months of age, the video deficit decreases for both immediate and delayed imitation (McCall et al., 1977). Altogether, imitation studies suggest a non-linear developmental progression of the video deficit effect from infancy to early childhood, and as children get older they are better able to take advantage of sources of support in the task.

Another task paradigm commonly used to study the video deficit effect, especially with toddlers, is object retrieval. In a typical object retrieval task, a child watches from an observation room as an experimenter hides a toy somewhere in a nearby hiding room. The child is then led to the hiding room and is instructed to find the toy. Researchers have manipulated whether children view the hiding event on screen or in person to study the video deficit effect. Studies show that 24-month-olds are worse at both retrieving a hidden object (Troseth & DeLoache, 1998) and placing an object in a specific place in the hiding room (Schmitt & Anderson, 2002) when information is conveyed over live-feed video compared to directly in person. This deficit persisted even when children watched a hiding event directly and over a screen simultaneously, as well as when they watched both a hiding and retrieving event demonstrated for them on screen (Troseth, 2003). The video deficit among this age group was not alleviated by a 2D version of the task that reduced transfer demands (Schmidt, Crawley-Davis, & Anderson, 2007) or by a goal-based version of the task (Deocampo & Hudson, 2005). However, 24-month-olds have been shown to improve in using screen mediated information to guide their search following a contingent, personalized, and interactive exchange with an experimenter over live-feed video (Troseth, Saylor, & Archer, 2006). Performance also improved in this task when opportunities for perseverative errors in retrieval were reduced (Suddendorf, 2003). In general the video deficit effect in object retrieval decreases with age (Schmitt & Anderson, 2002; Suddendorf, 2003; Zelazo et al., 1999), although it has been observed to persist in 30- to 36-month-olds when task complexity it high (Lauricella, Pempek, Barr, & Calvert, 2011). In sum, the video deficit in object retrieval is robust among 2-year-olds but generally decreases and is mostly gone by the age of 3.

In contrast to imitation and object retrieval paradigms, relatively few studies have directly compared learning from a video to learning from face-to-face interactions in the domain of language. Among the handful of studies that have explored infants' and toddlers' language learning from screen media, the tasks and methods used have been more varied than in the other domains reviewed above. The types of language learning tasks that have been studied range from making phonetic distinctions to learning new words. There are several reasons that studies looking at language learning from screen media are needed. One is that in order to harness technology to facilitate or improve language learning, we must better understand how children learn from screens in the first place. Another reason is that these types of studies can help elucidate the mechanisms that underlie language learning. Social interaction and contingency are thought to be especially important factors that support language learning. The use of screen media is a particularly apt way to manipulate and test accounts of social factors in language learning.

From birth, infants are constantly exposed to language, both from live and mediated sources, so it is reasonable to look at how even very young infants learn from speech from different sources. A developmentally early ability that paves the way for subsequent language acquisition is speech perception. Research shows that at six months of age infants can discriminate phonetic speech contrasts from their native language as well as from a non-native language to which they have never been exposed (e.g., Werker, Gilbert, Humphrey, & Tees, 1981). However, over the course of the first year of life, infants become attuned to the patterns of and distinctions between speech sounds in their native language. While this specialization facilitates learning of the native language, it also leads to an inability to recognize speech contrasts that do not exist in the native language. Kuhl, Tsao, and Liu (2003) tested what kinds of exposure to non-native speech distinctions can prolong 9-month-old infants' ability to perceive those distinctions. These researchers found a video deficit in maintaining phonetic distinctions: infants exposed to non-native speech on video lost their sensitivity to non-native speech distinctions, whereas those exposed to live speech were still able to discriminate between the same distinctions.

Once children are familiar with the patterns of sounds in their language, another major developmental step is learning the words that make up their language. Researchers have looked at how young children learn unfamiliar or novel words from various live and screen mediated sources. In a typical word-learning task, children are shown an object and told a label for that object. The label is usually repeated several times or the child is trained on the label over multiple sessions. To test for learning, the child is presented with an array of objects, including the trained object and distractor objects, and is asked to identify the one corresponding to the trained label. One study found a U-shaped developmental progression of the video deficit in word learning: while children between 13 and 20 months of age showed a significant deficit, younger children (aged 6 to 12 months) and older children (aged 21 to 24 months) did not show a deficit (Krcmar, 2010). Another word learning study found evidence for a video deficit among children between 15 and 24 months of age, which was actually stronger after children were exposed to a commercial video compared to a lab-created video (Krcmar, Grela, & Lin, 2007). While the studies mentioned so far were conducted in a lab setting, another study confirmed the video deficit in word learning among 12- to 18-month-olds exposed to a commercial video in their own homes (DeLoache et al., 2010). Finally, one study using a slightly older age group did not find any evidence for a video deficit in a word learning task among 30-month-olds (O'Doherty et al., 2011).

Altogether, studies in the domain of language learning show an early emerging video deficit effect that seems to diminish with age—by the third year of life, children are at least able to learn words as well from a screen as they can from a live person. In the studies that make up this dissertation, I focus on the upper end of this age range. Although evidence suggests that toddlers are able to learn more effectively from screen media, there are unresolved questions about the scope of this learning. In Study 1 I build on previous research in this area by comparing screen mediated and face-to-face word and category learning. Yet this emerging literature also suggests that children's screen mediated language learning may depend on characteristics of the screen mediated context itself, such as the kind of video used. Such a small sample of studies makes it difficult to draw broad conclusions about the video deficit in language learning, but these studies can still be informative to an account of the general video deficit effect.

Accounts of the Video Deficit Effect

Explanatory accounts of the video deficit effect fall into two main camps: cognitive/perceptual and social. Each of these perspectives focuses on a specific type of factor that differs between direct, in person learning and screen mediated learning. On one hand, cognitive accounts focus on the ways that cognitive processes like attention and perception are deployed in learning depending on the information that is available. On the other hand, social accounts focus on the ways that social contingency and joint attention, for example, influence learning. I will review the main proposed explanatory accounts that fall into each of these two categories in turn.

Cognitive/perceptual accounts. Some researchers have posited that the video deficit is caused by the perceptual differences that exist between screen mediated and live presentations of stimuli and events. The key difference is that live presentations are thought to be much more perceptually rich than those on a screen. Live presentations of stimuli allow the child to see the actual 3D objects and their features as well as spatial relations between objects and their surroundings, both when the items are static and in motion. Video presentations, on the other hand, can only convey an impoverished 2D representation of those objects, diminishing visual cues like texture gradients, motion parallax, and stereopsis (Schmitt & Anderson, 2002). Indeed, some neurophysiological evidence indicates that 18-month-olds process 3D visual information more quickly than 2D (Carver, Meltzoff, & Dawson, 2006).

A couple lines of evidence support this account, but entail slightly different ideas about how the differences between 2D and 3D perceptual information impact learning. One piece of support comes from the finding that repetition of screen mediated information diminishes the video deficit (e.g., Barr et al., 2007). This suggests that repeated exposure to screen mediated information can strengthen encoded representations. In this view, the video deficit is driven by differences in encoding perceptual information between video and live conditions. Children are able to access and recall information learned from a screen, but the quality of that information is what is detrimental to performance. However, other evidence suggests that encoding alone may not drive the deficit. The finding that children's performance suffers following 2D to 3D transfer (Zack, Barr, Gerhardstein, Dickerson, & Meltzoff, 2009) has led some researchers to posit that children have difficulty in understanding the relationship between 2D and 3D presentations (e.g., Barr & Hayne, 1999; Hayne et al., 2003). In this view, infants do not adequately perceive the similarity between screen mediated presentations of items and the real objects presented to them at test to be able to fully grasp how they are supposed to relate. This view shifts the focus from encoding to discrepancies in processing the information that is later retrieved in a new context.

The other main type of cognitive/perceptual account focuses on competing representations. One version of this account posits that children create separate representations of events that they observed on a screen and events in which they were directly involved. Importantly, children are thought to weight these representations differently, with direct experience being prioritized (e.g., Schmidt et al., 2007). These representations compete and direct experience tends to win out, causing the video deficit effect. Evidence for this account comes from studies exploring the patterns of errors that children make over multiple trials in object retrieval tasks. Several studies have demonstrated that children primarily make perseverative errors in the video conditions of this type of task (e.g., Deocampo & Hudson, 2005; Schmitt & Anderson, 2002; Schmidt et al., 2007; Suddendorf, 2003). Children tend to perform above chance on the first trial but quickly decrease in accuracy due to repeating the same behavior they had just carried out on the previous trial. This pattern of behavior supports the view that young children more strongly weight real life experience than that gained from a screen (Schmidt et al., 2007). Building on this, some also posit that the information children get from their experience with real items at test creates proactive interference for subsequent training involving screen media (Suddendorf, 2003). That is, direct experience in the task interferes with later encoding of new screen-mediated information.

Another version of the competing representations account is called the dual representation account. A key idea within this account is that screen mediated information itself presents two distinct representations that children have to deal with. A video of an event, for example, is both a concrete thing in and of itself as well as a symbol of something that has happened before or something that is currently happening elsewhere. In order to effectively use the information presented on a screen, children need to understand the link between the video they see and the actual events that it represents. This account has also been used to explain the development of children's use of other symbolic sources of information, such as photos and models (Troseth & DeLoache, 1998). One piece of evidence for this explanation comes from the interesting finding that sometimes there actually is an advantage for 2D presentations of information over 3D sources. For example, toddlers are better at using pictures to guide their subsequent search behavior than they are at using scale models (DeLoache, 1991). This view posits that a symbolic representation that is very interesting and engaging as an object on its own hinders children in making the link between symbol and reality. Screen media seem to be somewhere in between pictures and models in this regard, and in fact learning from a video can actually be transferred to and improve learning from a model (Troseth & DeLoache, 1998).

While these accounts each offer slightly different explanations for the video deficit effect, they all focus on the perceptual differences that exist between screen mediated and live presentations of information, and how cognitive mechanisms may operate differently on these kinds of information. The next type of explanatory account shifts the focus to social aspects of the screen mediated environment.

Social accounts. One type of social account of the video deficit effect is based on the idea that the deficit is largely driven by the discrepancy in social contingency that typically exists between video and live presentations of information (e.g., Richert, Robb, & Smith, 2011). In live, face-to-face interactions with infants and young children, adults use various tactics to convey information. For example, adults might adjust the timing or style of how they present information in order to keep children engaged in a task and to guide children's attention to what must be learned. Importantly, these kinds of adjustments are contingent on what the child is doing – that is, they are individualized to each child and situation. In contrast, screen mediated presentations completely lack this quality of contingency. While many studies comparing video and live conditions strive to keep the information conveyed to children in both conditions constant, it is difficult to completely rule out the presence and influence of these kinds of social cues in face-to-face interactions (Barr & Hayne, 1999).

A key piece of evidence for this account is the finding that making screen mediated presentations socially contingent can alleviate the video deficit. For example, having children learn information through an interactive, contingent computer game improves performance (Lauricella et al., 2010). Engaging in a socially contingent interaction with an experimenter over live-feed video seems to result in better subsequent performance in a screen-mediated task (Troseth et al., 2006). Even watching a video presentation that involves a socially meaningful actor (such as a child's own other) has been shown to reduce the video deficit effect (Krcmar, 2010). Social contingency may operate not only through immediately contingent interactions, but also through children's expectations about individuals that they have had contingent interactions with in the past. That is, children may prioritize the information they learn from a person with whom they have already established a social routine, or at least with whom they have had some prior experience interacting with.

The other key social account of the video deficit effect focuses on children's prior expectations about screen media. In this view, called the discounting hypothesis, children's previous experiences with screen media, such as TV and videos, have taught them to discount the information they see on screen. Children typically use screen media for entertainment, and so what they see on screen is not related to or informative about the things immediately going on around them. As a result, it is thought that young children struggle when they are asked to use information on screen to guide what they do in a task.

One piece of evidence for the discounting hypothesis is the U-shaped developmental progression of children's learning from videos (Strouse & Troseth, 2008; Troseth et al., 2006). Because very young infants have not yet learned to discount what they see on a screen, they should not show a video deficit. Over the first couple years of life, children are increasingly exposed to screen media for entertainment, and so they learn to discount screen sources over time. Another piece of evidence for this account is the Troseth and DeLoache (1998) finding that presenting screen-mediated information as if it is actually a real event currently taking place alleviates the video deficit. When these authors set up and presented the video as if it were actually a window into the hiding room, disguising the fact that a video was being used, about half of the children performed as well as a live condition. And a perhaps less direct piece of evidence for the discounting hypothesis is the fact that making a task goal-directed alleviates the video deficit (e.g., Deocampo & Hudson, 2003; Klein et al., 2006; Nielsen et al., 2008). Including a goal in a screen mediated task may help children realize that video information can be relevant to their actions and in turn help them override their tendency to discount that information.

In sum, social accounts of the video deficit effect focus on different aspects of what differs between screen mediated and live presentations of information. What these different social accounts have in common is their focus on children's expectations about the information they can get from a screen, or from a person shown on a screen, rather than the quality of that information itself. Children may expect socially contingent interactions from a person shown on a screen, or alternatively, they may expect someone shown on screen not to be a valid source of information. In either case the video deficit effect may be driven by a violation of these expectations: either the person on screen does not interact contingently, or there is information depicted on screen that actually is relevant to the learning task at hand.

It is clear that both social and cognitive processes contribute to screen mediated learning, as they do to any other type of learning. In this way, this dichotomy may not be useful to a unified account of the video deficit effect. However, thinking about social and perceptual factors at the level of the information available in the context of screen media may be helpful for better understanding the deficit and finding ways to help children learn from screens. In Study 2 of this dissertation, I drew on this dichotomy to design manipulations of social and perceptual information in screen mediated learning. This allowed for an investigation of how specific aspects of the screen mediated environment impacted word and category learning. The debate between these types of accounts also connects fundamentally to issues of language learning. In the next section, I will review the background literature on a specific area of language learning: word learning. I will focus on what is known about how children learn words and how word learning links to cognitive development, especially category learning. I will also review how the debate between social and cognitive/perceptual accounts is similar to a debate about learning mechanisms that has manifested in the word learning literature.

Language Learning and Cognitive Development

Acquiring language is a crucial feat of early childhood development. Children make this process look deceivingly simple, producing their first words around 1 year of age and then going through a spike in vocabulary development around 18 months of age (Goldfield & Reznick, 1990). There is a vast literature documenting the progression of word learning and investigating the mechanisms that drive this developmental process; I will focus my review on the aspects of this literature that are relevant to my project. First I will review what is known about two aspects of word learning: how children retain the words that they learn, and how children generalize newly learned words to categories. Retention and generalization of word learning are two key measures that I included in my studies. In the latter section I will focus on the literature that links word learning to category learning, as this is the theoretical position guiding my approach and study design. I will then briefly review the proposed mechanisms that support children's word learning and generalization, with particular attention to parallels between accounts of word learning and accounts of the video deficit effect.

Retention of word learning. As anyone who has spent time with a two-year-old can attest, young children are fast and efficient word learners, seeming to soak up the language around them. Research has established that children can learn and retain new words after brief exposures, a phenomenon called fast mapping (Carey, 1978). Carey observed fast mapping in preschoolers following naturalistic, incidental exposure to a new word. For example, children were asked by their teacher to "get the chromium tray, not the blue tray" in the course of their everyday classroom activities, thus not drawing special attention to the new word that was introduced. In the following weeks, children showed evidence of learning a new color category based on this limited exposure to a new word. This shows that children do not need to be explicitly taught new words but instead can infer the meaning of a word based on the context in which it is used. In this case, children seemed to go through a process of contrasting an unfamiliar word with a known color label to infer that the new word was in fact a new color label. Subsequent research has documented fast mapping in children as young as 17 months of age and across various domains in addition to color (e.g., Dollaghan, 1985; Halberda, 2003; Heibeck and Markman, 1987). Altogether, evidence shows that young children begin making inferences about the meanings of words following just a single exposure.

Once children identify a new word in the linguistic environment, they use the current context to start making inferences about the meaning of that word. How does this process unfold over time and what information do children retain? Researchers posit that fast mapping involves two distinct processes: referent selection and referent retention (e.g., Horst & Samuelson, 2008; Spiegel & Halberda, 2011). Referent selection is the process described above in which children make an initial mapping between a new word and an unfamiliar or unlabeled item. Referent retention is the ability to remember that word-item mapping over time and effectively use that information later on. Spiegel and Halberda (2011) tested 2-year-olds on referent retention following multiple trials of referent selection. Children showed high initial accuracy when given a choice between a familiar and novel item and asked to select one as the referent for a novel word. The authors then tested referent retention by displaying all of the novel items together and asking children to identify one by name. After only a single previous exposure to each word in the referent selection task, 2-year-olds chose the correct item at above chance levels during the retention test. This provides further evidence that young children can retain word-referent mappings learned from brief exposures.

However, other research has called this conclusion into question. For example, Horst and Samuelson (2008) argue that many studies of fast mapping only test learning over minimal delays and provide contextual support for referent retention. To support this argument, these authors conducted a carefully controlled study of referent retention to investigate how 2-year-olds retained newly learned word-referent mappings following a single exposure. The authors included two key manipulations to create a stringent test of retention. Children were tested on word learning after a 5 minute delay, a longer time delay than has been used in most studies of fast mapping. The authors also controlled the contextual cues available at test by familiarizing children with all of the test items ahead of time. In this way, children could not rely on familiarity with items to differentiate trained items from distractors at test. The results showed that although children were accurate at the initial step of referent selection, they showed poor retention of word-referent mappings after a delay. This result suggests that referent retention may not be as robust as previously thought, particularly when children do not have helpful contextual cues available at the time of retrieval.

Other research has linked the process of referent retention in fast mapping to more general principles of learning and memory. For example, Vlach and Sandhofer (2012) examined the role of various memory supports on referent retention among 3-year-olds. Children were provided with either none, one, or multiple common types of memory support when initially learning new word-referent mappings. For example, a novel word might be repeated several times, or the child would be prompted to produce the novel word during referent selection. The results showed that referent retention improved proportionally to the amount of memory support provided at initial learning, even over delays of one week and one month. On the other hand, children showed forgetting curves similar to adults when no memory support was provided. Similarly, Childers and Tomasello (2002) investigated two-year-olds' word learning resulting from either massed or distributed exposure. The results showed that distributed exposure to a novel word over the course of several days resulted in significantly higher accuracy in children's production of that word compared to massed exposure concentrated on one day. Together these studies suggest that children's retention of fast-mapped words adheres to the same principles as those found more generally in the domains of learning and memory.

In sum, studies show that children are fast and effective word learners. Within the first two years of life, children develop the ability to quickly infer new word-referent mappings based on brief, incidental exposures. Children can also retain these mappings, but do so by taking advantage of various supports available to them, such as linguistic context and item familiarity. An important principle of children's retention of newly learned words is that they seem to rely on domain-general learning and memory abilities to support word learning. Although studies of the video deficit effect in word learning suggest that toddlers are able to fast-map words to referents in the context of screen media, it is unknown how robustly children retain these mappings. The studies that make up this dissertation include measures of retention in order to investigate this question. Another novel contribution of the current work is the approach of studying word learning as category learning, a topic I turn to next.

Word learning and categorization. Developmental research has established that from early in life children show a taxonomic constraint in word learning, specifically in the context of nouns (Markman & Hutchinson, 1984). That is, children infer that a noun refers to multiple different items that all belong to the same taxonomic category. What is striking about this finding is the fact that children organize items differently depending on whether or not they are provided with a label, and even depending on what kind of label it is. For example, Markman and Hutchinson (1984) tested for the taxonomic constraint in word learning among 2- to 5-year-olds. Children were shown a picture of a familiar target item and either given a novel label (e.g., "This is a *dax*") or no label (e.g., "Look at this one") for the item. Next, an experimenter presented a pair of test items to choose from: one that belonged to the same taxonomic category as the trained target item, and one that was thematically related to the trained target item. Children were asked to choose one of the test items either by name in the label condition (e.g., "Find another *dax*") or by similarity in the no label condition (e.g., "Find another one that is the same"). Children who had been given labels for the target items were more likely to choose members of the same taxonomic category at test compared to children in the no label condition. This suggests that the presence of labels leads children to override a tendency to attend to thematic relations and instead leads them to think about categories of items.

Since this early finding of a taxonomic bias in word learning, many studies have replicated and further tested the effect. Research has shown that preschool-aged children extend novel labels to other members of the same taxonomic category specifically in the context of nouns and not adjectives (Waxman, Philippe, & Branning, 1999). The same study also demonstrated that this noun-specific taxonomic bias persists over delays in testing of up to one hour. The presence of labels also helps preschoolers classify items into taxonomic categories at the superordinate level (Waxman & Gelman, 1986). This finding held true even among 3-year-olds, who struggled to identify superordinate categories when no labels were provided. While many studies on the taxonomic bias in word learning focus on preschool-aged children, there is evidence for this bias among 2-year-olds (Markman & Hutchinson, 1984; Waxman & Kosowski, 1990) and even younger children. One study found evidence for increased attention to taxonomic categories in the context of nouns compared to no label among 12- to 13-month-old infants (Waxman & Markow, 1995). Another study showed that hearing two distinct labels led 9-month-olds to expect to see two distinct objects, suggesting an early precursor for understanding that words denote different kinds of objects in the world (Xu, 2002). Together these studies show that from early in life children appreciate the link between words, especially nouns, and categories of things in the world.

Why would it be useful to have this connection between language and categories? Categories help us organize our knowledge of the world, and words are powerful representational tools that can help in this process. The facilitative effect of words on categorization continues even into adulthood. Research shows, for example, that the presence of linguistic, pseudoword labels helps adults learn novel image categories better than the presence of non-linguistic cues to category membership (Lupyan, Rakison, & McClelland, 2007). This is true even though the labels themselves do not provide any extra information about the categories to be learned. Rather, Lupyan and colleagues propose that just the presence of linguistic labels acts as a cue for categorization. This account is similar to that in the developmental literature that labels help children cue into taxonomic categories. But what kinds of information do children use when there is no taxonomic information available, that is, when they are learning novel words and categories?

Research shows that children deploy their attention in skilled and targeted ways when they generalize a new word to other novel category members. In the absence of familiar taxonomic categories, children attend to specific features of items that should be predictive of category membership. This pattern of behavior has led researchers to posit broader, feature-based constraints on word learning in addition to the taxonomic bias. One such constraint seen in children's noun learning is the shape bias. The shape bias refers to young children's tendency to generalize newly learned nouns to other objects based on similarity in shape (Jones & Smith, 1993; Landau, Smith, & Jones, 1988). This pattern of word learning has been demonstrated using a novel noun generalization paradigm. In this type of task, a child is taught a novel noun for a novel item (e.g., "This is a *dax*"). Then the child is presented with various test items that each match the trained item in specific features, and is asked to generalize the newly learned noun to the test items (e.g., "Is there a *dax* here?"). The shape bias is evident when a child extends that noun to other objects matching the original in shape, even if that shape match differs from the original in texture, color, or size. Conversely, that child will not extend the noun to objects that match the target in other features but not in shape. The shape bias captures a pattern of behavior in which children generalize newly learned words in targeted ways, with clear limits to the types of objects that they will consider as potential category members.

While the shape bias is relevant to children's word learning in the context of novel solid objects, attentional biases in general are useful for learning new words and categories of various kinds. For example, children show a material bias when given a word learning task involving non-solid substances. That is, children attend to the material that a substance is made of over and above other features like shape or size in the context of learning labels for non-solid substances (e.g., Soja, 1992; Soja, Carey, & Spelke, 1991; Yoshida & Smith, 2005). Other work has investigated how children generalize words to novel items that are presented as animate things, for example by adding eyes to the stimuli or providing verbal cues that indicate animacy (e.g., Booth & Waxman, 2002; Booth, Waxman, & Huang, 2005; Jones & Smith, 2002; Jones, Smith, & Landau, 1991; Yoshida & Smith, 2003). Together these studies show that rather than preferentially attending to one specific feature or another, children selectively attend to multiple features to guide their label generalizations in the context of animate items. Overall, studies of novel noun generalization show that children flexibly and strategically attend to different features of the items they are learning about depending on various cues about the nature of those items. In other words, children are skilled word learners—but how do they get to be that way?

Mechanisms of word and category learning. The evidence strongly indicates that young children deploy attention and infer categories strategically when learning new words. There is debate, however, as to why children show these abilities from quite early in life. On one hand is the argument that children develop these skills due to exposure to the rich structure of the environment around them. On the other hand is the argument that children's word learning skills are based on conceptual knowledge that children have in place, either based on prior knowledge or even naïve theories. I will focus on these two explanatory accounts of children's word and category learning, and link these accounts to those reviewed earlier on the video deficit effect.

The first account of word and category learning is based on the idea that there is rich structure in the world and in the linguistic environment in which children develop. Labels are useful indicators of categories, as reviewed earlier, and according to this account that is true in part because words and categories consistently co-occur in the world. Children become skilled word learners because they are sensitive to these co-occurrences and structured patterns. Children are able to use domain-general learning mechanisms over time to recognize that, for example, nouns tend to label groups of items that are similar to each other in shape. A key characteristic of this account is that it conceptualizes children's skilled word and category learning as a developmental process that arises out of children's natural interactions with the environment.

This perceptual account of word and category learning has been supported by several lines of evidence. For example, two longitudinal studies have shown a link between word learning in general and children's skilled attention to shape. In one study, Gershkoff-Stowe and Smith (2004) longitudinally tested children on their attention to shape in generalizing a novel label. The researchers also collected diaries tracking children's vocabulary growth. The results showed that children's attention to shape increased as the number of nouns in their vocabularies increased. In another study, Smith, Jones, Landau, Gershkoff-Stowe, and Samuelson (2002) intensively trained 17-month-old children on labels for novel shape-based categories. The children exposed to this training not only showed targeted attention to shape earlier than is typically seen, they also showed a dramatic increase in vocabulary size compared to a control group. This suggests that learning to attend to shape helps accelerate children's learning of object names outside of the lab. Together these studies show that as children add more nouns to their growing vocabularies, they show an increasing preference to attend to shape in the context of naming objects. This preference to attend to shape in turn facilitates subsequent word learning. That is, this pattern of skilled attention to a particular feature in generalizing new words to categories arises out of the process of word learning itself.

Another line of evidence for this account comes from a computational modeling approach. Colunga and Smith (2005) pioneered this approach by training a neural network on input structured like an early child vocabulary. That is, the network input contained the same kinds of correlations between perceptual features, such as shape similarity across categories of solid objects, as is seen in the linguistic environment of young children. When given this realistically structured input, the network developed the same kinds of skilled attentional strategies as those observed in young children. This and subsequent studies have used this approach to accurately model the development of children's skilled word and category learning over time and in different populations (e.g., Colunga & Sims, 2012; Sims, Schilling, & Colunga, 2012). This line of research further supports the idea that children use general learning mechanisms to guide their word and category learning. Children are skilled at this kind of learning because they are able to learn reliable and predictive relationships in the world over time. In sum, the key idea of this first account of word and category learning is that children use domain-general learning mechanisms to benefit from the rich structure available in the environment around them.

