Young Children's Word Learning through Skype

By

Bryan Matthew Koppa

Department of Psychology and Neuroscience, University of Colorado at Boulder

Defense Date: Tuesday November 4th, 2014

Defense Committee:

Eliana Colunga, Department of Psychology and Neuroscience, Thesis Advisor

Richard Olson, Department of Psychology and Neuroscience, Honors Council Representative

Janet Casagrand, Department of Integrative Physiology

## Abstract

Children today are heavily exposed to screen media and beginning to interact with it at a very young age. However, it has been shown that young children do not learn as well from videos as they do from equivalent in person interactions. This study examines how the addition of a live interaction to screen media affects young children's word learning. Thirty-six children between the ages of thirty and thirty-six months were taught novel words either through Skype or in person. Children were then tested to see how well they learned the words and how well they retained the information over a week delay period. Results suggest that learning through Skype and in person may not be all that different.

## Introduction

We live in an age where we are surrounded by screen media in the form of television, computers, phones, tablets, and much more. Screen media has become a highly integrated part of our culture, and it is not just adults that use these devices on a daily basis. Children under the age of two spend approximately one hour (:58) with screen media daily, and children between the ages of two to four average almost two hours (1:58) a day (Common Sense Media, 2013). The American Academy of Pediatrics (APP) suggests that parents should avoid television and video viewing for children under the age of two (APP, 2010). There are many research studies providing support for this claim that children under two years of age do not learn as well from screen media as they do from in person interactions. There is also other research indicating that this learning impairment persists past the age of two, especially for more complicated learning tasks (Scofield & Williams, 2009; Lauricella, Pempek, Barr, & Calvert, 2010;

Sims, 2013). Either way, even though very young children can learn from videos to some degree, the extent of their learning is almost always better from an in person interaction.

Some researchers have theorized that this impairment is due to a lack of social interaction, which can be provided through videoconferencing (Roseberry, Hirsh-Pasek & Golinkoff, 2013). Videoconferencing or video calling is not a recent technological advancement; however, the ease and virtually free access to programs like Skype and FaceTime are a relatively new component to our society. Just like videos and television, these programs have become widely accepted and are used all the time. Therefore, it is not surprising that children are being to be exposed to this technology for wide array of reasons such as communicating with grandparents and family members. In a recent survey taken by thirty-four parents in the Boulder, Colorado area of their thirty to thirty-six month old children, 94.1% had been exposed to Skype or some other form of videoconferencing program. It is worth exploring how the addition of a live interactive component to screen media (e.g., videoconferencing) affects children's screen learning and whether or not this video learning impairment disappears or becomes pronounced.

**The Video Deficit**

Many studies have been performed to assess how well young children learn from videos. The vast majority of these studies have come to the same conclusion that children do not learn as well from videos compared to equivalent in person interactions. This difference in performance has come to be known as the video deficit (Anderson & Pempek, 2005). However, it is not entirely clear how long the video deficit persists developmentally, and the exact degree to which it is present. The majority of research indicates that the video deficit is most

prominent in children under the age of two and that the video deficit is less pronounced as the children get older. However, after the age of two the video deficit may persist in more subtle ways. I will discuss first, the video deficit and its implications, and then secondly, how it relates to children's learning through videoconferencing sources such as Skype.

Imitation has been used as a tool to assess whether or not young children can learn from videos. Barr and Hayne (1999) performed an experiment in which three groups of thirty-six infants ages twelve, fifteen, and eighteen months attempted to imitate three simple actions from different sources. The children watched as the experimenter demonstrated the same three target actions involving a puppet in one of three ways: in person, from a video, or a control group in which no actions were demonstrated. When tested twenty-four hours later, by being given the opportunity to imitate the actions, the infants in each of the three age groups performed better when the initial demonstration was in person rather than from a video or the control. In the twelve and fifteen month old age groups only one child in each group was able to imitate any of the actions from the video and this was no better than the control group. Only in the eighteen month old age group were the children able to imitate actions from the video better than the control; however, children in this age group still performed best from the live demonstration. These results suggest that infants cannot learn a simple imitation task from a video nearly as well as they can from an in person demonstration, clearly indicating a video deficit in these age groups.

Another task used to evaluate young children's learning is the item retrieval task in which the child is asked to retrieve a hidden toy after being given information regarding its location from different sources. In a study by Troseth and DeLoache (1998), twenty-four and

thirty month old children participated in an item retrieval task where they watched on a monitor as a toy was being hidden in the next room which was designed to look like an ordinary living room. Afterwards the children entered the room and attempted to find the toy. The older children were more successful in finding the toy on the first attempt (79%) compared to the younger children (44%). In a second experiment twenty-four month old children were 100% successful in retrieving the toy on the first attempt when they viewed the toy hiding event through a window (i.e., no video). An item retrieval task is more difficult than pure imitation tasks and gives a more in depth measure of learning. Older children were able to perform significantly better on the task, and children who watched the event in person performed the best, again indicating a video deficit and that the video deficit is reduced with increasing age.

Word learning is also another valid measure used to assess how well children learn from videos. Kcrmar, Grela, and Lin (2007) performed a study looking at how well children ages fifteen to twenty-four months learn words from a video. Children were shown objects that they might or might not have seen before, but were unlikely to know their names like a periscope or a spatula. The objects were then given new novel names, and the children were taught the words in one of five conditions: in person, in person while being distracted, a video of the experimenter, a television program (Teletubbies), or a no word control. The testing task required the children to choose the target object from a group of distractor objects, and as predicted the children all performed significantly better from the in person condition. Children chose the correct object 67% of the time in the in person condition compared to the experimenter video (53%), television program (40%), or the in person while being distracted condition (43%). There was also a significant interaction between age and condition on

performance, indicating that the younger children (fifteen to twenty-one months old) had a much harder time identifying the target object in the videos conditions than the older children (twenty-two to twenty-four months old). These results also suggest a video deficit in young children when it comes to word learning that becomes less pronounced as the children get older.

Scofield and Williams (2009) have also investigated word learning from video in young children. Eighteen month to two year olds (mean age = thirty-one months) participated in a study in which they learned novel words from a video and were tested for word learning in various ways onscreen. Their results showed that these children were able to link the novel words/labels to the referent objects (word-to-referent mapping) by choosing the correct object, out of two choices, when asked for it by the novel name that had just been assigned to it at a rate significantly above chance. However, results were different on the disambiguation task which involved giving the child a choice between two objects, one that has been labeled and one which has not been labeled. The child was then asked for the unlabeled object with a different novel label, with the expectation that the child will extend the new novel label to the unlabeled object. On the disambiguation task children did not score above chance indicating that they were unable to disambiguate between two items on a video. This is significant because previous research indicates that children of this age and younger can normally disambiguate between items (Merriman, Bowman & MacWhinney, 1989). Children of this age were able to map the novel names to the novel objects when the information was presented on video, suggesting that they can learn the words from a video; however, these children were not

successful on the disambiguation task, indicating there may some type of video deficit still present with more difficult tasks.

Sims (2013) conducted important research regarding toddlers' word and category learning via screen media. I will