The second account of word and category learning shifts the focus from the environment in which learning takes place to the knowledge that children bring to the task of learning. According to this account children possess domain-specific prior knowledge about the kinds of things that exist in the world. Children use this knowledge to help guide their attention in word and category learning. For example, children know that 'animal' is a kind of thing in the world and that certain types of features, like non-obvious insides, help to differentiate between members of this kind (e.g., Gelman & Wellman, 1990). A key characteristic that separates this account from the one described above is that children's expectations in word and category learning are not driven solely by information in the environment like perceptual features, but that knowledge about kinds and essences plays a significant role (Gopnik & Nazzi, 2003).

Support for this account comes from studies of novel noun generalization. For example, Booth and Waxman (2002) investigated whether children would shift their attention to different features in targeted ways based on conceptual rather than perceptual cues. These authors provided conceptual information about objects to be learned through short vignettes in which objects were described as either artifacts or animates. The authors found that children deployed attention to different features in a way consistent with how the object was described in the story, over and above perceptual cues indicating what type of object it was. This result suggests that children prioritize conceptual information (for example, knowledge they have about animate things) over perceptual cues available in the context of word and category learning. In another word learning study, toddlers generalized novel words based on the function of novel objects rather than perceptual similarity (Kemler Nelson, Russell, Duke, & Jones, 2000). Together these examples show how even very young children use prior knowledge of kinds to guide them in learning novel words and categories.

These two accounts of word and category learning have some parallels with the two types of proposed accounts of the video deficit effect reviewed earlier. The perceptual account of word learning overlaps in some ways with cognitive and perceptual types of accounts of the video deficit effect. In both cases the focus is on properties of the environment and the context of learning. In word learning this means the focus is on linguistic patterns and perceptual features of objects and categories. In screen mediated learning these types of accounts focus on the quality of the perceptual information conveyed on a screen. On the other side, the conceptual account of word learning shares some similarities with the social types of accounts of the video deficit effect. In both of these accounts, the focus is on knowledge that children already have and bring to bear on the learning task. In one case this relates to children's knowledge about kinds and categories that exist in the world, and in the other this has to do with children's expectations about how to interact with and get information from other people. In both cases the key component driving learning (or inhibiting learning, in the case of the video deficit) is children's prior knowledge coming into the task.

In sum, word and category learning are intricately related areas of cognitive development. When children learn and remember words, they also infer whole categories that words refer to. The studies included in this dissertation were designed to test different aspects of word and category learning, including word-referent mapping, retention, and generalization to categories. This approach of using multiple measures will add to understanding of the scope and characteristics of screen mediated learning in young children. As in accounts of the video deficit effect, there is debate over the mechanisms that support children's word and category learning. The current studies bypass these debates, instead using them as a starting point for thinking about the kinds of information that are available in the context of screen mediated learning. This allows for an investigation of how changes to this context may facilitate learning from a screen. The current work also takes one other approach to explore the relationship between the screen mediated context and young children's learning: investigating parent-child co-viewing of screen media.

Parental Influences

So far I have reviewed literature about how the form of information (i.e., screen mediated or in person) can impact how young children learn. I have also reviewed what is known about how children learn a specific type of content: words, and by extension, categories. In this final background section I will review literature on one common context of children's screen mediated learning: co-viewing with an adult. Survey data indicate that most parents report watching TV with their child either all or most of the time (Rideout & Hamel, 2006). Parental co-viewing can provide a social context for children's screen media use and help scaffold children's learning.

Exploring the attributes and impacts of co-viewing is vital for connecting research on children's screen mediated learning to real world applications.

Several studies of co-viewing have focused specifically on the nature of parent-child interactions in the context of screen media. These studies do not directly assess children's learning of any specific content from screen media, but rather explore how parents and children interact either while watching together or even when there is only background screen media present. For example, Kirkorian, Pempek, Murphy, Schmidt, and Anderson (2009) investigated the impact of a background adult-directed video on parent-child interactions during free play. The authors found consistent differences in both the quantity and quality of parent-child interactions across 1-, 2-, and 3-year-old children. Parental verbal interactions decreased when the video was on compared to when it was off. Further, when parents did interact with children when the video was on, these interactions tended to be less active and responsive. Overall, these results show that the presence of background screen media aimed at adults can be detrimental to parent-child interactions, particularly through influences on parental behavior.

Using a similar paradigm, Courage, Murphy, Goulding, and Setliff (2010) observed parents and infants during periods of free play both with and without an infant-directed video playing in the background. The authors found that while 6- and 18-month-old infants were primarily interested in playing with toys, they frequently shifted their attention to the screen for short intervals when the video was on. Similarly, parents were much more focused on their infants than on the video, but also vocalized less often and initiated play for shorter durations when the video was on. These results show that while the presence of screen media in the background may not completely draw infants' attention away from other activities, it is disruptive on several measures of both child and parent behavior. These studies also begin to suggest that the content of screen media plays a role in how it impacts parent-child interactions.

What about the screen media that parents and children are actively attending to? Nathanson and Rasmussen (2011) compared parent-child interactions across three contexts: toy play, book reading, and TV viewing. The authors found that among both toddlers and preschoolers, both the frequency and responsiveness of parental speech was lowest during the TV viewing period compared to the other two activities. This result suggests that co-viewing screen media fosters less rich and contingent interactions compared to other common social activities for young children. Does all screen media lead to decreases in the quantity and quality of parent-child interactions? As suggested by the studies of background screen media, content may make a difference. For example, Stoneman and Brody (1982) observed naturalistic, in-home interactions as parents co-viewed sitcoms and *Sesame Street* with their preschool-aged children. The authors found that parents used both programs as a teaching opportunity, but for different kinds of content. Parents talked about and asked comprehension questions about educational content while watching *Sesame Street*, but focused more on providing information about narrative and character motivations while watching a sitcom. This result suggests that, although parental interactions differ depending on the content of screen media, active co-viewing seems to involve parents explicitly teaching information no matter the content.

Another study compared co-viewing interactions using different screen media content, but this time involving different types of infant-directed media. Pempek, Demers, Hanson, Kirkorian, and Anderson (2011) investigated this question by having different groups of parent-infant dyads watch one of two videos at home over a two week period and then come into the lab. The videos were both aimed at infants, but one of them incorporated explicit modeling of quality parent-child play behaviors. Parents and their 12- or 18-month-old infants were observed both in free play and while co-viewing the videos they had been familiarized with at home. The results showed that, consistent with other evidence, parent-child interactions decreased during co-viewing compared to free play for both of the video exposure groups. However, parents who had co-viewed the video that modeled quality interactions more at home showed more frequent and better quality interactions with their infants during free play in the lab. In sum, these results further confirm the general finding that parent-child interactions suffer during co-viewing, but also suggest that high-quality content can have positive effects for parent-child interactions beyond the context of screen media viewing.

If screen media can be used as an effective model of parent-child interactions in everyday play activities, perhaps it can be used to promote better quality co-viewing interactions as well. Fisch and colleagues (2008) investigated this question in preschoolers. In this study, parent-child dyads co-viewed a child-directed video in one of three conditions in which the authors manipulated parent-directed subtitles appearing on the screen. Parents either saw subtitles containing jokes and general parenting tips, subtitles containing prompts and hints for content-related comments they could make while watching, or no subtitles. The subtitles related to the content included suggestions of questions parents could ask to help children actively think about what was going on in the video. The authors found that the presence of content-related prompts led to greater interactions between parents and children. These interactions were also more closely related to the content of the videos. Together the results show that targeted on-screen information can help improve the quantity and quality of co-viewing interactions.

All of the studies reviewed so far focus on how screen media, whether it is the focus of attention or not, influences parent-child interactions. The results show consistent decrements in measures of interaction in the context of screen media, especially compared to situations when there is no screen media present. However, a couple of studies also suggest that screen media content can have an effective and positive impact on parent-child interactions both while co-viewing and in free play. The next question I turn to is how the nature of parent-child co-viewing interactions are associated with children's learning from screen media. That is, even if we know that better quality of co-viewing interactions can be elicited from parents and children, does this improvement in quality have any link to an improvement in learning? Researchers have manipulated and quantified the type and quality of co-viewing interaction to investigate it's impact on outcome measures in children.

Co-Viewing and Learning. Several studies have experimentally tested the impact of different types of co-viewing interactions on children's subsequent learning. For example, Collins, Sobol, and Westby (1981) manipulated the kind of commentary provided by adults while co-viewing a narrative TV show with second-grade children. An experimenter either provided explanatory comments about implicit information in the show (e.g., character motivations or inferences about the story) or neutral commentary. The key result showed that explanatory comments during co-viewing led to better comprehension, even on segments of the show that were not commented on, compared to children who heard neutral comments. In other words, targeted adult commentary facilitated children's comprehension of implicit narrative information that they otherwise would have missed. Another set of studies tested the impact of co-viewing on learning using content that is more commonly watched by younger children: *Sesame Street* episodes. In one study, Reiser, Tessmer, and Phelps (1984) manipulated whether or not adults asked content-specific questions and provided feedback and encouragement while watching segments teaching letters and numbers with preschool-aged children. Post-test results showed that children learned the content of the video more effectively in the experimental condition compared to the control condition, in which adults did not provide any commentary. To further explore this effect, Reiser, Williamson, and Suzuki (1988) tested several additional co-viewing conditions, including questions only, questions with feedback, and one in which adults simply directed children's attention to the screen when educational content was being shown. The results showed that asking questions resulted in better learning among children, whether or not feedback was provided. Together these studies show that adult commentary and questions during co-viewing can directly facilitate children's understanding of and learning from real screen media content.

Fewer studies have linked co-viewing interactions to outcome measures among infants and toddlers. These studies mostly look at certain qualities of parental behavior or parent-child interactions and link them to child behavior in the context of co-viewing. For example, Demers, Hanson, Kirkorian, Pempek, and Anderson (2012) investigated the contingencies between parent and infant gaze directed at a screen. In the co-viewing context, children ranging from 12 to 21 months of age looked more often to the screen following their parent's looks to the screen. Children also looked longer at the screen when they were following a parent's gaze. This study shows that even parent's looking behavior can modulate children's attention to screen media. Another study investigated the quality of parent-infant interactions during co-viewing and linked this to infant behavioral measures (Fidler, Zack, & Barr, 2010). In this study, the authors coded co-viewing interactions for shared focus and turn taking behaviors, qualities that indicate sensitive and reciprocal interactions between infants and parents. The authors also measured parental verbalizations during co-viewing. Results showed that the interaction quality measures predicted 6- to 18-month-olds' looking time to an infant-directed video, even controlling for parental verbalizations. A related study approached this topic slightly differently by classifying parental interaction styles based on co-viewing behaviors (Barr, Zack, Garcia, & Muentener, 2008). The authors found that parents could be classified into different levels of scaffolding, and that these clusters of co-viewing behaviors predicted looking time and responsiveness to infant-directed videos among 12-, 15-, and 18-month-old infants. The parents classified as providing a high level of scaffolding tended to use verbalizations that oriented their children to the video and focused on the content therein. Together these studies indicate several effective strategies for getting infants and toddlers to attend and respond to screen media. Parents that used eye gaze toward the screen, high-quality, responsive interactions, and content-focused verbalizations were most effective in establishing joint attention to the screen and getting their children actively involved in co-viewing.

One study so far has started to link observations of co-viewing to measures of learning specifically among this younger age group. In this study, parents co-viewed a video with their 12- to 25-month-old children that was specifically intended to teach words (Fender, Richert, Robb, & Wartella, 2010). The authors observed parental co-viewing behaviors as well as child verbalizations, and, importantly, also asked parents which of the words in the video their child was unfamiliar with. The authors observed how often children produced words that they had been unfamiliar with prior to seeing the video and used this as their measure of learning. Parents tended to cluster into different groups depending on how much their co-viewing behavior was focused on teaching the words in the video to their children. It was found that children produced more words that they were previously unfamiliar with when their parents had a higher teaching focus during co-viewing. Further analyses showed that these parents tended to focus specifically on the words that they knew their children were unfamiliar with. This result is particularly interesting in light of the literature on the video deficit effect in word learning. This study shows that sensitive parental scaffolding during co-viewing with infants and toddlers may be able to ameliorate the video deficit effect.

The literature on adult-child co-viewing suggest that parents can effectively scaffold children's attention to and learning from a screen. In the final study included in this dissertation, I used a quasi-experimental approach to investigate how parental speech during co-viewing is related to children's word and category learning and retention. Drawing from prior studies of co-viewing with infants and toddlers, the current study involved observation of naturalistic interactions between parents and children while watching a video. However the current work builds on prior research by analyzing more fine-grained characteristics of parental speech during co-viewing and by including more stringent, controlled tests of learning. The literature on co-viewing with young children will benefit from future studies that use a variety of approaches, both observational and experimental, working to understand what kinds of strategies are effective and why in the context of learning from screen media.

The Current Studies

The studies presented here investigate how children learn words and categories from video compared to in person presentations. This work directly adds to the literature on young children's screen mediated learning by measuring word learning and retention, as well as the related cognitive domain of category learning. No other studies so far have investigated screen mediated word learning from this perspective. Within this framework, there are three main goals of the current studies. First, to establish the scope and characteristics of screen mediated word and category learning when directly compared to in person learning in the same task. Second, to explore the impact of perceptual and social information available in the screen mediated learning environment. Third, to investigate whether and how co-viewing screen media with a parent is related to children's word and category learning.

To achieve the first goal, Studies 1a and 1b are comprised of experiments that directly compare children's learning, retention, and generalization of words and categories from screen mediated and in person presentations. These studies demonstrate that although children can learn word-object mappings just as well from a screen as from a person, there is a video deficit related to the categories that children infer and retain over time. Study 2 was designed to address the second goal stated above; this study further investigates the effect observed in Studies 1a and 1b by exploring how the addition of perceptual or social information impacts learning from a screen. Study 2 reveals that different perceptual and social manipulations have different effects on screen-mediated learning, but ultimately these factors must be used together to support learning. Study 3 addresses the third and final goal of the current work by relating parental speech during co-viewing to children's learning from a video. This final study shows that certain things parents do while co-viewing a video are indeed associated with their children's learning from that video, in both positive and negative ways.

Chapter 3: Study 1a

The goal of the first study was to explore the scope and characteristics of children's learning from screen media compared to learning from a person. Rather than simply confirming that learning from one source is better than learning from another, I wanted to better characterize the differences in these types of learning. To this end, in this study I made two key extensions to prior work in the video deficit literature. First, this study made use of an experimental word learning task that tested more than basic, one-to-one word-referent mappings. In Study 1a, children were also tested on generalization and retention. These measures have not been previously tested in a study comparing learning words and categories from a screen to learning from a person. The measure of generalization in particular goes beyond word learning to capture children's category learning. As will be described shortly, testing generalization gives a measure of what children infer about whole categories of items when they are taught a name for a single exemplar. By testing generalization at two different times, I explored the nature of children's learned categories immediately and after a delay.

Second, the sample for this study consisted of a slightly older age group than has previously been examined in the majority of video deficit studies, and especially those related to language and word learning. Previous work has suggested that the video deficit effect diminishes by about two years of age (but see Lauricella et al., 2011, for evidence that the deficit persists for more difficult tasks). One reason for choosing the age group of 2½- to 3-year-olds was to explore the characteristics of screen-mediated learning once children should be beyond the point of the deficit. Another reason for using this age group was because of the nature of the task: by this age, children show consistent patterns of generalization in word learning. Because of this, there is sufficient evidence in the literature to make specific predictions about how children should generalize newly learned words when trained directly in person. In sum, the use of new measures and a new age group in Study 1a represents a new perspective in investigating screen mediated language and category learning.

Predictions

If the video deficit effect is gone by two years of age, then there should be no difference in word learning performance between children trained in person and those trained through a video. Prior work using a similar word-referent mapping paradigm has suggested that children no longer show a deficit in this ability over the age of two years (O'Doherty et al., 2011). On the other hand, if the deficit has not completely disappeared then differences between groups of children trained in person and by video should emerge. If this is the case, a deficit effect should be most apparent in the more difficult tasks of the current study: generalization and retention. Both of these tasks are more demanding than one-to-one word object mapping. Children have been shown to maintain a video deficit on harder versions of a task, even if they do not show the effect on an easier version. This has been shown, for example, in the object retrieval paradigm (Lauricella et al., 2011). If the video deficit is still present to some extent in the third year of life, it should manifest when children are given relatively challenging tests of learning.

From this prediction comes the question of what a video deficit would look like in measures of retention and generalization. The answer is fairly straightforward for retention: although some decrement in retention is expected no matter how children are initially trained, the children who are trained by video should show relatively poorer performance following the delay compared to children trained in person. In terms of the word-referent mapping task, children trained by video should become less accurate in their identification of trained objects over the delay. Children trained in person should show relatively less of a decrement in accuracy. A video deficit effect in the retention of word-referent mappings could indicate that screen mediated learning disrupts long-term encoding of information, or that representations formed from screen mediated information are more susceptible to interference or forgetting over time.

To predict how a video deficit effect may impact generalization, it is informative to first consider what “correct” performance on a novel noun generalization task should look like. As reviewed in the introduction, by about 24 months of age children show a robust preference to attend to shape when generalizing names for solid objects, even if other feature similarities are present (Landau et al., 1988). Over time children learn that shape is a particularly useful feature to pay attention to when they have to decide whether a new object they encounter belongs to the same category as an object they have already learned. That is, children’s inferences about category membership are largely guided by similarity in shape in the context of solid objects. Therefore, I expect that the typical toddler would also show this pattern in a laboratory test of generalization. If the current task gets at the same underlying concepts as previous work on novel noun generalization, my results should replicate previous findings of a preference to learn shape-based categories in the in person learning group.

If toddlers experience a video deficit effect in inferring and remembering words for categories of objects, I should observe less robust generalizations among children trained by video compared to those trained in person. Behaviorally, this may manifest in different ways. Children trained by video may still show a preference for extending words to shape-based categories, but this preference may be less robust compared to the group of children trained in person. This might suggest that children still use a typical approach to word learning even in the context of screen media, but are perhaps less confident or more conservative in how far they can generalize those words. Children trained by video may alternatively show a reversal of the expected pattern, generalizing words more based on other features relative to shape. This might suggest that screen mediated presentations act to highlight other kinds of features relatively more than they highlight the shape of objects. Finally, another possible behavioral pattern that may be seen among children trained by video is overgeneralization. This would be seen if, relative to children trained in person, those trained by video generalize words equivalently to various different objects regardless of feature similarity. This might suggest that the typical, shape-biased approach to word learning is disrupted by screen media. Perhaps screen mediated presentations of information do not support the robust, detailed representations of objects needed to discern which features matter and which do not in a generalization task. The exact way in which generalization of information learned on a screen differs from learning in person will provide insight into the characteristics of screen mediated learning.

Method

Participants. Twenty-nine children (M_{age} = 33.4 mo., *SD* = 1.8 mo., 14 girls) were recruited for participation from the Boulder, CO area. Of these children, 23 (M_{age} = 33.4 mo., *SD* = 1.8 mo., 11 girls) completed both sessions of the experiment and are included in the analyses presented here.

Materials. Children were taught six novel words for six novel objects in the experimental task (see Figure 1a). There were six object sets each consisting of four items: a target object and three other objects that matched the target in one specific feature, but differed in other features. These three feature matching objects consisted of one shape match, one color match, and one texture match object. Thus each object set was made up of an exemplar and three potential category members (see Figure 1b). Each target object was labeled with one of the following novel words: *elg*, *ife*, *nork*, *gub*, *zeb*, and *lug*. The novel words were chosen to be simple, single-syllable words that would be phonologically allowable in English.

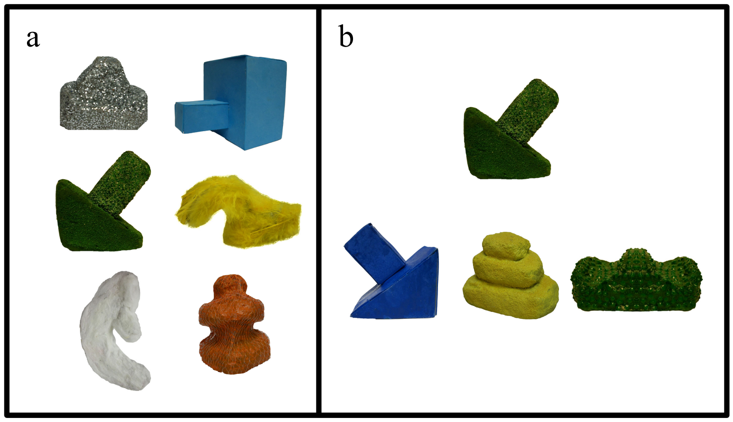

Figure 1. a) Novel target objects created for Study 1a. b) An example of one of the object sets consisting of a target object and, from left to right, a shape match, a texture match, and a color match.

The video condition on Study 1a involved video clips of the experimenter presenting and labeling each novel object in turn. The video clips were filmed in the same lab setting that all children were tested in, and were set up to capture the training events from the child's perspective. The experimenter shown in the video was always the same experimenter presenting the tasks at the first session.

Additionally, parents were given a brief survey on their children's screen media habits (see Appendix A).

Design. Children completed two sessions in the lab approximately one week apart. Children were randomly assigned to one of two between-subject training conditions: in person or video. These two conditions represent how the novel words and objects were initially presented to children.

Procedure.

Practice trial. The first experimental session began with a practice item and question to familiarize children with the generalization task. The experimenter presented a tennis ball, either in person or on the video, and told the child “This is my ball. Do you see my ball? That’s my ball.” Then the experimenter would either put the tennis ball away or pause the video on a blank screen, and retrieve a tray of four items. The tray included a golf ball, a colorful rubber ball, a plastic clip, and a wheel-shaped toy. Children were asked to find a ball, and were praised for answering correctly. If a child chose one of the other items, the experimenter corrected them by saying “Nope, that’s not a ball, so that stays on the tray.” The practice trial was meant to demonstrate to children that not all items presented at test were necessarily members of the same category as the original exemplars. Once the practice trial was completed children proceeded to training.

Training. Children were trained on the novel words and novel objects by an experimenter either in person or through a video. In both conditions, the objects were presented two times each. At each presentation of an object, the experimenter provided three instances of the corresponding novel label (e.g., “This is my *elg*. See my *elg*? That’s my *elg*.”). Therefore, each object was labeled six times across training.

In the in person condition, the experimenter sat across a table from the child and caregiver. The experimenter retrieved an object from behind a shelf and placed it on the table in front of the child. The experimenter rotated the object as they labeled it. The child was then given a moment to look at or explore the object in front of them. The experimenter placed the object out of view before repeating this process with the next training item.

In the video condition, the procedure was similar except for instead of presenting objects, the experimenter showed the child a video. A laptop was placed in front of the child at a comfortable viewing distance. The video consisted of clips of the experimenter retrieving an object, placing it on the table, and rotating it as they said the object label. A blank interstimulus screen appeared for 3 seconds between each clip. As the video played, the experimenter present in the room sat quietly and did not interact with the child. Once training was complete, all children moved on immediately to testing.

Session 1 testing. All children were tested in person to allow for direct comparisons of the measures across groups.

Generalization: free choice task. The first testing task that children completed was a free choice generalization test. Children were presented with a tray of four items: a target object and an object that matched the target in shape, an object that matched in color, and an object that matched in texture. The child was asked to choose an item by trained label (e.g., “Can you show me an *elg* here?”). After the child chose an object, the experimenter took it and placed it out of sight. The experimenter then asked whether there was another object of the same kind on the tray (e.g., “Is there another *elg* here?”). This continued until the child answered no or had chosen all of the objects. The experimenter recorded the order of children’s object choices. Children completed six trials of the free choice generalization test.

Learning: target identification. Next, children were given a two-alternative forced choice target identification test. On each of six trials, children were presented with pairs of items: a target exemplar object and a distractor (a shape match from a different object set). Children were asked to identify the target item (e.g., “Which one of these is an *elg*?”). The experimenter recorded responses as correct or incorrect. The target identification test concluded the first session of the experiment.

Session 2 testing. When children returned to the lab for their second session they were not retrained on any of the items. The second session began with the "ball" practice question, followed immediately by the free choice generalization test. Finally, children completed two types of forced choice tasks. One of the tasks was the same target identification test as was administered at the first session. Because the free choice generalization and target identification tasks were administered at both visits, both tasks offer a measure of retention. The other task was a two-alternative forced choice test of generalization.

Generalization: forced choice task. Children were presented with a shape match from one set and a target item from another set, and asked for the shape match by name. The experimenter recorded whether children selected the correct shape match. The forced choice generalization task was included for two main reasons. First, it offers another way to get at children's category learning. It is a more constrained test of whether children extend a name learned for an exemplar object to another object matching that exemplar in shape. Second, it provides a direct test of the prediction that children tend to extend object names by similarities in shape. If children consistently choose the shape match in the forced choice generalization task, then that supports the prediction that preference for shape can be used as a measure of category learning in the free choice generalization test.

Results

Age and gender were analyzed with respect to the dependent measures reported below, and no significant effects were found. Therefore, these variables are not included in the following analyses.

Learning: target identification. The first analyses investigated learning and retention of simple word-to-referent mappings. Did children correctly learn the novel labels given for the exemplar objects? Children's proportions of correct choices in the forced choice target identification task were first compared to chance. Across groups, children mapped the labels to the correct objects as shown at training, both at initial testing ($M = .67$, $SD = .19$, $t(21) = 4.06$, $p < .001$, Cohen's $d = 0.87$) and at delayed testing ($M = .68$, $SD = .26$, $t(21) = 3.26$, $p < .01$, $d = 0.69$). A paired t-test comparing testing times showed that performance did not differ across visits ($t < 1$, $p > .05$). The next question was whether performance on this task differed depending on training condition. Children who originally learned the words in person did not perform significantly differently than those who learned from video, either at immediate or delayed testing ($t < 1$ and $p > .05$ for both two-tailed t-tests).

Upon inspection of the target identification data, I found that there were three subjects who did not effectively learn the words at their first visit, as evidenced by accuracy performance below 50%. I excluded these subjects, two from the video training condition and one from the in person training condition, and ran the above analyses again. Overall accuracy increased slightly both at initial testing ($M = .72$, $SD = .15$) and at delayed testing ($M = .71$, $SD = .25$), and remained above chance at both times ($t(18) > 3.50$, $p < .01$, and $d \geq 0.85$ for both comparisons to chance). Performance did not differ between visits, and there were still no differences between the training condition groups at either visit ($t < 1$ and $p > .05$ for all comparisons). Because these results remained stable, regardless of whether children who did not learn the words were included, I decided to exclude these three children from all subsequent analyses.

In sum, the results of the forced choice target identification task show that the majority of children accurately learned and remembered the novel word-object mappings introduced at initial training. Importantly, there were no differences on this task between children trained in person and children trained by video. In other words, these 30- to 36-month-olds did not show the video deficit effect in the context of learning a single word for a single object. What about generalization?

Generalization analyses. The next set of analyses explored the results of the generalization tests, first in the forced choice task and then in the free choice task.

Forced choice task. The dependent measure in the forced choice generalization task was children's proportions of correct shape match choices. Across groups, children showed that they could correctly extend a novel word to an object that matched the shape of the trained exemplar of that word ($M = .68$, $SD = .23$, $t(21) = 3.39$, $p < .01$, $d = 0.78$ when compared to chance). Children originally trained in person did not differ significantly from those trained by video ($t < 1$, $p > .05$). However, comparing each group separately to chance, only those children who learned the words in person chose the correct shape match significantly above chance ($M = .68$, $SD = .20$, $t(9) = 2.91$, $p = .02$, $d = 0.92$); children who learned by video only performed marginally above chance ($M = .67$, $SD = .26$, $t(8) = 1.90$, $p = .09$, $d = 0.63$). Although not a conclusive result, these comparisons to chance start to hint at some differences in generalization between video and in person learning. Yet overall, these analyses show that in a constrained test of generalization, children extended the trained novel labels to objects that match the trained exemplars in shape.

Free choice task. The next set of analyses concern the free choice generalization task administered both immediately after training and after a week-long delay. To get a broad picture of how children behaved in this task, first I looked at how often children chose any of the test objects presented during free choice generalization. With four items presented per set, children had the opportunity to extend the novel words to up to 24 objects at each testing time. This measure of total number of objects chosen was submitted to a 2 (Training Condition: in person or video) × 2 (Visit: first or second) mixed models analysis of variance (ANOVA). Although the total number of choices were numerically different between conditions, with children trained by video tending to choose more objects on average (M = 23.42, SD = 1.24) than those trained in person (M = 21.05, SD = 4.90), this effect did not reach significance ($F(1, 18) = 2.26$, $p = .15$). The effect of visit and interaction between condition and visit likewise did not reach significance. This analysis shows that overall children tended to choose many of the objects presented in the free choice generalization task. This suggests that overall number of choices is not sensitive enough to detect possible differences between training conditions or across generalization test items.

To better explore generalization patterns, I used a measure that has been shown to be sensitive to young children's representations of categories: sequential choices (Sugarman, 1982). This dependent measure was based on the order in which children selected different types of objects at test. For each type of object within a set (target, shape match, texture match, and color match), the order of children's choices were recorded, or an item was scored as zero if it was not chosen. Order was then reverse coded to give greater weight to initial relative to later choices. The item chosen first was given a score of 3, second choice was 2, third choice was 1, and fourth choice or not chosen was 0. Finally, for each type of item (e.g., all shape matches across sets) an average weighted choice score was calculated across all object sets for each child. Therefore, each child ended up with four values ranging between 0 and 3 expressing their relative preference for choosing target objects, shape matches, texture matches, and color matches, respectively.

My first prediction was about what "typical" generalization performance should look like, based on previous studies of novel noun generalization. First I looked at the group of children trained on novel words in person, in order to see whether these children showed a preference for generalizing words based on similarities in certain features more than others. Average weighted choices in the in person training condition were submitted to a 4 (Object Match Type: target, shape, texture, or color) × 2 (Visit) repeated measures ANOVA. Children generalized novel words to the different generalization test objects to different extents, as shown by a main effect of match type ($F(3, 27) = 7.62$, $p < .001$, partial $\eta^2 = .46$). No other effects were significant.

As can be seen in Table 1, children trained in person showed a preference for selecting the target objects (matching in all features) earliest. The next most preferred type of test item were those that matched the trained targets in shape, followed by texture and finally color matches. It is not surprising that children tended to first extend novel words to the target objects; this result further confirms that children indeed learned the novel word-object mappings. However, shape match objects were the next most commonly preferred type of generalization test item. Of note, children in the in person group tended to extend novel words to shape matches earlier than they did to either texture matches ($t(19) = 1.76$, $p = .09$, $d = 0.63$) or to color matches ($t(19) = 2.84$, $p = .01$, $d = 0.64$). This confirms the predicted pattern that children trained in person preferred to extend novel names to objects similar to a target in shape over and above extensions to objects with similarities in other kinds of features.

Table 1. Average weighted choices for children in the in person training condition for each of the object match types presented during the free choice generalization task.

Target	Shape Match	Texture Match	Color Match
2.09 (0.58)	1.55 (0.57)	1.16 (0.60)	0.93 (0.61)

Note. Numbers in parentheses are standard deviations.

The next question is whether this pattern of preferences is different depending on how children were initially trained. First, average weighted choices were submitted to a 2 (Training Condition) × 4 (Object Match Type) × 2 (Visit) mixed models ANOVA. Across training conditions and visits, children preferentially extended words differently depending on feature, as shown by a main effect of match type ($F(3, 54) = 12.81$, $p < .001$, $\eta_p^2 = .42$). However, this pattern was qualified by a significant three-way interaction between training condition, match type, and visit ($F(3, 54) = 5.35$, $p < .01$, $\eta_p^2 = .23$). No other effects were significant. This interaction suggests that children's preferences for different kinds of feature matches did differ depending on training condition, but also changed over time.

To better understand the nature of the differences indicated in the three-way interaction, I separated the data by visit and conducted two additional 2 (Training Condition) × 4 (Object Match Type) ANOVAs. The analysis of immediate testing yielded only a main effect of object match type ($F(3, 54) = 10.45$, $p < .001$, $\eta_p^2 = .37$). As can be seen in the left half of Figure 2, at immediate testing children in both training conditions tended to preferentially extend novel names earliest to target objects. This was followed by shape match objects and a lesser preference for objects that matched the trained target in texture and in color.

The same analysis at delayed testing again yielded a main effect of object match type ($F(3, 54) = 8.20$, $p < .001$, $\eta_p^2 = .31$). On average across both conditions, children tended to choose the target item first most often ($M = 1.85$, $SD = 0.58$), further indicating that they effectively learned the novel word-object pairings. Shape match objects were preferred next most often ($M = 1.69$, $SD = 0.44$), followed by texture matches ($M = 1.22$, $SD = 0.51$), and finally color matches ($M = 1.13$, $SD = 0.57$). However, children's generalization patterns over items also depended on how they were trained, as shown by an interaction between match type and training condition ($F(3, 54) = 4.78$, $p < .01$, $\eta_p^2 = .21$). As can be seen in the right half of Figure 2, and confirmed in a one-way ANOVA of object match type, the pattern of choice preferences seen in the main effect is maintained in the group of children who learned the words initially in person ($F(3, 27) = 12.15$, $p < .001$, $\eta_p^2 = .57$). However, this pattern is disrupted among children who learned the words via video presentation. The same one-way ANOVA run on the children who learned words by video did not yield a significant effect of object match type ($F < 2$, $p > .05$). That is, by the second visit children in the video training group did not show any distinction in their preference among the different types of feature match objects at test.

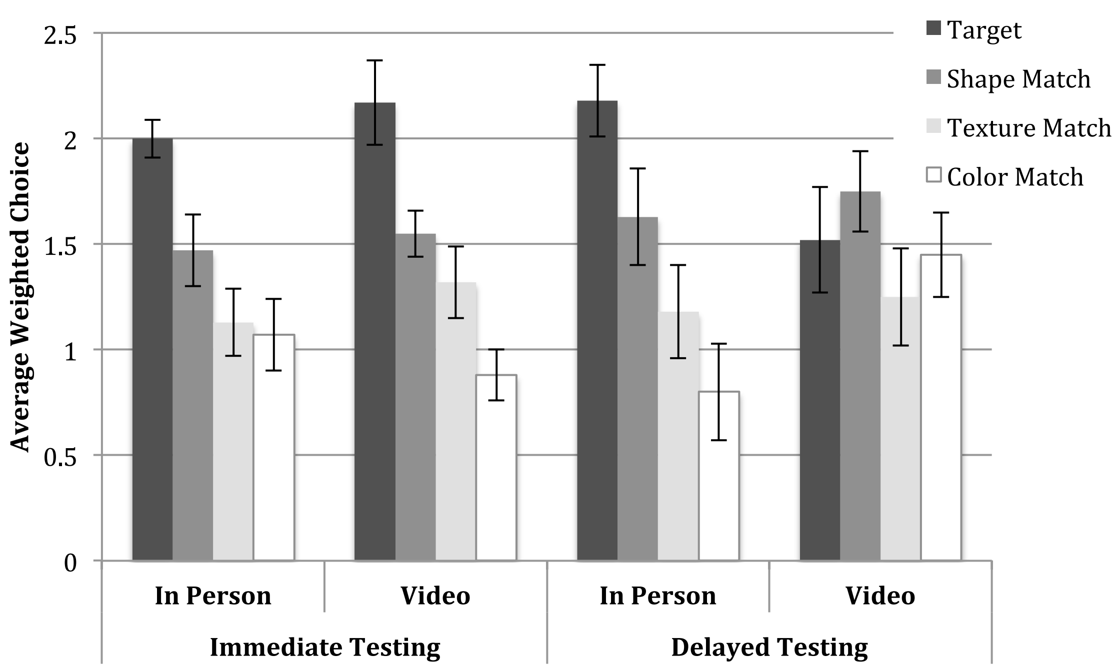

Figure 2. Children's average weighted choice scores at immediate (Visit 1) and delayed testing (Visit 2) separated by training condition and object match type.

Discussion

Children across both groups learned the correct word-object mappings and retained that information over a week-long delay. Children's initial accuracy in these mappings confirms the prediction, and is consistent with prior findings, that there is no longer a video deficit in simple word-referent mapping for children over 30 months of age. However, children's accuracy after a week-long delay, particularly in the video training group, does not fit with my predictions. Instead, this result indicates that the information that children encode about objects presented and labeled in a video is robust enough for children to recognize the correct word-object pairings a week later. Although information learned from a video proved sufficient for word-object mapping, generalization was another story.

Children's performance in the forced choice generalization task started to hint at differences between the in person and video training groups. Although the video group did not differ significantly from the in person group, children trained by video tended to choose the correct shape matches less often, and their performance did not differ from chance. At the same time, this task represents a very constrained test of generalization. In order to succeed at this task, children had to recognize the general shape of an earlier learned target object when it was asked for by name. As shown in the target identification task results, children accurately recognized the target objects themselves at the second visit, so perhaps the forced choice generalization task further reflects that learning. The design of this task also builds in an assumption that shape is the key feature by which children generalize.

A more targeted test of generalization was administered in the free choice task by letting children choose among multiple objects that matched the target in various different ways. The results of this task showed a more striking difference between the in person and video groups.

Children trained in person consistently chose target objects earlier, follow by shape match objects, and chose texture and then color match objects later on. In contrast, children trained by video no longer differentiated the target objects from other potential category members in the same set by their second visit to the lab. This failure to differentiate between object match types at test, and especially between target objects and other types of feature match objects, shows that there was some degradation in the categories that children trained by video retained over time.

What is the best way to characterize this difference between groups at the second visit? Overall, children's generalization choices tended to be more uniform across objects in the video group than in the in person group. Children in the video group did not differentiate target objects in the free choice task, but their other results indicate that they did not forget the trained target items altogether. This hints that the difference is not necessarily in children's retention of the trained target objects, but perhaps in their retention of the inferred categories that those target items represent. In the next study I set out to clarify this difference in generalization through several changes to the experimental design.

Chapter 4: Study 1b

Rationale

Study 1b was designed to further explore the video deficit effect in toddlers' retained categories found in Study 1a. Several changes were implemented to address issues from the first study. One issue with Study 1a was the presence of trained target objects in the free choice generalization task. Children's preference to extend novel names to these objects earlier than they did to any other type of test object could have been due either to their learning of the trained target objects or to a preference for the only familiar object in the array. To control for a confounding effect of familiarity, the free choice generalization task in Study 1b consisted of novel test objects only. Additionally, a new, more structured measure of generalization was added in Study 1b. This measure comes from a forced choice generalization task in which an object matching a trained target in shape was paired with an object matching the same target in material. This constrained test of feature preference should help further test predictions about how children extend novel words. Together, the changes implemented in Study 1b introduce stronger experimental controls and help further focus on toddlers' patterns of novel word generalization. The results will shed more light on the characteristics of toddler's screen mediated word and category learning and retention.

Predictions

Because children in Study 1a accurately learned and retained one-to-one word-referent mappings, there should not be a video deficit in this measure of retention in Study 1b. Consistent with the first study, all children in Study 1b should accurately learn and remember the trained pairings between novel objects and novel words.

If children continue to experience a video deficit effect in certain aspects of word learning, particularly categorization, then similar generalization differences as observed in Study 1a should be observed between the in person and video training groups in Study 1b. The group of children trained on the novel words and objects in person should show a preference to generalize by shape in the free choice and forced choice tasks. In contrast, the group of children trained by video should use shape less consistently as a basis for generalization. These children may still show some preference for shape in the forced choice generalization tasks. These tasks are closely related in content, structure, and time of administration to the forced choice target identification task. In these simple and constrained tasks children may be able to leverage their successful target identification performance to accurately identify objects that match those trained target exemplars in shape. However, the free choice generalization task is expected to capture more subtle differences in children's inferred category structures and thus is expected to show differences between training groups.

Method

Participants. Thirty-seven children (M_{age} = 33.9 mo., *SD* = 1.9 mo., 15 girls) were recruited for participation from the Boulder, CO area. Of these children, 4 were not able to return for a second visit within the required time window, and 1 was fussy or non-compliant at the second visit. Thus, 32 children (M_{age} = 33.8 mo., *SD* = 1.8 mo., 14 girls) are included in the analyses presented here.

Materials. As in Study 1a, children were taught six novel words for six novel objects. The words remained the same (*elg*, *ife*, *nork*, *gub*, *zeb*, and *lug*), but new novel objects were created for Study 1b (see Figure 3a). Each of the six new sets of novel objects consisted of five items: a target object, two objects matching the target in shape, and two objects matching the target in material (see Figure 3b). These objects were designed to focus more directly on shape-based and material-based categories.

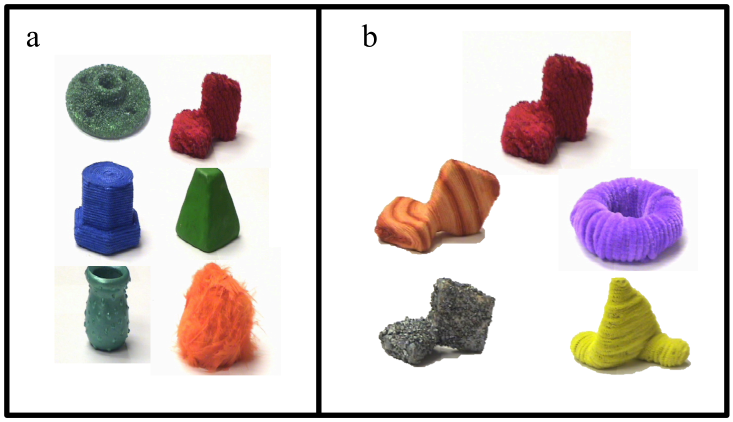

Figure 3. a) Novel target objects created for Study 1b. b) An example of one of the object sets consisting of a target object at the top, two shape matches on the left, and two material matches on the right.

Study 1b also included a more extensive survey as well as a standardized vocabulary measure. The survey included sections on screen media and print media to get an idea of the frequency and nature of children's uses of various media (see Appendix B for questions and results of the screen media survey). The vocabulary measure was the MacArthur-Bates Communicative Development Inventory (CDI-III; Fenson et al., 2007), a standardized checklist of 100 vocabulary words that parents completed to indicate which words their child knows.

Procedure.

Practice Trial. The same practice trial described in Study 1a was used as an initial warm-up for all children in Study 1b.

Training. As in Study 1a, children in the current study were trained on the novel words and novel objects by an experimenter either in person or through a video. Each object was presented two times each, and labeled six times total. At each presentation of an object, the experimenter provided the label in three different ways (e.g., “This is an *elg*! Do you see my *elg*? There’s the *elg*!”). The use of each word in these three kinds of sentences was meant to establish the novel word as a count noun.

Training in both conditions proceeded as in Study 1a. For children in the video presentation condition, the experimenter depicted in the video always matched the experimenter present to administer the testing tasks. Once training was complete, all children moved on immediately to testing.

Session 1 Testing. As in Study 1a, all children were tested in person.

Generalization: free choice task. First children completed an updated free choice generalization task. Children were presented with a tray of four items from a set: two objects that matched the target of that set in shape (but differed in other features) and two objects that matched the target of that set in material (but different in other features; see the lower four objects in Figure 3b for an example). Note that, unlike in Study 1a, the target object was not present during free choice generalization. Therefore, all objects presented on the tray in Study 1b were equally unfamiliar to the child. Free choice generalization proceeded otherwise as it did in the first study. Children completed six trials, with order of choices recorded as the measure.

Learning: target identification. Next, children were given the forced choice target identification task. To better control for familiarity, children were presented with pairs of trained target items, and asked to identify one of the objects (e.g., “Which one of these is an *elg*?”; see Figure 4a). Children completed six trials, with responses recorded as either correct or incorrect. This concluded the first session of the experiment.

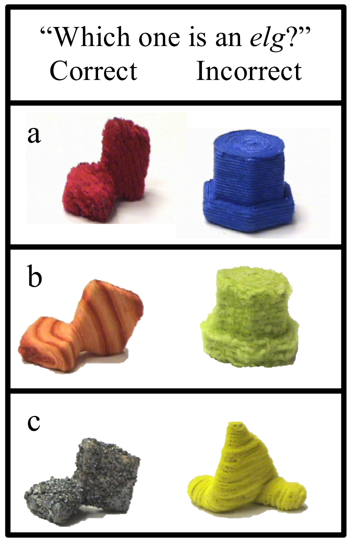

Figure 4. Examples of the forced choice tasks implemented in Study 1b. a) The target identification task paired two of the trained target items. b) The shape vs. shape task paired shape match items from two different sets. c) The shape vs. material task paired one shape and one material match item from the same set.

Session 2 Testing. When children returned to the lab for their second session they were not retrained on any of the items. Children began with the "ball" practice trial and then were given the same free choice generalization and forced choice target identification tasks as at their first visit. Children also completed two additional forced choice tasks; the order of these two final tasks was counterbalanced across subjects.

Generalization: forced choice shape vs. shape task. One of the additional forced choice tasks was similar to that used in Study 1a, and was meant to be a constrained test of generalization specifically to shape matches. Children were presented with pairs of shape match objects from two different item sets, and asked to choose one by a trained name (see Figure 4b). This shape vs. shape forced choice generalization task gives another measure of whether children will extend a trained label to a shape match.

Generalization: forced choice shape vs. material task. The other additional forced choice task was also meant to be a constrained measure of generalization, and paired shape and material match items from the same set at test (see Figure 4c). Children were presented with these pairs and asked to choose one object by trained name. This shape vs. material forced choice generalization task provides another way to establish whether children tend to show a shape or material preference in each condition.

Parents completed the survey and vocabulary checklist at the second session as well, either while their child completed testing or at the end of the session. At the end of participation, parents were fully debriefed on the purpose of the study.

Results

Age and gender were analyzed with respect to the dependent measures reported below, and no significant effects were found. Therefore, these variables are not included in the following analyses.

Learning: target identification. As in Study 1a, the first analysis focused on whether children in Study 1b accurately learned and retained the one-to-one word-object mappings taught to them at training. Children's proportions of correct target choices were first compared to chance. On average across training conditions children accurately identified the trained target objects both at immediate testing ($M = .57$, $SD = .18$, $t(31) = 2.14$, $p = .04$, $d = 0.38$) and at delayed testing ($M = .63$, $SD = .18$, $t(31) = 3.94$, $p < .001$, $d = 0.70$). There was no difference in accuracy between immediate and delayed testing ($t < 1$, $p > .05$). As in Study 1a, children's target identification accuracy did not differ depending on whether they were trained on words in person either at immediate or at delayed testing ($t < 2$, $p > .05$ for both comparisons).

Inspection of the target identification data revealed that five subjects did not effectively learn the novel words, as shown by immediate testing accuracy below 50%. I excluded these subjects, two from the in person training condition and three from the video training condition, and ran the above analyses again. Compared to the full dataset, overall accuracy increased slightly at immediate testing (M = .62, SD = .13), stayed about the same at delayed testing (M = .62, SD = .17), and remained above chance at both visits ($t(26) > 3.50$, $p < .01$, and $d > 0.65$ for both). Accuracy across visits remained equivalent as in the full dataset ($t < 1$, $p > .05$). Also consistent with the full dataset, there were no differences in accuracy between training conditions at either visit ($t < 2$, $p > .05$ for both comparisons).

The results of Study 1b replicate those found in Study 1a in that toddlers correctly identified and remembered trained objects no matter whether children learned about the objects from a person or from a video. Also consistent with the first study is the finding that about 15% of the sample did not effectively learn these word-object mappings. Removing these subjects did not alter the results, and so this constrained dataset will be used for all subsequent analyses. The results of the target identification task also demonstrate that the effects found in Study 1a were not specific to the novel object stimuli created for that experiment. Study 1b included a different set of novel objects and yielded the same results. Having again found no evidence for a video deficit in children's target identification performance, the next question is how children generalize those newly learned words.

Generalization analyses. The following analyses focus on children's generalization performance, first in the two forced choice tasks newly implemented in Study 1b, and then in the free choice generalization task.

Forced choice shape vs. shape task. For the first forced choice generalization task, the dependent measure was children's proportions of correct shape match choices when presented with two shape match objects from two different sets and asked to pick one by a trained label (see Method section and Figure 4b). As this task was only administered at delayed testing, there are no analyses comparing visits. On average across training conditions children accurately chose the correct shape match object above chance levels ($M = .57$, $SD = .17$, $t(26) = 2.28$, $p = .03$, $d = 0.44$). In a direct comparison between training condition groups there was no difference in accuracy between children trained in person and children trained by video ($t < 1$, $p > .05$). Comparing each group of children separately to chance showed that those trained in person performed at chance levels ($M = .56$, $SD = .20$, $t < 2$, $p > .05$), but those trained from a video accurately identified shape matches above chance ($M = .59$, $SD = .13$, $t(12) = 2.30$, $p = .03$, $d = 0.69$).

Together these results suggest that children correctly remember and identify objects that match the trained targets in shape. Comparing each training condition separately to chance hints at subtle differences in the average performance of each group. This is similar to the results of the forced choice generalization task in Study 1a, but in this case appeared in the opposite direction: children trained from the video performed above chance, but those trained in person did not. This change may be due to the changes in the task itself: in Study 1b, familiarity with the shape match test items is better controlled by presenting two test items that had been seen an equal number of times before. The results of this task may hint at differences in children's retention of learned categories, but they do not represent a very direct measure of generalization. The results of the remaining generalization tasks speak more directly to this issue. By comparing children's choices of and preferences for shape matches relative to material matches, the remaining results can shed light on the question of whether screen mediated word learning changes the structure of the categories that children learn.

Forced choice shape vs. material task. The other forced choice generalization task paired shape match and material match objects from the same set against each other at test (see Method section and Figure 4c). The dependent measure was children's proportions of shape match choices when presented with these kinds of pairs and asked to pick one by a trained label. As with the previously reported task, this task was only administered at the second visit. Across both training conditions children significantly preferred the shape match objects ($M = .72$, $SD = .17$, $t(26) = 6.68$, $p < .001$, $d = 1.28$). This preference did not differ between children trained in person and children trained by video ($t < 1$, $p > .05$), and both groups preferred shape matches significantly above chance (in person: $M = .71$, $SD = .18$, $t(13) = 4.50$, $p < .001$, $d = 1.20$; video: $M = .73$, $SD = .17$, $t(12) = 4.78$, $p < .001$, $d = 1.33$).

In this constrained test of shape preference administered one week after initial learning, all children exhibited a strong preference for objects that matched the originally learned targets in shape. This preference for shape is consistent with patterns demonstrated in the literature and with the results of Study 1a. Therefore, the results of this task establish that children show an expected preference for shape over material in a constrained test of word learning. This is true even after a delay and no matter how children were originally trained. The final set of analyses investigated whether children used this preference for shape to extend newly learned words to coherent categories of objects.

Free choice generalization task. The final set of analyses explore the results of the free choice generalization task administered immediately after training and after a week-long delay. This task is meant to evaluate the kinds of categories that children infer from being trained on individual exemplars of novel words. The free choice task measures this by giving children the opportunity to extend those words to various items that they had not seen before, each of which matched the target in one specific feature. First, I analyzed this data with respect to the same measure as that used in Study 1a: averaged weighted choices of each type of test item, based on the order in which each object was chosen in the task. Weighted choices of the two shape matches in each set were averaged, as were the choices of the two material matches in each set. Then, all shape match values and all material match values, respectively, were averaged across sets. Therefore, each child ended up with a value for weighted shape match choices and weighted material match choices for each testing time.

In analyzing children's average weighted choices of the two types of generalization test objects used in Study 1b, I found that this measure did not capture meaningful differences between the two training conditions. For example, average weighted choices were submitted to a 2 (Training Condition: in person or video) × 2 (Feature: shape or material) × 2 (Visit: first or second) mixed models ANOVA. The results showed that all children chose shape match test items consistently earlier than material match items, as shown by a main effect of feature ($F(1, 25) = 66.98$, $p < .001$, $\eta_p^2 = .73$). There was also a marginal trend suggesting that all children chose more test objects overall at delayed testing relative to immediate testing ($F(1, 25) = 3.20$, $p = .09$, $\eta_p^2 = .11$). No other effects were significant. As can be seen in Table 2, the data from Study 1b showed a numeric trend somewhat similar to the pattern seen in Study 1a. At immediate testing, children in both training conditions strongly differentiated between different types of test objects in their novel word generalizations, and particularly preferred shape over material matches. At delayed testing, children trained in person maintained this strong preference but children trained by video showed a somewhat weaker shape preference and relatively less discrimination between object types. However, this measure did not reveal any significant differences in performance between children in the two training conditions.

Table 2. Average weighted choice data from the free choice generalization task. Children's average weighted choices of shape and material match objects are separated by visit and by training condition.

	Immediate Testing (Visit 1)		Delayed Testing (Visit 2)	
	In Person	Video	In Person	Video
Shape Matches	1.88 (0.29)	1.88 (0.29)	1.95 (0.33)	1.87 (0.28)
Material Matches	0.89 (0.43)	0.97 (0.43)	0.88 (0.52)	1.08 (0.36)

Note. Numbers in parentheses are standard deviations.

To get another perspective on how children performed in the free choice generalization task, I next calculated a dependent measure that captured the extent to which children generalized words to shape and material match objects at test. To get a measure of shape match choices, for each child I calculated a proportion from the number of times that child extended a word to a shape match object relative to the total number of shape match objects that were available to choose across all sets. I calculated an analogous value for children's material match choices. Therefore the dependent variable was the proportion of times that children chose objects in the free choice generalization task, regardless of the order of their choices, with each child having a value for shape match objects and a value for material match objects for each testing time.

Using this proportion of choice measure, I first tested the prediction that children trained on novel words in person should demonstrate a preference to extend those words based on similarities in shape over and above similarities in material. Proportions of object choices of children in the in person training condition were submitted to a 2 (Feature: shape or material) × 2 (Visit: first or second) repeated measures ANOVA. Children who had originally learned the words in person showed a consistent preference for extending those words to objects matching in shape (M = .96, SD = .08) compared to objects matching in material (M = .66, SD = .37), as shown by a main effect of feature ($F(1, 13) = 12.74$, $p < .01$, $\eta_p^2 = .50$). There was no main effect of or interaction with visit, showing that this preference was evident immediately and persistent over the delay between testing sessions. This again confirms the typical pattern of a preference to use shape as the basis to generalize newly learned words, and further demonstrates that children trained in person learned and remembered coherently shape-based categories.

Next I explored whether children trained on words in the video condition showed a similar pattern of generalization performance as the in person group, or showed a loss in feature discrimination in their retained categories as was observed in Study 1a. First, proportions of object choices from all children were submitted to a 2 (Training Condition: in person or by video) × 2 (Feature) × 2 (Visit) mixed models ANOVA. A significant main effect of feature confirmed children's overall preference to generalize by similarities in shape ($F(1, 25) = 18.50$, $p < .001$, $\eta_p^2 = .43$). The effect of feature was qualified by two marginal interactions: a two-way interaction between feature and condition ($F(1, 25) = 3.17$, $p = .09$, $\eta_p^2 = .11$) and a three-way interaction between feature, condition, and visit ($F(1, 25) = 3.00$, $p = .096$, $\eta_p^2 = .11$). Although marginal, these interactions are consistent with the findings of Study 1a that children's preference to extend words by shape more than material both depended on how those words were originally learned and changed over time.

To further explore children's generalization performance in the different training conditions and across time, I conducted two additional 2 (Condition) × 2 (Feature) ANOVAs, one at each visit. Using the immediate testing data only, this analysis revealed a main effect of feature ($F(1, 25) = 17.59$, $p < .001$, $\eta_p^2 = .41$). When tested immediately, all children extended the novel words to more of the shape match objects ($M = .95$, $SD = .08$) compared to the material match objects ($M = .74$, $SD = .32$). No other effects were significant. The same analysis conducted using only the delayed testing data also yielded a main effect of feature (*F*(1, 25) = 13.81, *p* = .001, .36). There was also a main effect of condition (*F*(1, 25) = 5.27, *p* = .03, η_p^2 = .17) such that children who had learned the words from a video tended to choose more objects presented at test overall more often (*M* = .96, *SD* = .13) compared to children trained in person (*M* = .81, *SD* = .33). This pattern of choosing more of both types of object matches was further qualified by a significant interaction between condition and feature (*F*(1, 25) = 4.64, *p* = .04, η_p^2 = .16).

As can be seen in Figure 5, at delayed free choice testing children in both training conditions consistently generalized the learned words to all of the shape match objects presented at test. However, only the children who were trained on the words in person maintained the pattern of coherently categorizing by shape significantly more than material (*t*(13) = 3.27, *p* < .01, *d* = 0.87). While children who were trained on the words from the video do consistently choose the shape match objects at test, they select the material match objects just as often (*t* < 2, *p* > .05). In other words, children in the video training condition do not retain the same coherent, shape based categories that are seen among children who originally learned the words in person.

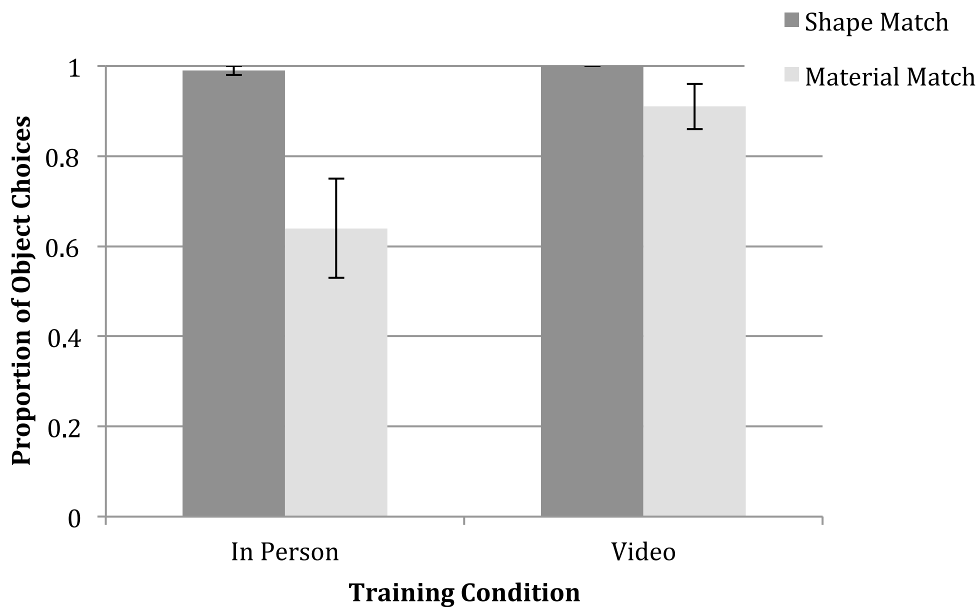

Figure 5. Children's proportions of shape match compared to material match object choices at delayed free choice generalization testing. Error bars represent standard errors.

Discussion

Study 1b both confirms and adds to the findings of Study 1a. The results of the target identification task in Study 1b confirm that children can learn and remember one-to-one word-object mappings using a different set of novel object stimuli. This learning was equally accurate no matter whether children learned the information from a person or from a video. This adds to evidence from Study 1a and from the literature showing that children have overcome the video deficit in word learning by 2½ years of age, at least in the context of forced choice target identification measures. The forced choice generalization measures in Study 1b, both administered at delayed testing only, provide an additional measure to those used in Study 1a and tell a similar story. In general, children can accurately extend a novel word to an object matching the originally learned target in shape, and prefer to extend words based on shape rather than material similarities within a constrained, two-alternative forced choice task. As in the target identification task, there were no differences in these results comparing children trained in person to those trained by the video. Altogether, the results of the forced choice tasks show that children accurately learn, remember, and retain shape similarity preferences for novel word-object pairings learned either from a person or from a video.

The free choice generalization results reveal differences that are not captured by the forced choice measures. These results, which explore the types of categories that children infer and retain based on learning word-object pairs, also further confirm the patterns seen in Study 1a. At immediate testing, all children showed evidence of inferring coherent, shape based categories. This was seen in children's tendency to extend the newly learned words to generalization test objects that matched the original target in shape much more often than they extended words to test objects that matched the original in material. This coherent pattern of generalization based on shape over and above material was also observed at delayed testing, but only among children who had learned the words in person. Children who had originally learned the words via video presentation no longer showed evidence of coherent categories at delayed testing; instead, they extended words indiscriminately to test objects matching the original target object in either shape or material.

These observed differences in category coherence can arguably be seen as another form of the video deficit effect. The ability to infer categories based on learning the label for an item in the world is an important aspect of language development. A key part of this type of inference is that it has limits—children do not indiscriminately extend the labels that they learn. Instead, as reviewed in the introduction, children seem to be biased word learners. They attend to specific features in the context of word learning and generalization, depending on the nature of the item that they are learning about. As has been shown in the case of the shape bias for solid objects, this kind of constrained attention is both an outcome of early word learning and contributes to further vocabulary growth (Gershkoff-Stowe & Smith, 2004; Smith et al., 2002). In other words, as children learn words they also learn about the structure of different kinds of categories in the world. Once these patterns are learned, children can more effectively attend to the features that are important for different kinds of items, thus accelerating further word learning and language development.

The ability to deploy attention selectively to certain features is clearly at work in the children in the current studies. The forced choice generalization task results show that children remember and prefer to extend words to objects that match the originally learned target in shape. The immediate free choice generalization task results build on this by showing that children not only prefer shape, but consistently use it as the basis for inferring coherent categories of named objects organized by their similarity in a single feature. Yet the process of learning new words and categories through screen media somehow changes the way that children attend to object features over time. The object representations that children retain following screen mediated learning seem to be enough to support a shape preference in the context of simple, two-option choices. However, these representations are not sufficient to guide clear, coherent category choices in the more complex, and perhaps more common context of identifying categories when faced with multiple possible choices.

An interesting point to emerge from the results of both studies 1a and 1b is how different kinds of measures capture different aspect of word learning. It is especially striking that certain measures reflect proficiency while others suggest a deficit. For example, all children can effectively learn and retain one-to-one novel word-object mappings, but there is evidence for a video deficit effect when it comes to retained category coherence. This highlights the fact that word learning involves various underlying cognitive skills, and different measures are needed to capture these different abilities. This also indicates that the video deficit effect does not have a unitary, global effect on children's performance, even within specific cognitive domains.

Together, Studies 1a and 1b establish that toddlers still experience some extent of a video deficit effect, specifically when generalizing novel words to novel categories after a delay. Next I investigated whether it is possible to ameliorate children's category retention following screen mediated word learning. Study 2 draws from explanatory accounts of the video deficit effect to test how specific aspects of the screen mediated learning context impact performance in the current tasks. In the next study, I added either extra perceptual or social information to the screen mediated context and explored how these changes impacted children's word and category learning and retention.

Chapter 5: Study 2

Studies 1a and 1b established and confirmed that 2½- to 3-year-old children learn and retain concepts differently from a video presentation compared to an in person presentation. Children who were taught novel words for novel objects in person consistently inferred that those words referred to coherent, shape-based categories, both immediately following training and after a delay. In contrast, children who learned the original exemplars by video did not retain coherent categories over time. Rather than inferring shape-based categories, as has been seen before in generalization measures in the literature and in the in person group, children in the video group overgeneralized and chose all objects that matched the trained exemplar in any way. This can arguably be seen as a video deficit in category learning; the ability to generalize labels is important, but children must also recognize that there are limits to these generalizations. Labels lose their usefulness if they are extended indiscriminately to all sorts of items. What might help children learn, and especially retain, more coherent categories from screen media?

Study 2 explores two possible ways to ameliorate the differences in category learning observed in studies 1a and 1b between in person and screen mediated presentations of novel words and novel objects. Drawing on the two main proposed explanatory accounts of the video deficit effect, Study 2 explored specific changes to social and perceptual information present in the context of screen mediated word and category learning. Both of these factors represent important differences that exist between learning from a person and learning from a video. In Study 2 I manipulated the screen mediated learning context by adding contingent social interactions in one training condition, and by adding exposure to the physical objects to be learned in another training condition. An additional control condition further explored a slight variation on one of these manipulations. Together, Study 2 included two specific kinds of manipulations within the broader domains of social and perceptual factors. Therefore, this study is not intended to be an exhaustive investigation of the social and perceptual accounts of the video deficit effect. Rather, Study 2 tests how specific changes to the social and perceptual informational content available in the context of screen media influence toddlers' word and category learning performance.

Rationale

The video and in person conditions of Studies 1a and 1b differed from each other in at least two important ways. First, the video condition provided children with different and arguably impoverished perceptual information about the exemplar objects to be learned. Children in the video condition saw 2D representations of the objects on screen, both as stills and in motion. In contrast, children in the in person condition saw the actual 3D objects in front of them, and could even touch the objects if they wanted to. Second, the video condition lacked the interactive, contingent social interaction during training that was present in the in person condition. Recall that in the video condition even though the experimenter was present to start and stop the video, they sat quietly and did not engage the child in any way while the video was playing. Further, the video always presented objects and labels with the same timing, and was not contingent on individual children's attention or interest. For example, if a child became distracted and looked away from the video, they would simply miss the information presented in that moment. If this happened in the in person condition, the experimenter could get the child's attention again before continuing with training.

The social and perceptual characteristics of the two training conditions used in Studies 1a and 1b represent variations in the amount of information provided in each of these domains. When the dimensions of social and perceptual information are crossed, the training conditions used in Studies 1a and 1b can be seen as comprising opposite ends of both dimensions (see Table 3). In person training is high on social contingency and interaction, and offers plenty of perceptual information about the objects to be learned. Video training lacks any social contingency or interaction and involves relatively impoverished perceptual representations of the objects to be learned. Crossing these dimensions also creates two other possible ways to present information about the objects at training: one presentation that is high on social but low on perceptual information, and one that is low on social but high on perceptual information. In Study 2 I tested these two additional training conditions and compared them to the results of Study 1b. This comparison should offer insights into how manipulating the kinds of social and perceptual information available while learning from screen media may play a role in the category coherence differences observed in my task.

Table 3. Training conditions used in Studies 1 and 2 according to relative amounts of social and perceptual information provided.

		Social Information	
		High	**Low**
Perceptual Information	**High**	In person with objects present (Study 1)	Video *plus* objects present (Study 2)
	Low	Video *plus* in person labeling (Study 2)	Video only (Study 1)

Predictions

If children have overcome the video deficit effect in word learning, then toddlers in Study 2 should learn and retain novel one-to-one word-object mappings when tested in the target identification task. Studies 1a and 1b showed that 2½- to 3-year-old children can learn object labels equivalently from a person and from a video. Therefore children should also successfully learn the object labels in the two new conditions of Study 2.

If children have a general tendency to infer shape-based categories for the novel words that they learn, and are provided with enough information to accurately do so, this pattern should be evident in their forced choice generalization task performance. The key result will again be in how children generalize newly learned words to the arrays of objects matching the trained targets in either shape or material. One way to think about the two training conditions included in Study 2 is that each represents a specific improvement to being trained by video alone. The question is whether either improvement will influence generalization performance in a way that makes it equivalent to performance in the in person training condition, and more coherent than the video training condition, of Study 1b. In the end, this will indicate whether adding social or perceptual information can ameliorate the video deficit effect in category coherence.

Method

Participants. Thirty-four children (M_{age} = 32.7 mo., *SD* = 1.6 mo., 15 girls) were recruited for participation from the Boulder, CO area. These children were recruited from the same participant database as those in Study 1b, and thus were presumably drawn from the same population. Of these children, 3 were not able to return for a second visit within the required time window. Therefore, 31 children (M_{age} = 32.7 mo., *SD* = 1.7 mo., 13 girls) were included in the analyses presented here.

Materials. The novel object sets were the same as those used in Study 1b. For the condition in which social information was added to the video only training, new videos were shot so that the experimenter in the video was not talking (so as not to interfere with the physically present experimenter providing labels during training). Otherwise the actions performed in the video and the timing of those actions were the same as the training video from Study 1b.

The same survey and vocabulary measure as used in Study 1b were also given to parents of participants in Study 2 (see Appendix B).

Procedure.

Practice Trial. The same practice trial as was used in Studies 1a and 1b was again used as an initial warm-up for all children in Study 2.

Training. Children were trained on the novel words for novel target objects in one of two conditions. In the Video + Person condition, the physically present experimenter provided novel labels as each target object was shown in the video on screen. In the Video + Objects condition, children watched the same video that was used in the video training condition of Study 1b. As each object first appeared on screen, the physically present experimenter unobtrusively placed the corresponding physical object on the table in front of the child. Once the presentation of one object had completed in the video, the physically present experimenter removed that object from the table, and replaced it with the next one appearing in the video. Experimenters did not speak or interact with children during this training.

Testing. Both immediate and delayed testing were administered the same way as in Study 1b (see Study 1b Method section).

Results

Age and gender were analyzed with respect to the dependent measures reported below, and no significant effects were found. Therefore, these variables are not included in the following analyses.

Learning: target identification. In the first set of analyses I explored whether children accurately learned to map the novel names to the novel objects in the two new training conditions. To asses immediate learning, children's proportions of accurate target object choices at the first visit were compared to chance. Across both new conditions, children learned the novel word-object mappings marginally above chance ($M = .57$, $SD = .20$, $t(30) = 1.94$, $p = .06$, $d = 0.34$). The same test conducted on children's performance at the second visit showed that all children accurately retained those word-object mappings over the delay ($M = .61$, $SD = .19$, $t(30) = 3.32$, $p < .01$, $d = 0.60$). However when comparing immediate and delayed testing, there was no difference in accuracy ($t < 1$, $p > .05$). Comparisons between the two new training conditions at both immediate and delayed testing revealed no differences in accuracy. Children who learned from a video with objects present and children who learned from a person labeling a video were equally accurate at both time points ($t < 2$, $p > .05$ for both comparisons). A final comparison across studies showed that accuracy among all children and across time points in Study 2 did not differ from overall accuracy among all children in Study 1b ($t < 1$, $p > .05$).

As in Studies 1a and 1b, target identification accuracy at immediate testing was inspected to identify any children who did not learn the words. Eight subjects were excluded from further analyses because of immediate testing accuracy below 50%; five subjects from the Video + Person condition and three subjects from the Video + Objects condition. Target identification accuracy data without these excluded subjects were submitted to the above analyses. Accuracy increased slightly both at immediate ($M = .67$, $SD = .12$) and at delayed testing ($M = .63$, $SD = .17$), and was above chance at both times ($t(22) > 3.60$, $p < .01$, and $d > 0.75$ for both). Overall accuracy did not differ between visits, and accuracy among children in the two training conditions did not differ between groups either immediately or after the delay ($t < 2$, $p > .05$ for all comparisons). Children's accuracy in Study 2 still did not differ from accuracy levels observed in Study 1b ($t < 1$, $p > .05$).

The results of the forced choice target identification task confirm that children can learn and retain the novel word-object mappings presented to them in the two training conditions implemented in Study 2. This makes sense given the earlier results that even children who were trained by a video alone performed accurately on this task, and performed just as well as children trained in person. This lack of a video deficit effect for the target identification task shows that 2½- to 3-year-old children can effectively learn and retain word-object mappings when information is provided in various modalities, including screen mediated contexts. The training conditions used in Study 2 both have a component of screen-mediated learning, and so the results further confirm that children can learn words from screens. Next I turn to the generalization results.

Generalization analyses. Children in Study 2 completed the same forced choice and free choice generalization tasks as were administered in Study 1b, so I will analyze the results similarly.

Forced choice shape vs. shape task. For the forced choice generalization task in which children were presented with two shape matches from different sets and asked to identify one by name, the dependent measure was the proportion of correct choices over six trials. On average across the two training conditions in Study 2, children chose the correct shape match object at chance levels ($M = .56$, $SD = .23$, $t < 2$, $p > .05$). Children were not significantly different in accuracy between the training conditions ($t < 1$, $p > .05$). Looking at each group separately, children in both the Video + Person ($M = .58$, $SD = .20$) and Video + Objects ($M = .54$, $SD = .26$) training conditions performed no differently than chance ($t < 2$, $p > .05$ for both comparisons).

Overall, children in Study 2 were not able to accurately recognize and choose the correct shape match when asked for it by name.

These results contrasts with those found in Study 1b, particularly when compared to the group of children trained by video alone. While those children from Study 1b correctly identified shape match objects at above chance levels, neither of the new training conditions led to above chance performance. This was the case despite the fact that both of the new conditions added some extent of information to the screen mediated training context.

Forced choice shape vs. material task. In the remaining forced choice task, which was a constrained test of generalization, the dependent measure was the proportion of shape match choices across trials. On average across both training conditions children chose the shape match objects significantly more than chance ($M = .70$, $SD = .18$, $t(22) = 5.25$, $p < .001$, $d = 1.09$). Children in the two training conditions did not differ significantly from each other in this task ($t < 1$, $p > .05$). Analyzed separately, the group of children who were trained by a person labeling the video chose the shape matches significantly more than material matches ($M = .68$, $SD = .20$, $t(9) = 2.91$, $p = .02$, $d = 0.92$). Children trained by watching a video with objects present also significantly preferred shape matches in this constrained test of generalization ($M = .71$, $SD = .17$, $t(12) = 4.38$, $p < .001$, $d = 1.22$).

This result shows that in a constrained test of generalization, children in both of the training conditions of Study 2 preferred to extend novel words to objects that matched the trained targets in shape rather than those that matched in material. Performance in this task was similar to that observed in both training conditions of Study 1b. Next, analyses of the free choice generalization task can help identify whether either of the training conditions of Study 2 differed from either of the generalization patterns observed in Study 1b.

Free choice generalization task. The final set of analyses concerns the free choice generalization task, which was designed to measure the quality of the categories that children infer and retain in the novel word learning task. First, children's choices of shape and material match objects in both of the new training conditions of Study 2 were compared to each other and across visits. The dependent measure was the same as that used in Study 1b: proportions of choices for different object match types at test. Each child's responses were processed to produce a proportion of shape and material match choices at both immediate testing and at delayed testing. These proportions were first submitted to a 2 (Training Condition: Video + Objects or Video + Person) × 2 (Feature: shape or material) × 2 (Visit: first or second) mixed models ANOVA. Across visits and training conditions, children tended to select shape match objects (*M* = .93, *SD* = .16) more often than material match objects (*M* = .76, *SD* = .30), as shown by a main effect of feature ($F(1, 21) = 12.66$, $p < .01$, $\eta_p^2 = .38$). This analysis also revealed a trend in which children in the Video + Person training condition tended to choose all objects more often (*M* = .93, *SD* = .17) compared to children in the Video + Objects condition (*M* = .78, *SD* = .29; $F(1, 21) = 3.37$, $p = .08$, $\eta_p^2 = .14$). No other effects were significant.

In order to compare the results of Study 2 to those found in Study 1b, first children's proportions of object choices were submitted to a 4 (Training Condition: Video + Objects, Video + Person, In Person, or Video only) × 2 (Feature) × 2 (Visit) mixed models ANOVA. Across all four training conditions and both visits, children tended to select shape match objects (*M* = .95, *SD* = .13) more often than material match objects (*M* = .76, *SD* = .31), as shown by a main effect of feature ($F(1, 46) = 30.98$, $p < .001$, $\eta_p^2 = .40$). The only other effect to approach significance was an interaction between condition and feature ($F(3, 46) = 2.22$, $p = .098$, $\eta_p^2 = .13$).

In order to further explore patterns of similarities and differences among the four training conditions, each training condition of Study 2 was independently compared to both the in person and video training conditions of Study 1b. An interaction in this kind of analysis would indicate whether children's generalization performance in either of the new training conditions differed from either of the generalization patterns observed in Study 1b. These patterns can then be inspected to see whether the results of the new training condition are more in line with either the in person or video training condition of Study 1b. For these analyses, the data were also separated by visit because the results of the free choice generalization task in Study 1b differed qualitatively between immediate and delayed testing.

First, performance at immediate testing was analyzed. Free choice generalization data from the group of children in the Video + Objects training condition were submitted to a 3 (Training Condition: in person, video + objects, or video) × 2 (Feature) mixed models ANOVA. This analysis revealed only a main effect of feature ($F(1, 37) = 31.71$, $p < .001$, $\eta_p^2 = .46$). A similar analysis that compared performance of children in the Video + Person training condition to those in the in person and video training conditions of Study 1b also yielded only a main effect of feature ($F(1, 34) = 20.29$, $p < .001$, $\eta_p^2 = .37$). These results indicate that across all training conditions of both Studies 1b and 2, children preferred to extend newly learned words to objects matching the trained targets in shape (*M* = .95, *SD* = .13) more than those matching in material (*M* = .76, *SD* = .31) at immediate testing. The lack of any effects involving training condition indicate that this pattern was robust across all children.

Next, performance at delayed testing was analyzed in the same way. Free choice generalization data from the group of children in the Video + Objects training condition were submitted to a 3 (Training Condition) × 2 (Feature) mixed models ANOVA. Children in all three training conditions again preferred to extend novel words to shape matches more than material matches, as shown by a main effect of feature (*F*(1, 37) = 19.50, $p < .001$, $\eta_p^2 = .35$). In addition, free choice generalization at delayed testing also marginally differed by training condition (*F*(2, 37) = 2.71, $p = .08$, $\eta_p^2 = .13$). As can be seen in Figure 6, children in the video training condition tended to extend words to most of the test objects regardless of feature (*M* = .96, *SD* = .13), while children trained in person (*M* = .81, *SD* = .33) and with the video and objects (*M* = .79, *SD* = .33) generalized words to relatively fewer items overall. The interaction between training condition and feature did not reach significance.

These results suggest that performance among children trained by watching a video with the added perceptual information of having objects present is somewhere intermediate to the two distinct generalization patterns observed in Study 1b. This conclusion was reinforced by a post hoc comparison of generalization to different feature types within the Video + Objects training condition. Children in this group marginally tended to differentiate between shape and material matches at test (*t*(24) = 1.89, $p = .07$, $d = 0.67$), but this pattern was not quite robust enough to reach significance. If anything, children trained by Video + Objects tended to perform more similarly to children trained in person in Study 1b, but the results are not completely conclusive.

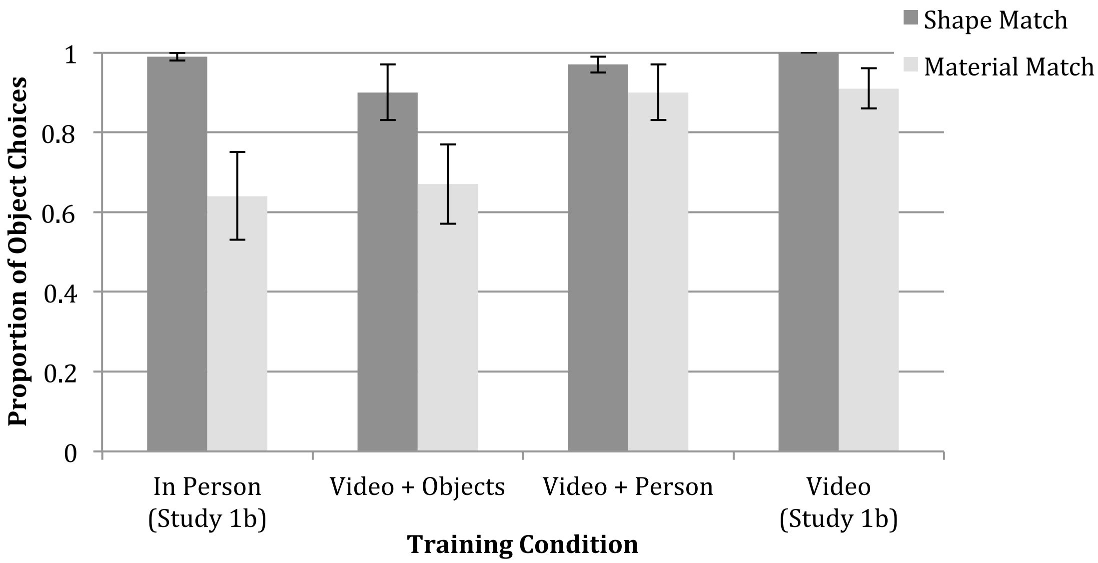

Figure 6. Free choice generalization performance at delayed testing (Visit 2) for children in the two training conditions of Study 2 (middle two pairs of bars) compared to children in the two training conditions of Study 1b (outer two pairs of bars).

Similarly, free choice generalization data from the group of children in the Video + Person training condition were submitted to a 3 (Training Condition) × 2 (Feature) mixed models ANOVA. The robust preference to choose more shape than material matches was again seen in a main effect of feature ($F(1, 34) = 15.04$, $p < .001$, $\eta_p^2 = .31$). The overall propensity to generalize words, regardless of test object type, was also found to differ between these three groups in a marginal main effect of condition ($F(2, 34) = 3.16$, $p = .06$, $\eta_p^2 = .16$). However, both of these main effects were also qualified by a significant interaction between training condition and feature ($F(2, 34) = 3.94$, $p = .03$, $\eta_p^2 = .19$). As can be seen in Figure 6, children's preferences to extend words to shape relative to material match objects in the free choice generalization task differed by training condition. Children trained by watching a video and receiving the extra social information of in person labeling performed similarly to children in the video training condition of Study 1b, and differed from the in person training condition.

The results of the free choice generalization test show that, first, immediate testing performance among children in the training conditions of Study 2 did not differ from that of children in the in person condition of Study 1b. This makes sense, given that immediate testing performance was also equivalent among children in both the in person and video only training conditions of Study 1b. Instead, and as seen in Study 1b, differences among the training conditions of Study 2 emerged at delayed testing. Comparing the delayed generalization performance of children trained by video with added objects to children in both conditions of Study 1b yielded inconclusive results. However, as shown in Figure 6, the generalization pattern of children in the Video + Objects condition is qualitatively more similar to children trained in person than those trained by video alone. On the other hand, having a physically present experimenter label objects as they are shown on video led children to generalize words similarly to those in the video training condition of Study 1b. That is, children in the Video + Person training condition exhibited a similar video deficit in retained category coherence as that observed in the prior experiment. These results indicate that increased social information, in the form of a physically present experimenter providing labels, did not ameliorate the video deficit effect in category coherence.

Control Condition

In an additional control condition I explored a slightly different way of manipulating the perceptual information available in the context of learning from a screen. A possible concern with the Video + Objects condition has to do with the procedure in which the experimenter places each target object in front of the subject as they learn about that object from the video. Although the experimenters did not speak to or interact with the subjects as they did this, the act of placing the objects on the table contingent with what was happening in the video may be seen as a sort of social interaction. To address this, I included a Video + All Objects training condition as a control for this implicit, unintended social interaction.

To control for the possibility of implicit social information I removed all experimenter involvement in training. Instead, children watched the video of an experimenter presenting and labeling the novel objects while having a tray of all six target objects in front of them. As in the video condition of Study 1b, the experimenter sat quietly during training and did not interact at all with either the subject or the target objects. Children in this condition were free to examine and manipulate the objects in any way they chose while watching the video.

Fifteen additional toddlers participated in the Video + All Objects training condition (M_{age} = 32.4 mo., *SD* = 2.2 mo., 7 girls). Of these, 13 showed above 50% accuracy in target identification at immediate testing, and thus were included in subsequent analyses. First I tested how well children learned the novel words for the target objects at their first visit to the lab. Children's accuracy in the target identification task administered immediately after training was significantly above chance ($M = .72$, $SD = .16$, $t(12) = 4.98$, $p < .001$, $d = 1.38$). However, unlike the other training conditions examined so far, children's accuracy decreased significantly over the delay ($t(12) = 2.42$, $p = .03$, $d = 0.62$). By the second visit, children's target identification accuracy was no different than chance ($M = .60$, $SD = .21$, $t < 2$, $p > .05$). However, although delayed accuracy was numerically slightly lesser than that observed in the other two training conditions of Study 2, there were no significant differences in these values ($t < 1$, $p > .05$ in independent t-tests).

Next I examined children's forced choice generalization performance. At delayed testing, children's accuracy in identifying the correct shape match to the target items was also at chance levels ($M = .53$, $SD = .10$, $t < 2$, $p > .05$). However, in the delayed forced choice task pairing a shape match and a material match from each set, children consistently selected the shape match objects (*M* = .67, *SD* = .19, $t(12) = 3.12$, $p < .01$, $d = 0.87$). Overall, among children who did learn the names for the trained target items initially in the Video + All Objects training condition, there was a decrease in accuracy over the delay. These children did not completely retain the correct word object mappings, either in identifying the trained target items or in identifying objects that matched the targets in shape. However, in a forced choice test of generalization, these children did consistently choose the shape match objects over the material match objects.

How did children in the Video + All Objects training condition perform in the free choice generalization task, particularly in comparison to children trained in person and by video in Study 1b? Data from immediate generalization testing were first submitted to a 3 (Training Condition: in person, video + all objects, or video) × 2 (Feature: shape or material) mixed models ANOVA. As in similar analyses reported above, this yielded only a main effect of feature ($F(1, 37) = 19.43$, $p < .001$, $\eta_p^2 = .34$). Children in all three training conditions extended words preferentially to shape match objects more than material match objects when tested at the first visit.

Next, data from delayed free choice generalization testing were submitted to a similar 3 (Training Condition) × 2 (Feature) ANOVA. This analyses revealed an overall preference for shape over material match objects ($F(1, 37) = 17.85$, $p < .001$, $\eta_p^2 = .33$) as well as a difference between training conditions in how much children chose objects at test overall ($F(2, 37) = 4.13$, $p = .02$, $\eta_p^2 = .18$). However, these two main effects were also qualified by a significant interaction ($F(2, 37) = 4.43$, $p = .02$, $\eta_p^2 = .19$). As can be seen in Figure 7, children trained by watching a video with all objects present in front of them performed similarly to children trained by video alone, and differently than those trained in person, in Study 1b. This pattern of results is qualitatively similar to that seen for the Video + Person training condition reported above. Having all objects present while watching a video to learn about novel words did not help children retain coherent categories, based only on one characteristic feature, over a delay.

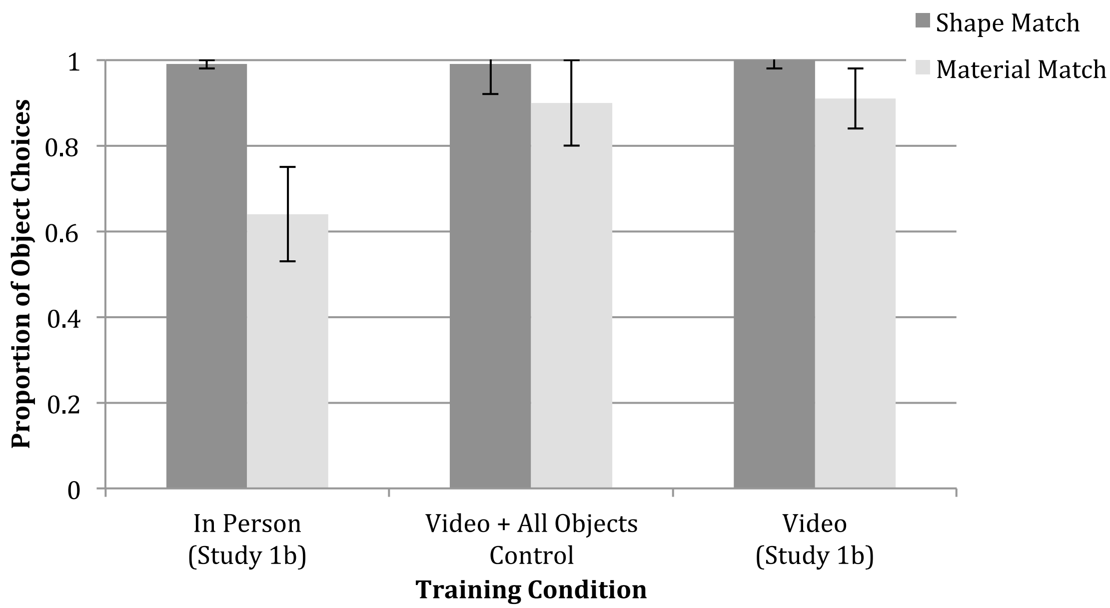

Figure 7. Free choice generalization performance at delayed testing (Visit 2) for children in the control condition of Study 2 (middle pair of bars) compared to children in the two training conditions of Study 1b (outer two pairs of bars).

Discussion

The goal of Study 2 was to explore the impact of specific forms of social and perceptual information available in the screen mediated learning context on children's learning, retention, and generalization of words and categories. To this end, I created two additional novel word training conditions that added information back to the video only training condition of Study 1b. Extra perceptual information was added in the Video + Objects training condition and extra social information was added in the Video + Person training condition. As predicted, children in both of these training conditions were able to effectively learn and retain the trained novel words for the novel objects. These children were also all able to generalize words coherently by similarities in shape over and above similarities in material when tested immediately. However, performance differences emerged in delayed generalization testing, particularly when each new training condition was compared to the two distinct generalization patterns observed in Study 1b.

In analyses of children's free choice generalization at delayed testing, I found that having each physical target object present at the same time that it was being shown and labeled on screen helped children to perform similarly to children who had learned the words in person. Although this result was not completely conclusive, children given increased perceptual information tended to differentiate shape from material matches in their novel word generalizations, and performed qualitatively more like children in the in person than the video training condition of Study 1b. In contrast, having a present experimenter contingently label objects appearing in a video on the screen led children to perform more similarly to children trained by video alone in the previous experiment. Children's performance resulting from increased social information at training was suggestive of the video deficit in retained category coherence. That is, these children did not show a clear preference for shape match objects over and above material match objects. This result suggests that having a physically present person label a video on screen does not alleviate the video deficit in retained category coherence. Instead, the results tentatively suggest that adding physical objects to the screen learning context may help ameliorate this specific deficit.

The implications of these results can be further developed by considering the results of the control condition. In the control condition I addressed the fact that having an experimenter place target items in front of children during training (as was done in the Video + Objects training condition) might be considered a kind of social interaction. To better isolate the influence of the added perceptual information of having physical objects present during training, I removed this social aspect. Children in this control condition initially learned the new words and generalized them consistently by similarities in shape after being trained by a video with all objects present in front of them. However, by the second visit, two key changes were observed. First, children did not accurately retain the one-to-one word objects mappings, as seen by a drop in target identification accuracy. Second, children no longer generalized the words consistently by one feature over another. Their retained category coherence looked similar to the children who had been trained by video alone in Study 1b.

Taken together, the experiments of Study 2 start to give some insight into the roles of both social and perceptual information in screen mediated word and category learning, and also lead to further questions. The results of the Video + Objects training condition and the Video + All Objects control condition suggest that the facilitative effect of physically present objects on retention of coherent categories is mediated by other factors. That is, having objects present during learning only tended to help when the experimenter placed each target object in front of the child at the moment that it was being presented in the video. This suggests that having access to richer, fuller perceptual information is only helpful when there is some measure of social information to help guide learning. On the other hand, the results of the Video + Person training condition showed that adding more social information did not help children retain more coherent categories when learning from a screen. This adds to the above conclusion: social information may only helpful in guiding learning when there is adequate perceptual information available as well. Together, these results suggest that social and perceptual information are both important and work together in word and category learning from a screen, but what role specifically does each factor play?

A big part of learning in this task, especially learning categories, depends on perceptual information. For children to coherently categorize by shape, they need to encode a good representation of the physical shape properties of the novel objects. Perhaps it is also helpful to encode a good representation of the material properties in order to rule those out as a basis for categorization. Yet to be able to use this perceptual information, children need to allocate their attention in specific ways. One way is in attending to the physical properties of the objects, especially shape and material. Another way is in attending to this perceptual information at the right time – that is, when the object that is at the focus of attention is being labeled.

The results of Study 2 suggest that social information may play a role in helping children allocate attention. This is suggested especially in the results of the training and control condition that involved having physical objects present. The subtle social information conveyed by having an experimenter place each object in front of children as they were learning about it on screen may have aided retained category coherence because it signaled to children to attend to what was important in a given moment. On the other hand, having all of the target objects available while watching the training video may have been too attentionally demanding. Without some sort of guidance on which physical object to attend to at any given moment, children may have missed out on encoding important information at training. This possibility is supported by the fact that the Video + All Objects control condition was the only condition in which children did not accurately retain the correct names for the target objects. Overall, social and perceptual information work together. Social signals, either in speech or action, help children guide their attention to what is being labeled and what is important to attend to at a given moment. Once attention has been deployed effectively, the quality of perceptual information available has a strong influence on how children learn, generalize, and retain words and categories.

A question that may be raised about all of the experiments included in Studies 1a, 1b, and 2 has to do with how children's learning is impacted by the transfer from training to testing. While children in the in person training condition showed the clearest pattern of retained category coherence, these children also experienced the least amount of change when transitioning from the training to testing phases of the experiment. That is, these children learned about physical objects directly from interactions with an experimenter and were subsequently tested on physical objects through interactions with that same experimenter. In all of the other training conditions there was some extent of transfer in context from training to testing. Are the observed effects merely due to the process of transfer? There are at least two reasons why transfer may not be the primary force at work within the current experiments.

First, the various training conditions included in the current experiments represent differing degrees of transfer from the training to testing contexts. The training conditions of Study 1b represent two extremes. As noted above, the in person training condition did not require any transfer because the training and testing contexts were the same. In the video only condition, children had to transfer their learning about 2D representations from a video of an experimenter to a testing context involving physical objects and direct, in person interactions with the experimenter. This change in context could provide a possible explanation for the observed deficit in retained category coherence seen in the video training condition. However, the training conditions included in Study 2 also involved context transfer, but to differing degrees. The change from training to testing in Study 2 could be thought of as partial transfer because some aspects did change but some stayed the same. For example, in the training condition involving increased social information, the social context remained constant from training into testing. Similarly, this was the case with the perceptual context in the conditions involving increased perceptual information. Although all of the training and control conditions of Study 2 involved some partial transfer in context from training to test, there was not a uniform deficit in performance across all conditions.

Second, there was a delay between when the training to testing context transfer took place and when differences in learning were observed in the current experiments. If children experienced a change in context when transitioning from training to testing, this only happened at their first visit to the lab. Yet across conditions, children accurately learned the novel word-object mappings and inferred coherent shape-based categories at initial testing. It may be possible that context transfer only has a delayed effect in the current task, but it is worth noting that there are no transfer effects immediately after the transfer actually takes place. Altogether, it seems that the context of learning, particularly the combination of perceptual and social information available, has more of an impact on outcome than the process of transfer itself. However, this issue of context transfer is interesting in its own right and will be returned to in the general discussion.

In sum, Study 2 builds on Study 1b by showing how both perceptual and social information play a role in children's learning, generalization, and retention of words and categories from screen media. In the current task, rich and contingent social information seems to play a role of helping children efficiently allocate their attention to what matters in the context of learning. But the effectiveness of this learning also depends on the quality of the perceptual information available to children during learning. Although the addition of physical objects to the screen learning context was the only manipulation that even started to help children maintain coherent categories over time, social factors might have played a role as well. The final study focuses more on the relationship between social factors and screen mediated learning. In Study 3

I investigated the more naturalistic context of children learning from screens while co-viewing with a parent, exploring how different aspects of parental behavior relate to children's learning outcomes in my task.

Chapter 6: Study 3

In the third and final study I investigated factors that may be associated with 2½- to 3-year-old's screen mediated learning in the current experimental task. I focused on a common context for young children's screen mediated learning: co-viewing screen media with an adult. As reviewed earlier, research has shown that the quality of parental interactions during co-viewing has an influence on children's attention to and learning from screen media. Although both the quantity and quality of parent-child interactions tend to suffer in the presence of screen media when compared to other common interactive contexts, certain kinds of interactions can have a positive impact in the context of co-viewing. Several studies have demonstrated that sensitive, reciprocal, and content-focused co-viewing interactions can facilitate infants' and toddlers' attention to screen media (e.g., Barr et al., 2008; Demers et al., 2012; Fender et al., 2010; Fidler et al., 2010), but only one of these studies has actually attempted to look at word learning outcome. Fender and colleagues (2010) found that children of parents who focused more on teaching vocabulary words presented in a video tended to produce more of those key vocabulary words. However, the authors used a broad measure to capture how much parents focused on teaching the target vocabulary words, simply counting any utterance that included one of the words. This measure does not take into account more fine-grained qualitative differences in how parents may use vocabulary words or talk about the content of a video. Further, the authors measured children's word learning based only on the words that children produced during co-viewing. More structured tests of word learning could better assess, for example, how well children retain and recognize newly learned words at a later time.

The literature on the relationship between co-viewing and young children's learning from screen media is still in it's infancy, and thus many questions remain. One question is how specific characteristics of parental speech during co-viewing relate to children's learning of screen media content. Another question is how co-viewing is associated with various controlled and structured measures of learning. Study 3 aims to address these two questions by investigating the fine-grained characteristics of parental speech during co-viewing, and linking that to children's word and category learning, retention, and generalization. In this way Study 3 makes a novel contribution to the literature and paves the way for future research on this topic.

Rationale

The goal of Study 3 was to connect parental co-viewing behaviors to learning outcomes in children, specifically in the measures of learning, generalization, and retention of novel words and categories used in the current studies. Parents were asked to actively watch the kind of novel word training video used in the previous studies with their children. Children were then tested on their learning of the video content, as well as on standardized language measures, and a control measure of real word learning. Parental co-viewing behaviors were coded, analyzed using principal component analysis, and examined as predictors of children's word learning, retention, and generalization. Standardized language measures and the control word learning measure were included as control variables in these analyses. Therefore, this study combines an analysis of children's linguistic environments, a standard technique in the field of linguistics (e.g., Cameron-Faulkner, Lieven, & Tomasello, 2003; Kavanaugh & Jirkovsky, 1982; Rondal, 1980), with an evaluation of learning outcome using structured, controlled measures. This cross-disciplinary approach should provide a unique perspective on the question of how co-viewing relates to learning. These results will add detail to what is known about the characteristics of parental speech during co-viewing and identify whether any of these characteristics are associated with toddler's word and category learning from screen media content. Because co-viewing speech is not directly manipulated in this study, it is not possible to draw causal conclusions about how co-viewing may influence learning outcome. However, the results will help inform future experimental work that can investigate such causal relationships.

Predictions

Although the previous studies reported here indicate that 2½- to 3-year-olds can learn and retain novel word-object mappings from screen media, parental behaviors during co-viewing may be useful predictors of these measures. A broad prediction is that the extent to which parents focus their speech on the key information to be learned on screen should predict children's learning outcomes. In line with prior research (e.g., Barr et al., 2008; Demers et al., 2012; Fender et al., 2010; Fidler et al., 2010), parents who are responsive and help their children focus on the content of the video should promote their children's attention to and learning from the screen. In this study that means talking about specific objects shown and novel words presented in the video.

More specific predictions can be made about the different types of tasks included. First, if labeling the novel objects as they appear in the video is helpful for word learning, then parents who focus more on the novel labels being taught in the task should be associated with children better learning the one-to-one word-object mappings. An emphasis on the novel labels to be learned for the novel objects in the video should help children establish and retain these mappings. Second, if drawing attention to information about the novel objects to be learned is helpful for category learning (as suggested by Study 2), parents who direct children's attention to details of the objects in the video should be associated with children better inferring coherent categories in the generalization tasks. Coherent categorization depends on children's differentiation between various features of the objects, and so co-viewing behavior that supports attending to those features should in turn support categorization.

The parental speech coding categories included in this study were designed to help test these predictions (see Table 4). For example, two of the codes captured how often parents used the novel labels that were taught in the video. These codes characterized whether parents elicited the labels from children in the form of a question, or simply produced the labels themselves. The inclusion of these codes will help test the prediction outlined above that focusing on the labels to be learned will be related to one-to-one word-object mapping. The results will also provide information about whether different uses of the target labels associate in different ways with children's learning.

Other codes help test the second main prediction about a link between speech that focuses attention on the objects to be learned and children's generalization performance. The extent to which parents directed attention to and talked about the content of the word training video was captured by several codes. These included explicitly directive attentional vocatives, descriptions of the specific objects shown on screen, and abstracting connections between the screen content and real world items or experiences. Other codes captured how parents directed attention more broadly by talking about the screen media itself or talking about something unrelated to the video content. Contrasting these different kinds of codes allows for testing whether parental focus on the specific content to be learned is associated with learning and generalization.

Table 4. Coding categories used to characterize parental utterances during co-viewing in Study 3.

Coding Category	Description	Examples
Questions		
Wh-	Questions beginning with what, who, when, where, or how.	"What is that?"
Yes-no	Questions meant to elicit a yes or no response.	"Did you see that?"
Label elicitation	Questions meant to encourage the child to produce the item label.	"Can you say ____?"
Tag	Questions appearing at the end of a statement.	"That's a nice zeb, isn't it?"
Labels	Single referents or phrases that only provided the novel label.	"Zeb."; "That's a zeb."
Descriptions	Utterances that included additional information about an item, such as adjectives.	"A green zeb."
Abstractions	Utterances that connected real world items or experiences to the content presented on screen.	"Do you have a toy like that?"; "That looks like a hat."
Attentional vocatives	Utterances that explicitly directed the child's attention to the screen.	"Look at that!"
Responsive feedback		
Confirmations	Positive feedback in response to something the child said.	"Yes, that's right, that's a zeb!"
Corrections	Negative feedback in response to something the child said.	"No, that's a lug."
Evaluations	Utterances that expressed a judgment about something presented on screen.	"That's a nice lug."
Interactive verbalizations	Utterances about how to interact or what to expect from the screen.	"It says we're going to learn about some new toys."
Verbalizations unrelated to media content	Utterances that were unrelated to the content on screen or the child's behavior in relation to the screen.	"We'll play with the train later."
Placeholders	Utterances that did not provide any new information.	"Okay."; "Oh."
Uncodeable verbalizations	Unclear speech or utterance that did not fit into any other category (< 0.5% of the data were uncodable).	

Method

Participants. Fifty children (M_{age} = 32.1 mo., *SD* = 1.3 mo., 28 girls) were recruited for participation from the Boulder, CO area.

Materials. The same novel object stimuli that were used in Studies 1b and 2 were used again in Study 3, with one addition: a distractor object was added to each novel object set. This object did not match the trained target object in any features (see Figure 8). Therefore, in the free choice generalization task children were presented with five objects: two that matched the trained target in shape, two that matched the trained target in material, and one non-match distractor. The distractor objects were included to provide another measure of how well children inferred and retained coherent categories, beyond that captured by a relative preference for shape over material matches.

Figure 8. An example novel object set from Study 3. Each set consisted of a target object (top), two shape match objects (middle left), two material match objects (middle right), and one non-match distractor object (bottom).

Training involved watching a video similar to that used in earlier video training conditions, but that was recorded with a research assistant who was not an experimenter in the study. This was done to be more naturalistic; children typically do not actually meet or interact with the people that they see on TV or in videos. The video was also edited to include several written cues as signals to parents (as will be described in the Procedure section).

An additional word learning task was added in Study 3 as a control measure. This involved training children on 5 real but unfamiliar words (*galoshes*, *amphibian*, *arachnid*, *canteen*, and *hexagon*). Training and testing materials consisted of laminated pictures of items.

The same survey and vocabulary measure as used in Studies 1b and 2 were also given to parents of participants in Study 3 (see Appendix B). Children were also tested on the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007) at the end of their first visit as an additional measure of language development. The PPVT is a standardized measure of receptive vocabulary.

Parent-child interactions during training video co-viewing were recorded and later transcribed in ELAN (Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands: http://tla.mpi.nl/tools/tla-tools/elan/; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). Parental verbalizations were subsequently coded using a scheme adapted from Barr et al. (2008). Barr and colleagues developed 11 coding categories based in part on studies of parent-child book reading and in part on parent-child interactions observed in their study on video co-viewing. Using these coding categories as a starting point, I developed a coding scheme for Study 3 that includes 15 coding categories (see Table 4 for descriptions and examples of each category). The main changes from the Barr et al. (2008) coding scheme were removing coding categories that were specific to their study (e.g., singing) and adding categories relevant to the current study (e.g., label elicitation questions). Twenty percent of the data were independently coded by two people, with 86% agreement. Disagreements in coding were resolved, and then one person completed the rest of the coding.

Procedure.

Training. The training and testing procedures were similar to those used in the video only training condition on Study 1b. The key difference in training was that parents were encouraged to actively co-view the video with their children. After the experimenter administered the "ball" practice trial, they explained that the video was meant to teach some new words and that the parent could talk to their child about the video as they would while watching at home. The experimenter left the room for the remainder of training. At the beginning of the video text appeared on the screen that said "Now let's learn about some new toys!" Halfway through training another screen of text appeared that said "Let's see those again!" This was meant to signal to parents that the presentation of the novel words and objects would repeat. At the end of training a screen appeared that said "Thanks for watching! Now let's play the game!" The experimenter re-entered the room at this point to administer the testing tasks.

Session 1 Testing. Children were tested on their immediate learning of the novel words and categories the same way as in Studies 1b and 2. Children completed the free choice generalization task followed by the forced choice target identification task. The PPVT was then administered and concluded the first session of the experiment.

Session 2 Testing. Children returned to the lab about a week after their first session and were tested in much the same way as in Studies 1b and 2. After completing the "ball" practice trial again, children were given the free choice generalization task and the target identification task. They also completed a forced choice shape vs. shape task and a forced choice shape vs. material task.

An additional word learning task was administered at the end of the second session as a control measure. In this task, children were trained and then immediately tested for learning on 5 real but unfamiliar words. Experimenters presented printed images of items while telling children about each word. The descriptions and pictures used to train on each word were drawn from *Sesame Street* clips (see Appendix C). Following this training, children were given two tests of learning. First they completed a generalization test in which they were presented with an array of images: the trained exemplar, two category members, two category non-member distractors, and a thematically related item. Children were asked to select items by trained name. The second testing task was a forced choice target identification test. Pairs of trained exemplar images were presented and children were asked to select one by name.

Results

Before conducting any analyses, several subjects were removed from consideration due to complications with their co-viewing observation data. One subject was excluded due to missing co-viewing data (a technical problem with the video recording equipment); one subject was excluded due to a very low amount of co-viewing interaction (the parent only produced one utterance during training); and four subjects were excluded due to their co-viewing interactions taking place primarily in a language other than English (the languages used were Korean, Swedish, and German). Because these circumstances interfered with coding the co-viewing data from these subjects, all of their data were excluded from the analyses reported here. Therefore, the final sample included in the following analyses consisted of 44 children (M_{age} = 32.1 mo., *SD* = 1.3 mo., 25 girls).

Behavioral data. First I analyzed the behavioral data collected on children's learning at both the first and second testing sessions. For Study 3 this included both the novel word learning tasks similar to those used in the other studies, as well as the additional word learning control task administered at the second visit that involved real but unfamiliar words. These analyses are meant to provide an overview of how children in Study 3 performed as a whole before exploring the ways in which parental co-viewing predicted learning across individuals. The overview of children's word learning and generalization performance will follow the same structure as that used in the previous studies. Also as in the previous studies, age and gender were analyzed with respect to the dependent measures reported below, and no significant effects were found. Therefore, these variables are not included in the following analyses.

Learning: target identification. As in Studies 1 and 2, my first question in Study 3 was whether children accurately learned and remembered the novel word-object mappings presented in the training video. First, children's proportions of correct target object choices at each testing session were compared to chance. Across the sample, children were accurate at above-chance levels at immediate testing ($M = .57$, $SD = .22$, $t(43) = 2.20$, $p = .03$, $d = 0.33$). These children also retained this level of accuracy over the weeklong delay between testing sessions ($M = .58$, $SD = .20$, $t(43) = 2.58$, $p = .01$, $d = 0.39$). Next, children's target identification accuracy was compared between the two testing sessions. A paired t-test revealed no difference in accuracy ($t < 1$, $p > .05$), confirming that children retained the word-object mappings they had learned initially. Overall, children in Study 3 accurately learned and retained the novel word-object mappings presented in the training video that they watched with a parent in the lab.

Generalization analyses. The next set of analyses examined the nature of the categories that children inferred based on the word-object mappings they learned from the training video. Generalization was assessed in two forced choice and one free choice task examining the extent to which children extended novel names to and preferred objects that matched the trained targets in either shape or material.

Forced choice shape vs. shape task. The first forced choice generalization task tested whether children accurately mapped the newly learned novel words to objects that matched the trained target items in shape. When presented with two shape match objects from two different item sets and asked to identify one by a trained novel word, children chose the correct shape match at above chance levels ($M = .60$, $SD = .20$, $t(43) = 3.10$, $p < .01$, $d = 0.47$). Overall, children were accurate in identifying objects that matched the original trained targets in shape based on the trained novel words.

Forced choice shape vs. material task. The second forced choice generalization task was a constrained test of children's preference to extend the novel trained words to other objects that matched the trained targets in shape over material. When presented with one shape match and one material match object from the same item set, children selected the shape match objects a significant majority of the time ($M = .63$, $SD = .27$, $t(43) = 3.10$, $p < .01$, $d = 0.46$). In this constrained test, with only two options available, children in Study 3 tended to prefer to extend the trained novel words to objects that matched the original trained target in shape specifically.

Free choice generalization task. The results of the free choice generalization task reflect the nature of the categories that children inferred based on the novel word-object mappings they learned from the training video. I analyzed children's performance in this task in two ways. First, consistent with Studies 1b and 2, I focused on how children generalized novel words to objects that matched the trained targets in either shape or material. The proportions of time that children chose the available shape and material match items at test were submitted to a 2 (Feature: shape or material) × 2 (Visit: first or second) repeated measures ANOVA. Children showed a consistent preference for shape match items, as shown by a main effect of feature ($F(1, 43) = 32.78$, $p < .001$, $\eta_p^2 = .43$). As can be seen in Figure 9, children selected a higher proportion of the available shape match objects (*M* = .81, *SD* = .25) compared to material match objects (*M* = .62, *SD* = .35) on average across both immediate and delayed testing. No other effects reached significance. In this unconstrained test of generalization, children in Study 3 as a group tended to infer shape-based categories based on the novel word-object mappings they had learned.

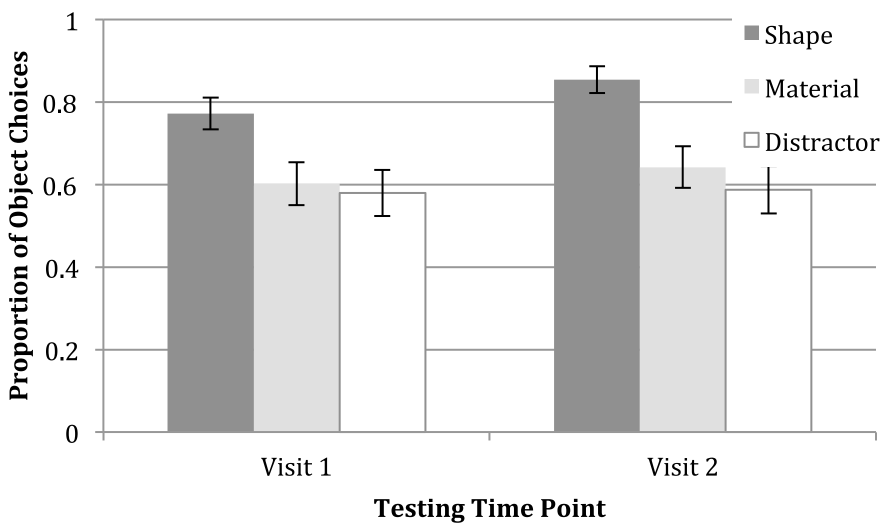

Figure 9. Free choice generalization performance at immediate and delayed testing for children in Study 3.

The other free choice generalization task analysis focused on children's choices of shape match objects compared to their choices of distractor non-match objects. Whereas the above comparison between shape and material match object choices established that children tended to infer shape-based categories, an analysis including distractor objects will help further establish whether these inferred categories were coherent in a more general way. Children's proportions of choice for each type of object were submitted to a 2 (Item type: shape match or distractor non-match) × 2 (Visit) repeated measures ANOVA. Children did indeed differentiate between shape match objects and distractors, as shown by a main effect of item type ($F(1, 43) = 30.65$, $p < .001$,

η_p^2 = .42; see Figure 9). Children's proportion of distractor object choices averaged .58 (*SD* = .38) across both testing sessions. No other effects were significant. Together, the results of the free choice generalization task show that children in Study 3, as a group, tended to infer and retain shape-based categories by extending the novel words to shape match objects more often compared to their choices of either material match objects or distractor non-match objects.

Word learning control. The final behavioral task administered at the end of the second testing session was a word learning task that involved real, unfamiliar words. This task was included as a baseline control measure to assess how effectively children were able to immediately learn and generalize real words trained in person. Although performance on this task will later be used only as a control variable (and not an outcome variable), an investigation of these measures will establish how effectively the current sample of toddlers are able to learn new words in the moment. First I analyzed children's target identification performance by comparing proportions of correct item choices to chance. A t-test comparison revealed that children accurately learned the one-to-one word-item mappings at above chance levels (*M* = .68, *SD* = .30, $t(43) = 4.02$, $p < .001$, $d = 0.61$).

Having learned these real but unfamiliar words, what kinds of categories did children infer? The proportions of time that children chose each type of generalization test item out of all the available items of each type were submitted to a one-way repeated measures ANOVA. There was a significant effect of item type ($F(3, 129) = 49.79$, $p < .001$, $\eta_p^2 = .54$) such that children tended to select the trained target items most often (*M* = .87, *SD* = .23), followed by untrained category members (*M* = .69, *SD* = .28), followed by category non-members (*M* = .46, *SD* = .36), and finally by thematically related items, which were chosen least often (*M* = .41, *SD* = .38). Post-hoc t-tests revealed that the trained target items were chosen significantly more often than any other type of item presented at test ($t(43) > 5.30$, $p < .001$, $d > 0.80$ for all comparisons). Children chose the untrained category member items significantly more than either the category non-members or the thematically related items ($t(43) > 5.50$, $p < .001$, $d > 0.85$ for both comparisons). Children also chose the category non-members more often than the thematically related items ($t(43) = 2.20$, $p = .03$, $d = 0.33$). Altogether these results confirm that children in Study 3, as a group, accurately learned the correct word-item mappings, as shown by their strong preference to extend the words to the target items, but also that they correctly generalized those words to other category members that had not been present at training. Further, children's choices showed an accurate distinction between category members and non-members. Somewhat surprisingly, children chose thematically related items least often, even less often than category non-members.

Behavioral data overview. The results thus far provide an overview of how children in Study 3 learned, generalized, and remembered words and categories. As a group, these children accurately learned the one-to-one novel word-object mappings, and retained this information with the same level of accuracy over a one week delay. Children in the current study also showed a preference to extend these newly learned words based on similarities in shape, both in constrained and unconstrained tests of generalization. At both immediate and delayed testing, children inferred shape-based categories rather than material-based categories, and consistently differentiated shape match objects from non-match distractors. The control word learning task additionally demonstrated that children in Study 3 overall accurately learned word-item mappings and generalized categories for real, unfamiliar words that were taught in person.

The key question of Study 3 is how this learning, generalization, and retention performance was associated with parental co-viewing behaviors that took place as children were learning about the novel words and objects. The next set of analyses will incorporate the co-viewing data collected during initial video training. Parental co-viewing behaviors and styles will be examined as potential predictors of how effectively children learned new words from the video.

Factor Analysis. An exploratory factor analysis using principal component analysis was conducted to reduce the dimensionality of the co-viewing data. A preliminary inspection of the coded co-viewing data showed that three of the coding categories were used by fewer than one third of parents in the sample: abstractions, corrections, and uncodeable utterances (see Table 5). These coding categories were excluded from the factor analysis. One other category, placeholders, was also excluded due to problems of interpretability. Placeholder utterances were defined as utterances that did not provide any new information, and so did not clearly fit into the analysis.

Table 5. Percentage of parental speech during co-viewing that fell into each of the coding categories (mean and standard deviation) and the percentage of all parents in the sample that used each category during the observed co-viewing interaction.

	Percentage of parental co-viewing speech		Percentage of parents that used this type of utterance
	M	*SD*	
Questions			
Wh-	14.38%	18.60%	75%
Yes/no	9.41%	10.15%	75%
Label elicitation	3.31%	5.98%	38.64%
Tag	1.49%	2.27%	34.09%
Labels	38.92%	16.78%	97.73%
Descriptions	8.38%	12.44%	52.27%
Abstractions	1.72%	3.97%	20.45%
Attentional vocatives	2.14%	3.45%	38.64%
Responsive feedback			
Confirmations	5.31%	7.52%	54.55%
Corrections	0.16%	0.80%	4.55%
Evaluations	2.10%	3.70%	40.91%
Interactive verbalizations	7.28%	6.24%	95.45%
Unrelated verbalizations	2.04%	4.39%	34.09%
Placeholders	3.12%	4.50%	47.73%
Uncodeable	0.24%	0.77%	9.09%

Therefore, 11 coding categories were entered into an initial factor analysis. Initial results revealed that three codes did not meet criteria for inclusion (Tabachnick & Fidell, 2007). Attentional vocatives did not load strongly enough on any of the resulting components (all loadings < .50). Yes/no questions and unrelated verbalizations each loaded onto their own independent components and did not associate with any of the other coding categories. These three categories were subsequently excluded.

A final factor analysis included 8 of the coding categories and resulted in four components with Eigenvalues above 1.0. Correlations between each of the included co-viewing coding categories are shown in Table 6. The first component resulting from the factor analysis explained 19.47% of the variance, the second 18.94%, the third 18.74%, and the fourth 18.01%

for a total explained variance of 75.15%. An orthogonal Varimax rotation was used to facilitate interpretation of the components. An oblique Promax rotation yielded the same components, but showed only weak correlations between factors (< .20), and so was not used for subsequent analyses.

Table 6. Correlations between parental co-viewing codes included in the final factor analysis.

	Wh- questions	Label elicitation questions	Tag questions	Labels	Descriptions	Confirmations	Evaluations	Interactive verbalizations
Wh- questions	--							
Label elicitation questions	-.16	--						
Tag questions	-.30*	.28*	--					
Labels	-.45**	-.10	-.01	--				
Descriptions	-.31*	-.09	.25†	-.13	--			
Confirmations	-.07	.27*	.01	-.13	-.34*	--		
Evaluations	-.13	-.10	-.07	-.11	.12	-.13	--	
Interactive verbalizations	-.01	-.06	-.12	-.19	-.27	-.11	.43**	--

Note. †$p < .10$. *$p \leq .05$. **$p \leq .01$.

The four components are shown in Table 7 with factor loadings on each coding category. The first component includes tag questions and descriptions of the items shown in the video. Tag questions, which are statements with a question appended at the end, were typically used by parents to talk about the items on screen. Many of the observed instances of tag questions were also descriptions (e.g., “it’s a green lug, huh?”). This component will now be referred to as describing objects because it captures the extent to which parents focused their speech on the individual objects depicted on the screen. The next component includes label elicitation questions and confirmations. These two types of utterances also appeared together often in co-viewing speech. This component will be referred to as label elicitation and feedback because it captures how often parents explicitly asked children to produce labels, including giving positive feedback for doing so. The third component includes evaluations and interactive verbalizations. Parents often used evaluations to make general comments about the video or item shown on screen (e.g., “that’s a cool gub”), and interactive verbalizations were comments about the video itself (e.g., “it says let’s see those again”). This component will be referred to as narrating because it captures parental speech about the video and screen viewing context more broadly. The fourth and final component includes wh- questions and explicit labeling. These codes loaded in opposite directions onto this component, so the component will be referred to as open-ended questions vs. explicit labeling. In the positive direction this component captures the extent to which parents provided the novel labels being taught on screen, and in the negative direction it captures the extent to which parents asked their children open-ended questions about the video (e.g., “what is that?”).

Table 7. Factor loading values and communalities (h^2) for the four components resulting from the factor analysis.

	Describing Objects	**Label Elicitation & Feedback**	**Narrating**	**Open-Ended Questions vs. Explicit Labeling**	h^2
Tag questions	**.664**	.389	-.057	.125	.61
Descriptions	**.853**	-.341	-.060	-.040	.85
Label elicitation questions	.189	**.777**	-.024	.012	.64
Confirmations	-.270	**.739**	-.112	-.066	.64
Evaluations	.141	-.146	**.831**	.007	.73
Interactive verbalizations	-.245	.020	**.846**	-.061	.78
Wh- questions	-.411	-.199	-.183	**-.774**	.84
Labels	-.179	-.190	-.198	**.903**	.92

Regression analyses. Hierarchical multiple regression analyses were conducted first on aggregate outcome variables in the novel word learning and generalization tasks, then on variables representing individual tasks or testing times. Aggregate outcome variables were computed by converting children's performance on each measure into z-scores, then averaging the z-scores to create three dependent variables: target identification, forced choice generalization, and free choice generalization. In each analysis, the independent variables were entered in three blocks in order to see how different kinds of variables predicted outcome over and above those entered previously. Demographic and standardized test variables were entered in the first block. These variables included child age in months, child gender (dummy coded), vocabulary percentile score (averaged over the CDI-III and PPVT), and average screen use per day in minutes. The next block included a measure from the control word learning task, children's accuracy on target identification. Finally, the four co-viewing components from the factor analysis were entered together in the third block of the hierarchical regression. This method of analysis allowed for evaluating the predictive value of the co-viewing components over and above the other included variables (Tabachnick & Fidell, 2007). The results will be presented in sections for each type of novel word learning task included.

Target identification. First, an analysis was conducted using a normalized aggregate score to capture overall forced choice identification performance. The dependent variable for this analysis was the averaged z-score of children's target identification accuracy across both visits. Table 8 displays unstandardized and standardized regression coefficients after entry of all independent variable blocks. Although the overall model was not significant ($R^2 = .20$, $F(9, 34) = 0.95$, $p > .05$ after entry of all independent variable blocks), target identification was marginally predicted by parents' use of label elicitation and feedback during learning. This component showed a positive relationship with target identification, suggesting that the more parents asked children to produce novel labels and provided positive feedback, the better children learned and retained novel word-object mappings.

Table 8. Multiple regression output for the forced choice target identification task (aggregate z-score).

	Unstandardized Coefficients		Standardized Coefficients
Predictors	B	SE	Beta
Age	-.032	.102	-.055
Gender	-.087	.254	-.057
Vocabulary Percentile	-.002	.006	-.054
Screen Time	-.001	.003	-.048
Control Task Measure			
Target ID	.501	.542	.198
PCA Components			
Describing Objects	-.019	.124	-.026
Label Elicitation & Feedback	.274	.138	$.362^{\dagger}$
Narrating	-.201	.124	-.135
Open-Ended Questions (-) vs. Explicit Labeling (+)	.114	.124	.150

Note. $^{\dagger}p \leq .10$. $^{*}p \leq .05$.

Follow-up regression analyses were conducted on children's target identification accuracy separately for each visit (see Table 9). None of the included variables significantly predicted immediate target identification accuracy ($R^2 = .17$, $F(9, 34) = 0.79$, $p > .05$ after entry of all independent variable blocks). Although the overall model for delayed target identification accuracy did not reach significance ($R^2 = .23$, $F(9, 34) = 1.31$, $p > .05$), delayed accuracy was significantly predicted by parent's label elicitation and feedback. This result confirms the outcome of the above analysis using the aggregate measure, and suggests that parents' use of this type of speech was particularly associated with retention of the novel word-object mappings. Additionally, the parental co-viewing component of narrating was marginally predictive of delayed target identification accuracy. This suggests that the extent to which parents talked about the video itself and the content in a general way was associated with worse performance in children's retained novel word-object mappings.

Table 9. Multiple regression output for the forced choice target identification task at immediate and delayed testing.

	Immediate Target Identification			Delayed Target Identification		
	Unstandardized Coefficients		Standardized Coefficients	Unstandardized Coefficients		Standardized Coefficients
Predictors	B	SE	Beta	B	SE	Beta
Age	-.016	.030	-.096	.002	.026	.011
Gender	-.094	.074	-.218	.050	.064	.129
Vocabulary Percentile	-.002	.002	-.231	.001	.002	.148
Screen Time	.000	.001	.036	.000	.001	-.110
Control Task Measure						
Target ID	.196	.158	.271	.024	.137	.037
PCA Components						
Describing Objects	.014	.036	.065	-.020	.031	-.105
Label Elicitation & Feedback	.042	.040	.196	.071	.035	.366*
Narrating	.012	.036	.055	-.052	.031	-.264†
Open-Ended Questions (-) vs. Explicit Labeling (+)	.035	.036	.161	.014	.031	.073

Note. †$p \leq .10$. *$p \leq .05$.

Forced choice generalization. The first regression analysis was conducted on the average of the z-scores of children's accuracy in identifying shape matches and children's preference for shape on the two forced choice generalization tasks administered at delayed testing. Although the overall model was not significant ($R^2 = .20$, $F(9, 34) = 0.98$, $p > .05$ after entry of all independent variable blocks), forced choice generalization was predicted by parents' use of narrating utterances during co-viewing (see Table 10). Parental speech during co-viewing that focused broadly about the content of the video and the media itself was negatively associated with children's performance on the delayed forced choice generalization tasks.

Table 10. Multiple regression output for the forced choice generalization tasks (aggregate z-score).

	Unstandardized Coefficients		Standardized Coefficients
Predictors	B	SE	Beta
Age	.036	.025	.250
Gender	.071	.063	.190
Vocabulary Percentile	.001	.002	.096
Screen Time	.000	.001	-.054
Control Task Measure			
Target ID	.075	.134	.120
PCA Components			
Describing Objects	-.036	.031	-.192
Label Elicitation & Feedback	.044	.034	.235
Narrating	-.062	.031	-.328*
Open-Ended Questions (-) vs. Explicit Labeling (+)	.001	.031	-.007

Note. †$p \leq .10$. *$p \leq .05$.

Follow up analyses conducted separately on each of the forced choice generalization measures confirmed and further clarified the above result (see Table 11). None of the included independent variables significantly predicted children's identification of shape match objects ($R^2 = .13$, $F(9, 34) = 0.54$, $p > .05$ after entry of all independent variable blocks). The full model for children's preference for shape over material matches did not reach significance ($R^2 = .20$, $F(9, 34) = 0.94$, $p > .05$). However, similarly to the aggregate analysis above, the narrating component was a significant negative predictor of children's preference for shape in the forced choice generalization task. This shows that the above result on the aggregate measure was driven by the relationship between parents' narrating characteristics during co-viewing and children's preference for shape matches in the constrained test of generalization. Additionally, age was marginally predictive of this measure as well, suggesting that children more strongly prefer shape matches with age.

Table 11. Multiple regression output for the forced choice generalization tasks (both administered at delayed testing).

	Forced Choice Shape vs. Shape			Forced Choice Shape vs. Material		
	Unstandardized Coefficients		Standardized Coefficients	Unstandardized Coefficients		Standardized Coefficients
Predictors	B	SE	Beta	B	SE	Beta
Age	.007	.030	.041	.065	.037	.314†
Gender	.077	.074	.183	.066	.091	.121
Vocabulary Percentile	.000	.002	.028	.001	.002	.110
Screen Time	-.001	.001	-.119	.000	.001	.018
Control Task Measure						
Target ID	.123	.157	.175	.027	.194	.030
PCA Components						
Describing Objects	-.024	.036	-.113	-.048	.045	-.177
Label Elicitation & Feedback	.052	.040	.249	.036	.050	.132
Narrating	-.034	.036	-.161	-.089	.045	-.329*
Open-Ended Questions (-) vs. Explicit Labeling (+)	-.027	.036	-.126	.029	.044	.108

Note. $^{\dagger}p \leq .10$. $^{*}p \leq .05$.

Free choice generalization. To attain an aggregate measure of free choice generalization performance, children's choices of shape match objects, material match objects, and non-match distractor objects were z-score normalized and averaged together across visits. Children's material match and distractor non-match choices were reverse coded to capture the extent to which children rejected these types of objects at test. Therefore the aggregate measure captures the extent to which children chose one type of object (shape matches) and discriminated any other type of object (material matches and distractors). Using this aggregate measure averaged over both immediate and delayed testing yielded a non-significant model overall (R^2 = .21, F(9, 34) = 1.02, p > .05 after entry of all independent variable blocks) with no significant predictor variables (see Table 12).

Table 12. Multiple regression output for the free choice generalization task (aggregate z-score).

	Unstandardized Coefficients		Standardized Coefficients
Predictors	B	SE	Beta
Age	.067	.062	.190
Gender	.029	.155	.032
Vocabulary Percentile	.002	.004	.108
Screen Time	-.002	.002	-210
Control Task Measure			
Target ID	.302	.331	.194
PCA Components			
Describing Objects	-.051	.076	-.109
Label Elicitation & Feedback	.046	.084	.099
Narrating	-.057	.076	-.122
Open-Ended Questions (-) vs. Explicit Labeling (+)	.118	.076	.253

Note. †$p \leq .10$. *$p \leq .05$.

Follow-up analyses were conducted on children's free choice generalization performance using the same aggregate scores but separated by visit. Again, this measure captures how much children chose shape match objects and rejected both material match and distractor non-match objects at each testing time. The overall model for immediate generalization performance did not reach significance (R^2 = .19, F(9, 34) = 0.90, p > .05 after entry of all independent variable blocks) and showed no significant predictor variables (see Table 13).

Table 13. Multiple regression output for the free choice generalization task, separated by visit (aggregate z-scores).

	Immediate Free Choice Generalization			Delayed Free Choice Generalization		
	Unstandardized Coefficients		Standardized Coefficients	Unstandardized Coefficients		Standardized Coefficients
Predictors	B	SE	Beta	B	SE	Beta
Age	.184	.133	.247	.016	.148	.019
Gender	-.055	.331	-.028	.111	.369	.050
Vocabulary Percentile	.006	.008	.152	.001	.009	.028
Screen Time	-.003	.004	-.137	-.006	.004	-.232
Control Task Measure						
Target ID	.300	.705	.092	1.104	.785	.296
PCA Components						
Describing Objects	-.113	.162	-.115	-.002	.180	-.001
Label Elicitation & Feedback	.098	.180	.100	.011	.200	.009
Narrating	-.015	.162	-.015	-.264	.180	-.236
Open-Ended Questions (-) vs. Explicit Labeling (+)	.258	.161	.261	.209	.179	.186

Note. $^{\dagger}p \leq .10$. $^{*}p \leq .05$.

The overall model for delayed generalization performance also did not reach significance ($R^2 = .23$, $F(9, 34) = 1.13$, $p > .05$ after entry of all independent variable blocks) but was significantly improved with the addition of the control word learning measure ($F_{change}(1, 38) = 4.45$, $p < .05$). Before the addition of the co-viewing component predictor variables, children's target identification accuracy in the control word learning task was a significant positive predictor of delayed novel word generalization performance ($b = 1.46$, $t(36) = 2.11$, $p < .05$). That is, the better children learned real but unfamiliar words, the more they tended to extend novel words exclusively to shape match objects and not to material match or distractor non-match objects. However, after the addition of the co-viewing components to the model, none of the predictors reached significance.

Discussion

The results of Study 3 showed that although children as a group effectively learned novel words and categories from a video presentation, certain aspects of parental speech were associated with differences in this learning. An exploratory factor analysis revealed several variables of parent speech that characterized the co-viewing linguistic environment and that predicted children's learning outcomes.

First, the extent to which parents elicited labels from their children and provided feedback while watching the training video predicted children's retention of word-object mappings. While it is not necessarily surprising that labeling predicted children's learning of word-object mappings, the specific form of this labeling is informative. Children's retention was predicted not just by hearing parents label the items on screen, but by parents cuing the children to produce the labels themselves. This interactive label elicitation went together with responsive feedback from parents, often in the form of confirming the labels that children produced while watching the video. This finding fits with prior work indicating that responsive behavior during co-viewing promotes children's attention to and learning from screen media (Barr et al., 2008; Fender et al., 2010), specifically showing that responsiveness in the form of confirmations is associated with word learning. This results also builds on prior work showing a link between parental label use and children's word production during co-viewing (Fender et al., 2010). Study 3 shows that a specific way of focusing on labels, using elicitation questions, predicted word learning. Overall, the use of interactive questions focused on label learning along with responsive feedback was related to children better remember newly learned words.

Another key result was that the extent of narrating that parents did during co-viewing was negatively associated with both retained word-object mappings and children's preference for shape over material match choices in the delayed constrained generalization test. This suggests that parent speech about the video itself or general, non-specific speech about what is shown on the screen may not be especially conducive to word or category retention. Perhaps when parents talk more about the screen media itself, children have a harder time focusing on the specific content that is important for retention. In other words, this result seems to partly capture a negative relationship with learning due to focusing on the form rather than the content of screen media. Yet the narrating co-viewing component also included parental speech that was about information on screen, but that was broad and non-specific. These results indicate that general evaluative speech about content on screen is not much more informative than talking about the video itself, and both of these together actually have a negative association with learning outcomes in two types of tasks.

An unexpected result to emerge from Study 3 was the fact that none of the included variables predicted free choice generalization performance. Only the measure of accuracy in the control target identification task came close, predicting delayed free choice generalization before the co-viewing components were added to the model. This suggests that the control word learning task may relate to children's ability to generalize novel words, but does not predict performance over and above any of the co-viewing components identified in this study. However none of the co-viewing components were predictive of free choice generalization either. It is possible that other characteristics of co-viewing that were not captured in this analysis may be more strongly associated with novel free choice generalization performance. Some ideas of other aspects of co-viewing to investigate will be discussed shortly.

An unexpected outcome of Study 3 was the fact that two of the co-viewing components, describing objects and open-ended questions vs. explicit labeling, did not predict any of the outcome variables in the study. The describing objects component seemed particularly apt to test the prediction that parent speech focusing specifically on the objects to be learned would be associated with children's generalization performance. A possible next step to explore this issue would be to take a closer look at parents' descriptive speech. Perhaps parents tended to describe features that were not relevant to the generalization task included here (e.g., talking about the color of objects rather than the shape). The fact that the component that included explicit labeling did not predict learning outcome may simply be due to the result outlined above in which questions focusing on the labels were linked to retained word-object mapping. It may be the case that not just any speech focusing on labels is associated with word learning, but that the form of that speech matters.

The results of this study represent an important first step in linking specific co-viewing behaviors to children's learning outcomes in the context of screen media. There are various ways to refine and build on this work. One future direction would be to incorporate measures of child behavior and parent-child interaction quality in the kinds of analyses presented here. The relative timing of utterances and responses between a child and parent may be predictive of word learning. For example, children may learn most effectively when parents respond promptly and provide information about the item that is the focus of the child's attention in that moment. A closer look at parent-child interactions could also help further investigate the role of label elicitation questions suggested in the current results. The current study showed a beneficial association between parents' label elicitations and children's word learning, but did not actually measure whether and how children responded to those elicitations. Label elicitations may have been helpful because they prompted children to produce the novel nouns, suggesting important links between speech production and language development. On the other hand, perhaps label elicitations were still helpful even if children did not produce the target word. It may be the case that activating a mental representation of a word form is sufficient to support learning. This question could be further investigated by focusing on the interactions and contingencies between parent and child speech during co-viewing. This kind of analysis would also resonate with research on the social aspects of screen mediated learning. Although social information was not manipulated directly in video training, it could be informative to test how different extents of social contingency in co-viewing link to learning outcomes.

Another future direction for this work would be to guide experimental investigations of co-viewing. The kinds of analyses presented here can be used to develop experimental manipulations of the linguistic environment surrounding screen media co-viewing. For example, the current results suggest that the specific way in which parents focus on labels during co-viewing associate with word and category learning in different ways. The type and extent of labeling, such as whether labels are provided to children or elicited from them and how many labels are used during co-viewing, could be manipulated in an experimental design. This would allow for greater control of other characteristics of the co-viewing context, randomized assignment of children to conditions, and causal conclusions about the role of labels in screen mediated word and category learning. Similar experiments could be designed to test the effects of specific, content-focused speech compared to broad, screen-focused speech during co-viewing. In sum, Study 3 represents a novel approach to the question of how parental speech during co-viewing is associated with children's in-the-moment, screen mediated learning. This study links naturalistic, fine-grained parental co-viewing speech characteristics to children's learning, retention, and generalization outcomes in a word learning task. The results reveal that certain characteristics of parental speech are positively associated with different aspects of children's word learning, and others are negatively associated, and have the potential to inform future research on co-viewing and screen mediated learning.

Chapter 7: General Discussion

In the current studies, I set out to investigate how toddlers learn words and categories from screen media. Prior work on screen mediated learning in young children has shown an early video deficit effect which diminishes, at least in the domain of word learning, by about 2½ years of age (Krcmar, 2010; O'Doherty et al., 2011). In comparing how children learned novel words from face-to-face and from screen-mediated training, my results were consistent with the current literature—as long as children were only tested on one-to-one word-object mappings. An exploration of how children generalize novel words to other novel objects—that is, how they learn novel categories—revealed another manifestation of the video deficit effect among 2½- to 3-year-olds. In the studies presented here I further investigated this video deficit in retained category coherence, looking at how the information available in the screen mediated learning environment impacts learning, and how parent-child interactions during co-viewing are associated with word and category learning.

Studies 1a and 1b explored the characteristics of 2½- to 3-year-olds' word and category learning from screen media compared to learning directly from a person. The results of Study 1a showed that children could accurately learn novel word-object mappings from either a video or in person presentation. Children also retained these mappings over a one week delay. Children in both training conditions were also able to generalize newly learned words to different extents across different kinds of test objects. However, after the week-long delay generalization performance among children in the two training conditions diverged. Children trained in person retained a clear preference to extend words based on certain feature similarities. Children trained by video no longer generalized the novel words in a structured way, and in fact their performance suggested that they no longer differentiated between any of the different object types at test. Together these results suggest that although toddlers have overcome the video deficit in learning one-to-one word-referent mappings, there is still a video deficit in retained category coherence when inferring novel concepts.

Study 1b refined and extended the results of Study 1a by focusing on children's preference for inferring shape-based categories specifically. Using a different set of novel objects, the results of Study 1b first confirmed that 30- to 36-month-olds can learn and retain word-object mappings just as accurately when learned from a video as from an in person presentation. Study 1b also further confirmed that children inferred coherent, shape-based categories immediately following training either in person or by video. However, the results showed that children trained by video retained less coherent categories than children trained in person, extending words indiscriminately based on both shape and material. This pattern was observed despite also finding that children across both training conditions accurately identified shape match objects when asked for by name and even tended to prefer shape over material match objects in a delayed forced-choice test of generalization. Taken together, the results of Study 1b suggest that children can glean enough information from a video presentation to learn and retain word-object mappings, and retain a preference for shape-based categories in the context of constrained, two-alternative tests. However, the quality of object representations that children retain over time after screen mediated learning does not seem to support clear, coherent categorization, particularly when children are faced with the relatively more complex and difficult task of selectively generalizing a word to multiple objects.

Having established a video deficit among toddlers in the measure of retained category coherence, the next step was to explore characteristics of the screen mediated learning environment that contribute to and influence this pattern of learning. Study 2 involved the manipulation of perceptual and social information present in the screen mediated learning context. Across two new training conditions, and an additional control condition, different extents of perceptual and social information were added to the screen mediated word learning context. The results suggest that increasing perceptual information by adding physical objects to the video learning context may help children retain more coherent, shape-based categories. On the other hand, increasing social information by having a physically present experimenter contingently label objects shown in a video led to a similar deficit in retained category coherence as observed in Studies 1a and 1b. However, an interesting finding to come out of Study 2 was that having objects present while learning from a video only started to make a difference in retained category coherence when an experimenter placed each individual object in front of the child as they were learning about that object from the screen. This suggests that perceptual and social information work together to impact children's novel category retention. Specifically, rich perceptual information may be helpful for children as they encode the features of novel objects that are likely to be important, but social cues may also be helpful for guiding attention to pertinent information (i.e., object properties) at the right time (i.e., as those objects are being labeled).

Study 3 continued to explore the relationships between screen mediated learning and the environment in which it takes place by considering a common naturalistic setting for this kind of learning: parent-child co-viewing of a video. Parent-child dyads were observed as they watched a training video presenting the novel words and objects to be learned, and parental speech was later coded and analyzed in relation to children's learning, retention, and generalization performance. The results showed that certain characteristics of parents' speech during co-viewing predicted some of children's learning outcomes. Parents' elicitation of labels from their children, and resulting positive feedback, predicted children's retention of word-object mappings. On the other hand, parents' use of broad, non-specific speech about the training video and novel items tended to predict worse learning outcomes in both word-object mapping and category learning tasks.

The studies presented here add to the literature on screen mediated learning in young children as well as point to questions for future investigation. In the remainder of this chapter I will discuss both the contributions and implications of the current work and some future directions for related research. The current results add to what is known about the video deficit effect and have implications for theories of screen mediated learning and language development more broadly. Much work also remains to be done to characterize the nature of young children's word and especially category learning from screen media. I will discuss how the current studies can help inform directions for future research in this area, as well as how this work has direct implications for practice.

Contributions and Implications

Together, the studies presented here make several contributions to the current literature. They document a video deficit effect in a domain that has not been explored before in this context, category learning. The ability to infer novel categories and generalize newly learned words is an important part of language acquisition and cognitive development. The fact that this ability is affected by screen mediated learning adds another piece to our understanding of what children can learn from screen media, and what they struggle with. The current studies also show that this video deficit in category learning is present in an older age group than has typically been shown to be susceptible to video deficit effects. This is particularly the case in the domain of word learning, in which the video deficit has been shown to diminish by the third year of life

(Krcmar, 2010; O'Doherty et al., 2011). This result underscores the importance of further investigating children's screen mediated learning in various domains of cognitive development, and in various experimental tasks. Using tasks that are already well understood in the context of face-to-face learning to investigate the video deficit effect will help further chart when and how children struggle with learning from a screen.

The current studies also have implications for theories of the video deficit effect. The results of Study 2 show that adding social information back to the screen mediated learning context was not sufficient to bring learning to the same level as in person training. Further, although perceptual information seemed to be somewhat helpful, the results also suggested that additional attentional cues were necessary to ameliorate the video deficit effect. This pattern of results does not definitively support either a perceptual or social account of the video deficit effect. Instead, the nuanced nature of these results reflects difficulty with a strong dichotomy between perceptual and social accounts of the video deficit. Perhaps separating the social and perceptual aspects of screen mediated learning is more useful in thinking about the learning task itself rather than theoretical accounts. Instead, considering the contributions of both social and perceptual information to screen mediated learning may be a better way forward toward a cohesive theory of the video deficit effect.

It may be beneficial in future work on the mechanisms of the video deficit effect to consider how and why more general learning mechanisms operate differently on the particular kinds of information that are conveyed through a screen. As has been discussed in the literature, and indicated in the results of Study 2, social and perceptual factors both play a role in young children's screen mediated learning. These factors may continue to provide a useful way to characterize the differences in information conveyed through a screen compared to face-to-face interactions. Thinking about screen mediated learning at this level may facilitate further investigations into how the quality of the information source may impact encoding, or how the demands of transferring between different contexts may impact retrieval of information learned from a screen. That is, a better approach may be one that moves away from the social/perceptual dichotomy, at the theory level at least, and instead focuses on more basic learning mechanisms applied to the unique quality of information presented in screen media.

The work presented here also adds to evidence indicating that although the video deficit effect has been observed across several domains of learning, performance seems to be impacted in different ways depending on the particular domain or task of interest. For example, in the current work toddlers were able to learn novel word-object mappings as well from a screen as from a person, but showed a video deficit in retained category coherence. Theoretical accounts of the video deficit effect must address the fact that the deficit is not global but instead seems to be task dependent. To better inform theory it would be useful to consider the particular characteristics or demands of various tasks, and how they are altered by screen mediated presentations. For example, the ability to infer coherent, shape-based categories may rely on encoding a perceptually rich representation more so than does the task of mapping a single word to a single object. Better understanding which aspects of various tasks are impacted by the screen mediated context can help contribute to a general account of the video deficit effect.

The current results also relate to theories of word learning in that they reflect on what information is sufficient for two aspects of learning in this domain. First, the results show that by the age of 2½ years toddlers can accurately learn novel, one-to-one word-object mappings just from watching a video. By this age, children seem to be such skilled word learners that they can learn and retain basic word-object mappings from a relatively impoverished source of information. This may indicate that the social convention of mapping a novel word to a novel object is so familiar to toddlers by this age that they can recognize and learn from a labeling act delivered by a non-interactive, 2D representation of a person. However, video alone is not sufficient to support toddler's retention of coherent lexical categories. Toddlers need rich perceptual information, and perhaps at least some subtle social cues to help guide attention, in order to form coherent and robust representations of inferred categories of novel objects. Interestingly, for this age group, information that is sufficient for one-to-one word-object mapping is not necessarily sufficient for the retention of coherent categories. This suggests that these are two distinct aspects of word learning with different developmental trajectories.

Finally, Study 3 makes novel contributions to the literature in a different way. The results of Study 3 add to what is known about how parents and children interact while co-viewing screen media. This study makes the novel contribution of linking parental speech during co-viewing to structured measures of children's word and category learning from screen media. Although Study 3 suggests that certain parental co-viewing behaviors predict children's learning from screen media, the design of the study does not allow for strong causal conclusions. This is because parents were allowed to co-view the video with their children in a relatively naturalistic, unstructured way. Further, there is not a strong basis to conclude that the co-viewing behaviors found to be predictive of learning in Study 3 would necessarily generalize to the population of toddlers as a whole. It is possible that parents in this sample used co-viewing behaviors that were individualized to their own children. Perhaps the results of Study 3 capture a sample of parents who are particularly in tune with the needs of their particular children in the context of watching a video. Perhaps some more fundamental quality of parent-child interaction, rather than type of co-viewing speech, is really causing differences in learning. As it is now, Study 3 does not allow me to fully rule out such alternative possibilities.

Although the results of Study 3 cannot be used to make causal conclusions about co-viewing and learning, they represent a novel approach to an empirical question that is still in the early stages of being investigated. Illuminating the links between adult-child co-viewing and young children's screen mediated learning has direct implications for future applied work. While the first group of studies included here provide new insights on the characteristics of toddlers' screen mediated learning and the video deficit effect, Study 3 has the most potential for guiding future applications that may help improve children's learning in this context. In the next section I will discuss several directions for future research based on the current findings, as well as how the current work connects to some specific applied uses of screen media with young children.

Future Directions

While the studies presented here add to the literature on screen mediated learning in young children, many questions remain to be investigated. A remaining issue that is fundamental to studies of screen mediated learning and the video deficit effect is transfer. Although transfer is not thought to be driving the current results (see Discussion section of Study 2), transfer is inherently a factor in how children learn from screen media. In naturalistic settings, for example watching television at home, transfer is often an integral part of learning from screen media. If children are to put concepts and lessons learned from a screen to use in real life, they must be able to effectively retain and transfer that information across time and context. Some have argued that the ability to transfer information learned from the screen mediated context to one's own environment develops with age and is a key factor underlying the video deficit effect (e.g., Barr, 2010). Following from this, a logical next step for research on facilitating young children's learning from screen media would be to investigate ways to reduce the transfer demands of this type of learning. Controlling for transfer of learning could also be useful for further testing and understanding the specific social and perceptual factors that are involved in screen mediated learning. Accounting for the effects of transfer could allow for more focused tests of social and perceptual information, and in turn help to better clarify the learning task that children face in the context of screen media.

In order to control for the context transfer that children typically experience going from training to testing in a screen mediated learning task, both parts of a task would need to be administered through a screen. Screen mediated testing could be implemented in several ways. For example, children could be asked to choose one or more test item images or videos on a touchscreen. Children could be asked questions about the screen at test either by a physically present experimenter or by a pre-recorded experimenter on video or voiceover. Socially interactive and contingent testing could be implemented through the use of videoconferencing. The combination of these testing methods paired with different screen mediated training designs would control for transfer but also allow for new tests of social contingency and perceptual information in language learning.

For example, one could manipulate the level of social contingency available at training by comparing a non-interactive training video and interacting with an experimenter through videoconferencing during training. Children in both conditions would then be tested through socially interactive videoconferencing, thus reducing transfer demands. If social contingency at the time of learning from a screen is key, then children would be expected to learn less well from the pre-recorded video compared to the videoconferencing training. To further explore the effect of perceptual information in screen mediated learning, one could manipulate whether or not physical objects are present as a child is trained by watching a non-interactive video. Testing in both conditions could be administered through pre-recorded questions on screen about the physical objects present in front of the child. Again this would reduce transfer demands as well as control for social contingency. If being able to see the physical objects to be learned about during training is beneficial, then children would be expected to learn better in the condition in which physical objects were present initially. These examples show how controlling for transfer may be an effective way to isolate other factors of the screen mediated learning context. Further, in combination with studies like those included in this dissertation, experiments that control for transfer can help clarify the role of context transfer itself in screen mediated learning.

The studies presented here suggest several specific ways to facilitate young children's word and concept learning from screen media. Study 3 has direct implications for practice in that the results demonstrate characteristics of co-viewing that are associated with children's word and category learning. Although possible causal explanations are limited, some general conclusions can be drawn from Study 3. For example, eliciting labels and giving positive feedback while co-viewing is positively related to children's learning and especially retention word-referent mappings. On the other hand, commenting on the screen media itself or talking broadly about what is on the screen is associated with inhibited lexical category learning. Experimental work on co-viewing is needed and the results of Study 3 will be useful in guiding this kind of research. The results of Study 3 can be used to design experimental manipulations and formulate hypotheses. For example, the results showed that parental label elicitations during co-viewing predicted children's retention of novel word-object mappings. A logical next step would be to design a study in which the ways that a co-viewing adult focused on labels during learning were manipulated. If this is a generalizable, causal relationship, then the result should be replicated in a sample of children randomly assigned to receive different types of labeling. Such an experiment could also help refine the nature of this relationship, for example by testing whether learning is also predicted by children's production of the labels during co-viewing. The results of Study 3 have the potential to help refine future studies of co-viewing and inform practice.

The findings of Study 3 have implications for the use of screen media in education. Screen media technology can be an important tool in the classroom, for example by tailoring content to individual students and providing extra, one-on-one time to work with that content. One area in which computer-based educational technology has been implemented and studied with children is in reading intervention programs (Wise et al., 2005). For example, Wise and colleagues describe a program that adapts to the needs of individual student users and provides structured practice with various reading skills, such as phonological encoding, word recognition, and passage comprehension. Another key component of this program is that the virtual tutor is interactive, appearing as an animated woman's face on the screen that exhibits realistic movements during speech. The approach used in Study 3 could have interesting implications for the development and testing of interactive, computer-based tutoring systems. It would be interesting to investigate whether children's interactions with parents while co-viewing are at all similar to children's interactions with a computer-based tutor. Perhaps findings on characteristics of parental speech that are associated with children's screen mediated learning could also help design effective strategies for tutoring systems.

The results of the current studies also have implications for therapeutic applications of screen media, such as speech-langauge therapy telepractice and video-based interventions for autism spectrum disorders. I will briefly discuss each of these applications in turn. Telepractice, or the delivery of speech-language therapy remotely through the use of communications technology, is a recently emerging area in the larger field of telemedicine (Dudding, 2009). Early diagnosis and intervention are important components of treatment in this field, for example in children with hearing loss, and some recent research has focused on the use of telepractice with young children. For example, videoconferencing has been investigated as a method for administering language disorder assessments to children (Waite, Theodoros, Russell, & Cahill, 2010) and connecting therapists, parents, and young children with hearing loss for therapy sessions (Constantinescu, 2012). Although such studies have shown some positive results, there are also concerns, particularly with treating very young children. Constantinescu (2012) surveyed therapists about a telepractice intervention program and found that almost half were dissatisfied with the level of engagement they attained with the children they were treating. This was particularly the case for therapists working with younger children in the study, including infants and toddlers. This result demonstrates why the study of young children’s learning from screen media could be useful for refining current approaches and developing new methods in telepractice.

Another therapeutic application of screen-mediated learning research can be seen in video intervention techniques for children with autism spectrum disorders (ASD). A growing literature of research indicates that video-based instruction is a useful method for teaching communication, social, and academic skills to children with ASD (Kagohara, 2010). This type of intervention involves having a child watch a video in which specific behaviors are modeled, such as a conversational exchange, with the goal of having that child imitate the target behavior later on. For example, one study investigated video-based modeling of several play behaviors with a 3-year-old child with ASD (D’Ateno, Mangiapanello, & Taylor, 2003). The child successfully imitated motor and verbal behaviors from the video, even without reinforcing feedback from an adult. Other research has shown that video-based modeling, along with feedback and reinforcement from a therapist, can help preschoolers with ASD learn and even generalize behaviors (e.g., Gena, Couloura, & Kymissis, 2005; Hine & Wolery, 2006). Interventions involving screen media may be particularly useful for children with ASD because they tap into strengths in visual learning while reducing potentially aversive social interactions (Kagohara, 2010). Research on screen mediated language learning has the potential to inform therapeutic approaches with children with ASD. At the same time, further research on video-based instruction with children with ASD may also inform theories about the mechanisms underlying the video deficit effect in language.

Research that adds to our understanding of how children learn from screen media, and how to improve that learning, has direct implications for these and other applications of screen mediated learning. In the case of telepractice, further work on how to promote social engagement through screen media would be particularly beneficial. Co-viewing may be an effective way to facilitate screen mediated interactions between a therapist and a young child. For example, a co-viewing parent may be able to act as a model of social engagement through videoconferencing and encourage their child to interact as well. In the case of video-based instruction for children with ASD, further research on retention and generalization of information learned from a screen would be beneficial. A key point of this kind of instruction is not only for children to learn the content of the video, but also to be able to apply and extend what they learn in other situations, such as in interactions with parents or peers. Some prior work has indicated certain conditions that seem to promote generalization of video-based learning (e.g., Hine & Wolery, 2006), and further research on screen mediated learning can help add to this. Future research on screen mediated learning in young children from a variety of populations will help educators and practitioners better harness the potential of screen-based technology.

Conclusion

Taken together, the studies presented here demonstrate some proficiencies but also some limits to 2½- to 3-year-old toddlers' word and category learning from screen media. These studies also suggest factors that may have the potential to help, for example additional perceptual information (perhaps with some social guidance) and responsive, informative co-viewing input (such as label elicitation questions and feedback). This research adds to what is known about how young children learn from screen media, and what can be done to facilitate this learning. This area of research is still in its' infancy, and so there is much to be done in future work. To better understand how children learn from screen media, why they sometimes experience a video deficit, and what can be done to facilitate learning, we need to continue to investigate various tasks and measures, and to look at learning on multiple timescales. The more informed we are about the scope of children's screen-mediated learning, the better we will be able to employ new technology for early learning. Integrating screen media into education, therapy, and day-to-day activities with young children should be guided by sound investigation. The work presented here represents several approaches to this kind of investigation in the domain of screen mediated word and category learning.

References

Allen, R., & Scofield, J. (2010). Word learning from videos: More evidence from 2-year-olds. *Infant and Child Development, 19*(6), 649-661. doi: 10.1002/icd.712

American Academy of Pediatrics. (2010). Media education. *Pediatrics, 126*(5), 1012-1017. doi: 10.1542/peds.2010-1636

Anderson, D. R., & Hanson, K. G. (2010). From blooming, buzzing confusion to media literacy: The early development of television viewing. *Developmental Review, 30*(2), 239-255. doi: 10.1016/j.dr.2010.03.004

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist, 48*(5), 505-522. doi: 10.1177/0002764204271506

Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated aggressive models. *The Journal of Abnormal and Social Psychology, 66*(1), 3-11. doi: 10.1037/h0048687

Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review, 30*(2), 128-154. doi: 10.1016/j.dr.2010.03.001

Barr, R., & Hayne, H. (1999). Developmental changes in imitation from television during infancy. *Child Development, 70*(5), 1067-1081. doi: 10.1111/1467-8624.00079

Barr, R., Muentener, P., & Garcia, A. (2007). Age-related changes in deferred imitation from television by 6- to 18-month-olds. *Developmental Science, 10*(6), 910-921. doi: 10.1111/j.1467-7687.2007.00641.x

Barr, R., Muentener, P., Garcia, A., Fujimoto, M., & Chávez, V. (2007). The effect of repetition on imitation from television during infancy. *Developmental Psychobiology, 49*(2), 196-207. doi: 10.1002/dev.20208

Barr, R., & Wyss, N. (2008). Reenactment of televised content by 2-year olds: Toddlers use language learned from television to solve a difficult imitation problem. *Infant Behavior & Development, 31*(4), 696-703. doi: 10.1016/j.infbeh.2008.04.006

Barr, R., Zack, E., Garcia, A., & Muentener, P. (2008). Infants' attention and responsiveness to television increases with prior exposure and parental interaction. *Infancy, 13*(1), 30-56. doi: 10.1080/15250000701779378

Barr, R., Zack, E., Garcia, A., & Muentener, P. (2008). Infants' attention and responsiveness to television increases with prior exposure and parental interaction. *Infancy, 13*(1), 30-56. doi: 10.1080/15250000701779378

Booth, A. E., Waxman, S. R., & Huang, Y. T. (2005). Conceptual information permeates word learning in infancy. *Developmental Psychology, 41*(3), 491.

Booth, A. E., & Waxman, S. (2002). Word learning is ‘smart’: Evidence that conceptual information affects preschoolers’ extension of novel words. *Cognition, 84*(1), B11-B22. doi: 10.1016/s0010-0277(02)00015-x

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science, 27*(6), 843-873. doi: 10.1016/j.cogsci.2003.06.001

Carey, S. (1978).The child as word learner. In J. Bresnan, G. Miller and M. Halle (eds.),Linguistic Theory and Psychological Reality. Cambridge, MA: MIT Press, 264-293.

Carver, L. J., Meltzoff, A. N., & Dawson, G. (2006). Event-related potential (ERP) indices of infants' recognition of familiar and unfamiliar objects in two and three dimensions. *Developmental Science, 9*(1), 51-62. doi: 10.1111/j.1467-7687.2005.00463.x

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38*(6), 967-978. doi: 10.1037/0012-1649.38.6.967

Collins, W. A., Sobol, B. L., & Westby, S. (1981). Effects of adult commentary on children's comprehension and inferences about a televised aggressive portrayal. *Child Development, 52*(1), 158-163. doi: 10.2307/1129225

Colunga, E., & Sims, C.E. (2012). Early-talker and late-talker toddlers and networks show different word learning biases. In N. Miyake, D. Peebles, & R.P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 246-251). Austin, TX: Cognitive Science Society.

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review, 112*(2), 347-382. doi: 10.1037/0033-295X.112.2.347

Constantinescu, G. (2012). Satisfaction with telemedicine for teaching listening and spoken language to children with hearing loss. *Journal of Telemedicine and Telecare, 18*(5), 267-272. doi: 10.1258/jtt.2012.111208

Courage, M. L., Murphy, A. N., Goulding, S., & Setliff, A. E. (2010). When the television is on: The impact of infant-directed video on 6- and 18-month-olds’ attention during toy play and on parent–infant interaction. *Infant Behavior & Development, 33*(2), 176-188. doi: 10.1016/j.infbeh.2009.12.012

D'Ateno, P., Mangiapanello, K., & Taylor, B. A. (2003). Using video modeling to teach complex play sequences to a preschooler with autism. *Journal of Positive Behavior Interventions, 5*(1), 5-11. doi: 10.1177/10983007030050010801

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development, 62*(4), 736-752. doi: 10.1111/j.1467-8624.1991.tb01566.x

DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., . . . O'Doherty, K. (2010). Do babies learn from baby media? *Psychological Science, 21*(11), 1570-1574. doi: 10.1177/0956797610384145

Demers, L. B., Hanson, K. G., Kirkorian, H. L., Pempek, T. A., & Anderson, D. R. (2012). Infant gaze following during Parent–Infant coviewing of baby videos. *Child Development, 84*(2), 591-603. doi: 10.1111/j.1467-8624.2012.01868.x

Deocampo, J. A., & Hudson, J. A. (2005). When seeing is not believing: Two-year-olds' use of video representations to find a hidden toy. *Journal of Cognition and Development, 6*(2), 229-258. doi: 10.1207/s15327647jcd0602_4

Diener, M. L., Pierroutsakos, S. L., Troseth, G. L., & Roberts, A. (2008). Video versus reality: Infants' attention and affective responses to video and live presentations. *Media Psychology, 11*(3), 418-441. doi: 10.1080/15213260802103003

Dollaghan, C. (1985). Child meets word: "fast mapping" in preschool children. *Journal of Speech & Hearing Research, 28*(3), 449-454.

Dudding, C. C. (2009). Digital videoconferencing applications across the disciplines. *Communication Disorders Quarterly, 30*(3), 178-182. doi: 10.1177/1525740108327449

Dunn, L., & Dunn, D. (2007). *PPVT-4 manual*. Bloomington, MN: NCS Pearson, Inc.

Fender, J. G., Richert, R. A., Robb, M. B., & Wartella, E. (2010). Parent teaching focus and toddlers' learning from an infant DVD. *Infant and Child Development, 19*(6), 613-627. doi: 10.1002/icd.713

Fenson, L. (2007). *MacArthur-bates communicative development inventories: User's guide and technical manual.* Paul H. Brookes Publishing Company.

Fenstermacher, S. K., Barr, R., Salerno, K., Garcia, A., Shwery, C. E., Calvert, S. L., & Linebarger, D. L. (2010). Infant-directed media: An analysis of product information and claims. *Infant and Child Development, 19*(6), 557-556. doi: 10.1002/icd.718

Fidler, A. E., Zack, E., & Barr, R. (2010). Television viewing patterns in 6- to 18-month-olds: The role of caregiver–infant interactional quality. *Infancy, 15*(2), 176-196. doi: 10.1111/j.1532-7078.2009.00013.x

Fisch, S. M., Akerman, A., Morgenlander, M., McCann Brown, S. K., Fisch, S. R. D., Schwartz, B. B., & Tobin, P. (2008). Coviewing preschool television in the us. *Journal of Children & Media, 2*(2), 163-173. doi: 10.1080/17482790802078680

Gelman, S. A., & Wellman, H. M. (1991). Insides and essence: Early understandings of the non-obvious. *Cognition, 38*(3), 213-244. doi: 10.1016/0010-0277(91)90007-Q

Gena, A., Couloura, S., & Kymissis, E. (2005). Modifying the affective behavior of preschoolers with autism using *in-vivo* or video modeling and reinforcement contingencies. *Journal of Autism and Developmental Disorders, 35*(5), 545-556. doi: 10.1007/s10803-005-0014-9

Gershkoff-Stowe, L., & Smith, L. B. (2004). Shape and the first hundred nouns. *Child Development, 75*(4), 1098-1114. doi: 10.1111/j.1467-8624.2004.00728.x

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language, 17*(1), 171-183. doi: 10.1017/S0305000900013167

Gopnik, A., & Nazzi, T. (2003). Words, kinds, and causal powers: A theory theory perspective on early naming and categorization. *Early category and concept development: Making sense* *of the blooming, buzzing confusion* (pp. 303-329). New York, NY, US: Oxford University Press, New York, NY.

Hains, S. M. J., & Muir, D. W. (1996). Effects of stimulus contingency in infant–adult interactions. *Infant Behavior & Development, 19*(1), 49-61. doi: 10.1016/S0163-6383(96)90043-0

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*(1), B23-B34. doi: 10.1016/S0010-0277(02)00186-5

Hayne, H., Herbert, J., & Simcock, G. (2003). Imitation from television by 24- and 30-month-olds. *Developmental Science, 6*(3), 254-261. doi: 10.1111/1467-7687.00281

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58*(4), 1021-1034. doi: 10.2307/1130543

Hine, J. F., & Wolery, M. (2006). Using point-of-view video modeling to teach play to preschoolers with autism. *Topics in Early Childhood Special Education, 26*(2), 83-93. doi: 10.1177/02711214060260020301

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*(2), 128-157. doi: 10.1080/15250000701795598

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development, 8*(2), 113-139. doi: 10.1016/0885-2014(93)90008-S

Jones, S. S., & Smith, L. B. (2002). How children know the relevant properties for generalizing object names. *Developmental Science, 5*(2), 219-232. doi: 10.1111/1467-7687.00224

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development, 62*(3), 499-516. doi: 10.2307/1131126

Kagohara, D. M. (2010). Is video-based instruction effective in the rehabilitation of children with autism spectrum disorders? *Developmental Neurorehabilitation, 13*(2), 129-140. doi: 10.3109/17518420903329281

Kavanaugh, R. D., & Jirkovsky, A. M. (1982). Parental speech to young children: A longitudinal analysis. *Merrill-Palmer Quarterly, 28*(2), 297-311.

Kemler Nelson, D. G., Russell, R., Duke, N., & Jones, K. (2000). Two-year-olds will name artifacts by their functions. *Child Development, 71*(5), 1271-1288. doi: 10.1111/1467-8624.00228

Kirkorian, H. L., Pempek, T. A., Murphy, L. A., Schmidt, M. E., & Anderson, D. R. (2009). The impact of background television on parent-child interaction. *Child Development, 80*(5), 1350-1359. doi: 10.1111/cdev.2009.80.issue-510.1111/j.1467-8624.2009.01337.x

Klein, A. M., Hauf, P., & Aschersleben, G. (2006). The role of action effects in 12-month-olds' action control: A comparison of televised model and live model. *Infant Behavior & Development, 29*(4), 535-544. doi: 10.1016/j.infbeh.2006.07.001

Krcmar, M. (2010). Can social meaningfulness and repeat exposure help infants and toddlers overcome the video deficit? *Media Psychology, 13*(1), 31-53. doi: 10.1080/15213260903562917

Krcmar, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? an experimental approach. *Media Psychology, 10*(1), 41-63. doi: 10.1080/15213260701300931

Kuhl, P. K., Tsao, F., & Liu, H. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100*(15), 9096-9101. doi: 10.1073/pnas.1532872100

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299-321. doi:10.1016/0885-2014(88)90014-7

Lauricella, A. R., Pempek, T. A., Barr, R., & Calvert, S. L. (2010). Contingent computer interactions for young children's object retrieval success. *Journal of Applied Developmental Psychology, 31*(5), 362-369. doi: 10.1016/j.appdev.2010.06.002

Liittschwager, J. C., & Markman, E. M. (1994). Sixteen- and 24-month-olds' use of mutual exclusivity as a default assumption in second-label learning. *Developmental Psychology, 30*(6), 955-968. doi: 10.1037/0012-1649.30.6.955

Linebarger, D. L., & Walker, D. (2005). Infants' and toddlers' television viewing and language outcomes. *American Behavioral Scientist, 48*(5), 624-645. doi: 10.1177/0002764204271505

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science, 18*(12), 1077-1083. doi: 10.1111/j.1467-9280.2007.02028.x

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16*(1), 1-27. doi: 10.1016/0010-0285(84)90002-1

McCall, R. B., Parke, R. D., & Kavanaugh, R. D. (1977). Imitation of live and televised models by children one to three years of age. *Monographs of the Society for Research in Child Development, 42*, 94-94. doi: 10.2307/1165913

McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. *Cognitive Development, 22*(3), 353-364. doi: 10.1016/j.cogdev.2007.01.001

Meltzoff, A. N. (1988). Imitation of televised models by infants. *Child Development, 59*(5), 1221-1229. doi: 10.2307/1130485

Nathanson, A. I., & Rasmussen, E. E. (2011). TV viewing compared to book reading and toy playing reduces responsive maternal communication with toddlers and preschoolers. *Human Communication Research, 37*(4), 465-487. doi: 10.1111/j.1468-2958.2011.01413.x

Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science, 11*(5), 722-731. doi: 10.1111/j.1467-7687.2008.00722.x

O'Doherty, K., Troseth, G. L., Shimpi, P. M., Goldenberg, E., Akhtar, N., & Saylor, M. M. (2011). Third-party social interaction and word learning from video. *Child Development, 82*(3), 902-915. doi: 10.1111/j.1467-8624.2011.01579.x

Okuma, K., & Tanimura, M. (2009). A preliminary study on the relationship between characteristics of TV content and delayed speech development in young children. *Infant Behavior & Development, 32*(3), 312-321. doi: 10.1016/j.infbeh.2009.04.002

Pempek, T. A., Demers, L. B., Hanson, K. G., Kirkorian, H. L., & Anderson, D. R. (2011). The impact of infant-directed videos on parent–child interaction. *Journal of Applied Developmental Psychology, 32*(1), 10-19. doi: 10.1016/j.appdev.2010.10.001

Reiser, R. A., Williamson, N., & Suzuki, K. (1988). Using “Sesame street” to facilitate children’s recognition of letters and numbers. *Educational Technology Research and Development, 36*(1), 15-21.

Reiser, R. A., Tessmer, M. A., & Phelps, P. C. (1984). Adult–child interaction in children's learning from "sesame street.". *Educational Communication & Technology Journal, 32*(4), 217-223.

Rice, M. L., Huston, A. C., Truglio, R., & Wright, J. C. (1990). Words from "sesame street": Learning vocabulary while viewing. *Developmental Psychology, 26*(3), 421-428. doi: 10.1037/0012-1649.26.3.421

Rice, M. L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development, 59*(2), 420-429. doi: 10.2307/1130321

Richert, R. A., Robb, M. B., & Smith, E. I. (2011). Media as social partners: The social nature of young children's learning from screen media. *Child Development, 82*(1), 82-95. doi: 10.1111/j.1467-8624.2010.01542.x

Rideout, V. J., Hamel, E., & Henry J. Kaiser Family Foundation. (2006). *The media family: Electronic media in the lives of infants, toddlers, preschoolers and their parents* Henry J. Kaiser Family Foundation.

Rondal, J. A. (1980). Fathers' and mothers' speech in early language development. *Journal of Child Language, 7*(02), 353. doi: 10.1017/S0305000900002671

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. doi: 10.1111/j.1467-8624.2009.01338.x

Schmidt, M. E., Crawley-Davis, A. M., & Anderson, D. R. (2007). Two-year-olds' object retrieval based on television: Testing a perceptual account. *Media Psychology, 9*(2), 389-409. doi: 10.1080/15213260701291346

Schmitt, K. L., & Anderson, D. R. (2002). Television and reality: Toddlers' use of visual information from video to guide behavior. *Media Psychology, 4*(1), 51-76. doi: 10.1207/S1532785XMEP0401_03

Scofield, J., & Williams, A. (2009). Do 2-year-olds disambiguate and extend words learned from video? *First Language, 29*(2), 228-240. doi: 10.1177/0142723708101681

Scofield, J., Williams, A., & Behrend, D. A. (2007). Word learning in the absence of a speaker. *First Language, 27*(3), 297-311. doi: 10.1177/0142723707077252

Sims, C.E., Schilling, S.M., & Colunga, E. (2012). Interactions in the development of skilled word learning in neural networks and toddlers. *Proceedings of the IEEE International Conference on Development and Learning/Epigenetic Robotics* (pp. 1-6). San Diego, CA.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*(1), 13-19. doi: 10.1111/1467-9280.00403

Soja, N. N. (1992). Inferences about the meanings of nouns: The relationship between perception and syntax. *Cognitive Development, 7*(1), 29-45. doi: 10.1016/0885-2014(92)90003-A

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition, 38*(2), 179-211. doi: 10.1016/0010-0277(91)90051-5

Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology, 109*(1), 132-140. doi: 10.1016/j.jecp.2010.10.013

Stoneman, Z., & Brody, G. H. (1982). An in-home investigation of maternal teaching strategies during sesame street and a popular situation comedy. *Journal of Applied Developmental Psychology, 3*(3), 275-284. doi: 10.1016/0193-3973(82)90020-X

Strouse, G. A., & Troseth, G. L. (2008). "Don't try this at home": Toddlers' imitation of new skills from people on video. *Journal of Experimental Child Psychology, 101*(4), 262-280. doi: 10.1016/j.jecp.2008.05.010

Suddendorf, T. (2003). Early representational insight: Twenty-four-month-olds can use a photo to find an object in the world. *Child Development, 74*(3), 896-904. doi: 10.1111/1467-8624.00574

Suddendorf, T., Simcock, G., & Nielsen, M. (2007). Visual self-recognition in mirrors and live videos: Evidence for a developmental asynchrony. *Cognitive Development, 22*(2), 185-196. doi: 10.1016/j.cogdev.2006.09.003

Sugarman, S. (1982). Developmental change in early representational intelligence: Evidence from spatial classification strategies and related verbal expressions. *Cognitive Psychology, 14*(3), 410-449. doi: 10.1016/0010-0285(82)90015-9

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson/Allyn & Bacon.

Troseth, G. L. (2003). Getting a clear picture: Young children's understanding of a televised image. *Developmental Science, 6*(3), 247-253. doi: 10.1111/1467-7687.00280

Troseth, G. L., & DeLoache, J. S. (1998). The medium can obscure the message: Young children's understanding of video. *Child Development, 69*(4), 950-965. doi: 10.2307/1132355

Troseth, G. L., Saylor, M. M., & Archer, A. H. (2006). Young children's use of video as a source of socially relevant information. *Child Development, 77*(3), 786-799. doi: 10.1111/j.1467-8624.2006.00903.x

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children’s retention of learned words. *Frontiers in Psychology, 3* doi: 10.3389/fpsyg.2012.00046

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010). Internet-based telehealth assessment of language using the CELF-4. *Language, Speech, and Hearing Services in Schools, 41*(4), 445. doi: 10.1044/0161-1461(2009/08-0131

Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development, 61*(5), 1461-1473. doi: 10.2307/1130756

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*(3), 257-302. doi: 10.1006/cogp.1995.1016

Waxman, S. R., Philippe, M., & Branning, A. (1999). A matter of time: Novel nouns mark object categories when delays are imposed. *Developmental Science, 2*(1), 59-66. doi: 10.1111/1467-7687.00055

Waxman, S., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1*(2), 139-156. doi: 10.1016/S0885-2014(86)80016-8

Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development,* , 349-355.

Wise, B., Cole, R., Van Vuuren, S., Schwartz, S., Snyder, L., Ngampatipatpong, N., Tuantranont, J., & Pellom, B. (2005). Learning to read with a virtual tutor: Foundations to literacy. *Interactive Literacy Education: Facilitating Literacy Environments Through Technology*. *Erlbaum Publishers, Mahway, NJ*.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., Sloetjes, H. (2006). ELAN: a Professional Framework for Multimodality Research. Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition, 85*(3), 223-250. doi: 10.1016/S0010-0277(02)00109-9

Yoshida, H., & Smith, L.B. (2003). Shifting ontological boundaries: How Japanese- and English-speaking children generalize novel names for animals and artifacts. *Developmental Science, 6*, 1-17. doi: 10.1111/1467-7687.00247_1

Yoshida, H., & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science, 16*(2), 90-95. doi: 10.1111/j.0956-7976.2005.00787.x

Zack, E., Barr, R., Gerhardstein, P., Dickerson, K., & Meltzoff, A. N. (2009). Infant imitation from television using novel touch screen technology. *British Journal of Developmental Psychology.Special Issue: Young People and the Media, 27*(1), 13-26. doi: 10.1348/026151008X334700

Zelazo, P. D., Sommerville, J. A., & Nichols, S. (1999). Age-related changes in children's use of external representations. *Developmental Psychology, 35*(4), 1059-1071. doi: 10.1037/0012-1649.35.4.1059

Zimmerman, F. J., & Christakis, D. A. (2005). Children's television viewing and cognitive outcomes: A longitudinal analysis of national data. *Archives of Pediatrics Adolescent Medicine, 159*(7), 619-625. doi: 10.1001/archpedi.159.7.619

Zimmerman, F. J., Christakis, D. A., & Meltzoff, A. N. (2007). Associations between media viewing and language development in children under age 2 years. *The Journal of Pediatrics, 151*(4), 364-368. doi: 10.1016/j.jpeds.2007.04.071

Appendix A

Study 1a Parent Survey Questions and Response Means

Measure is percentage of responses (rounded)
N = 28 parents completed the survey

1. How many hours, on average, does your child watch television a day?

0 to 30 minutes	11
30 minutes to 1 hour	32
1 to 2 hours	36
Over 2 hours	21

a. Weekend?

0 to 30 minutes	11
30 minutes to 1 hour	29
1 to 2 hours	32
Over 2 hours	29

2. What type of shows does your child typically watch (children educational, children entertainment, movies, adult entertainment)?

Children: Educational	93
Children: Entertainment	39
Movies	25
Adult: Entertainment	7
None	0

3. When did your child first start watching TV?

6 months or younger	18
7 to 12 months	25
13 to 18 months	29
19 to 24 months	11
Over 2 years	7
N/A	0

a. What was your child's favorite show or type of show during this time? (Note: parents reported many different shows; the following only includes shows watched by greater than 5% of the sample)

Sesame Street	11
Mickey Mouse Clubhouse	11
Signing Times	7

4. Does your child watch an entire show or does he/she tend to watch a few minutes and then get distracted easily? Explain.

Entire show	69
Distracted	18
Both	14
N/A	0

5. How often do you or another adult watch TV with your child? (Very Often, Often, Not Often, Never)

Very often	14
Often	54
Not often	25
Never	7

a. If an adult watches children's shows with the child do they or you encourage them to participate in the show? (Yes/No)

Yes	86
No	14
N/A	0

b. Explain if yes.

Sing or dance	29
Respond or encourage responding to the screen	25
Answer questions	21
Ask questions	18

Appendix B

Studies 1b, 2, and 3 Parent Screen Media Survey: Questions and Response Means

Measure is percentage of responses (rounded)
N = 166 parents completed the survey

1. On a typical day, which of the following activities does your child spend any time on:

	Yes
Watching TV	50
Watching a video or DVD	58
Listening to music (including while riding in the car)	98
Playing outside	98
Reading or being read to	98
Playing video games like X-Box, Playstation, or Wii	1
Playing inside with toys	99
Playing computer games	8
Using a computer for something other than games	8
Playing with hand-held electronic devices like iPhone, iPad, Leapfrog Platform, Gameboy, etc.	31

2. About how much time does your child spend on those activities in a typical day?

	More than 2 Hours	1-2 Hours	Less than 1 Hour	DK or blank	Mean Hours for Children Who Did This	Did Not Do This	Mean Hours For All Children
Watching TV	8	17	31	8	0:54	37	0:33
Watching a video or DVD	5	11	44	8	0:43	31	0:28
Listening to music (including while riding in the car)	14	17	67	0	0:55	1	0:54
Playing outside	48	40	11	1	1:49	0	1:49
Reading or being read to	6	37	54	2	0:53	0	0:53
Playing video games like X-Box, Playstation, or Wii	1	1	1	16	1:15	82	0:02
Playing inside with toys	70	22	7	1	2:17	0	2:17
Playing computer	0	0	7	14	0:22	78	0:02

games							
Using a computer for something other than games	1	0	11	13	0:18	75	0:03
Playing with hand-held electronic devices like iPhone, iPad, Leapfrog Platform, Gameboy, etc.	1	1	32	10	0:23	57	0:09

3. For most of the time your child watches TV on a typical day, is someone else watching TV or is he/she doing this alone?

Mostly with someone else	77
Mostly alone	11
Child doesn't watch TV/videos/DVDs	12
Don't know/blank	1

4. When your child watches TV or videos, is a parent typically in the room:

The whole time	28
Most of the time	40
About half the time	18
Less than half the time	7
Not at all	1
Child doesn't watch TV/videos/DVDs	6
Don't know/blank	1

5. How many televisions, if any, do you have in your household?

None	9
One	41
Two	30
Three	17
Four	2
Five	1
Six or more	1
Don't know/blank	1

6. When someone is at home in your household, how often is the TV on, even if no one is actually watching it?

Always	0
Most of the time	6
About half of the time	7
Less than half of the time	22
Hardly ever	32
Never	24
No TV in household	8
Don’t know/blank	1

7. Do you have any computers in your household, including laptops and desktops? If so, how many?

None	0
One	21
Two	46
Three	22
Four	5
Five	3
Six or more	1
Don’t know/blank	1

8. When you’re at home with your child and you have something important to do, how likely are you to sit him/her down with a video or TV show while you get it done?

Very likely	11
Somewhat likely	38
Not too likely	27
Not at all likely	19
No TV in household	4
Don’t know/blank	1

9. When your child is playing and the TV is on in the background, how frequently does it distract his/her attention from what he/she is doing?

Often	5
Sometimes	27
Hardly ever	14
Never	2
TV is never on in background	45
No TV in household	7
Don’t know/blank	1

10. At what age did your child first do each of the following things:

	Less than 6 Months	6-11 Months	1 Year	2 Years	DK or blank	Child Does Not Do This
Watch TV	11	19	28	26	3	13
Turn on the TV by themselves	0	1	7	23	2	64
Change the channels with a TV remote	0	1	3	11	2	80
Ask to watch a particular show or channel	0	1	11	58	2	28
Watch a video or DVD	5	16	25	43	2	8
Ask to watch a particular video or DVD	0	0	13	69	3	15
Put in a video or DVD by themselves	0	1	3	24	1	70
Use a computer while sitting on a parent's lap	0	1	14	36	2	45
Use a computer without sitting on a parent's lap	0	0	3	23	2	70
Turn on the computer by themselves	0	0	2	8	2	86
Use a mouse to point and click	0	0	1	24	2	72
Put a CD-ROM into the computer	0	1	2	7	2	86
Look at a website for children	0	1	10	25	3	60
Ask to go to a particular website	0	0	4	10	1	83
Play a video game	0	0	1	9	2	86

11. In total, about how many videos do you have at home for your child, counting both VHS tapes and DVDs and including any shared with brothers or sisters?

None	5
One or two	6
Three to five	14
Six to ten	19
Ten to 19	25
20 to 49	19
50 or more	6
Child doesn't watch videos/DVDs	3
Don't know/blank	1

12. Have they ever had any "Baby Einstein" videos like "Baby Bach" or "Baby Mozart"?

Yes	48
No	46
Child doesn't have any videos/DVDs	4
Don't know/blank	3

13. Which of the following television shows does your child watch regularly? (Note: parents reported many different shows; the following only includes shows watched by greater than 10% of the sample)

Sesame Street	39
Dora the Explorer	28
Mickey Mouse Clubhouse	25
Thomas & Friends	23
Dinosaur Train	21
Caillou	20
Curious George	19
Elmo	12
Bob the Builder	11
Child watches no television regularly	19

14. Which of the following electronic devices, games, and/or apps does your child use regularly?

Smartphone/tablet apps (e.g., Angry Birds, Monkey Preschool Lunchbox, videos, interactive books)	19
Educational games/products (e.g., LeapFrog, VTech, V.Reader)	14
Video games (e.g., Nintento Wii games)	1

15. In general, do you think the following activities mostly help or mostly hurt children's learning—or doesn't have much effect either way?

	Mostly Helps	Mostly Hurts	Not Much Effect	DK/blank
Watching TV	34	38	17	11
Using a computer	42	13	17	28
Playing video games	11	46	10	32

16. This question is about children's intellectual development—things like learning words and counting. How important, if at all, do you think each of the following is in helping the intellectual development of children the same age as your child:

	Very Important	Somewhat Important	Not Too Important	Not At All Important	DK/blank
Reading books	99	1	1	0	0
Building toys like blocks or Legos	83	16	1	0	2
Doing puzzles	75	23	1	0	2
Using educational toys like talking books	22	31	30	11	6
Watching educational TV shows like "Sesame Street"	10	36	37	15	2
Watching educational videos or DVDs	10	34	36	16	3
Playing educational computer games	7	28	37	17	11
Visiting educational websites	7	25	34	21	13

17. Which of the following do you think is **most** important to children's intellectual development?

Books	73
Building toys like blocks or Legos	22
Puzzles	13
Educational toys like talking books	1
Educational TV shows	3
Educational videos or DVDs	2
Educational computer games	1
Educational video games	1
Educational Websites	1
Don't know/blank	0
Other	11
Social interaction with others	8
Play	5
Playing and being outdoors	2
Creative activities (e.g., painting, drawing)	1

Appendix C

Study 3 Word Learning Control Task Materials

Note: descriptions were adapted from *Sesame Street* clips meant to teach words to children.

Word: amphibian
Description: I'm going to tell you about the word amphibian. (Show picture of toad) An amphibian is an animal who lives part of its life on land and part of its life in water, and breathes through its moist skin. This is a toad! A toad lives part of its life in water and part of its life on land, and a toad breathes through its moist skin. So a toad is an amphibian!

Word: Arachnid
Description: I'm going to tell you all about the word arachnid. (Show picture of spider puppet) This is an arachnid. An arachnid is a bug with 8 legs. and spiders are bugs. You can't completely see it, but this bug has 1, 2, 3, 4, 5, 6, 7, 8 legs. 8 legs! This bug has 8 legs! That means it's an arachnid!

Word: Canteen
Description: I'm going to tell you all about the word canteen! A canteen is a small container for water that you use to drink when you go hiking or camping. (Show picture of hose) Now this is not a canteen. This is a hose. (Show picture of fish tank) Now this is a fish tank, but it does hold water but it is not a canteen! (Show picture of bucket) This is a bucket not a canteen! (Show picture of canteen) Do you know what this is? It's a canteen!

Word: Galoshes
Description: I am going to tell you about a pretty fun word. The word is galoshes. Galoshes are waterproof boots you wear in the rain. (Show picture of sneakers) These aren't galoshes; these are sneakers. Galoshes are boots you wear in the rain. (Show picture of cowboy boots) These are boots, but they are cowboy boots. Galoshes are waterproof boots. (Show picture of galoshes) These are galoshes!

Word: hexagon
Description: I'm going to tell you about the word hexagon. A hexagon is a shape! (Show picture of circle) This is not a hexagon; this is a circle. Although a circle is a shape, it is round and without angles. A hexagon has sides and angles. (Show picture of triangle) Now this is not a hexagon. It is a shape and has sides and angles, but is has 1, 2, 3 sides and 1, 2, 3 angles. But it's a triangle, not a hexagon. A hexagon is a shape with 6 sides and 6 angles. (Show picture of hexagon) This is a shape, but not just any shape. This shape has 6 sides and 6 angles—1, 2, 3, 4, 5, 6 angles and 1, 2, 3, 4, 5, 6 sides. This is a hexagon!

Generalization test items

	Amphibian	Arachnid	Canteen	Galoshes	Hexagon
Trained exemplar	Toad	Spider puppet	Green canteen	Pink galoshes	Yellow hexagon
Category member 1	Red frog	Cartoon spider	Clear water bottle	Yellow galoshes	Hexagonal plate
Category member 2	Green frog	Cartoon spider	Red water bottle	Checkered galoshes	Hex nut
Distractor 1	Crocodile	Worm	Hose	Sneakers	Triangle
Distractor 2	Snake	Bird	Swimming pool	Cowboy boots	4 point star
Thematic match	Pond	Spider web	Tent	Umbrella	Crayon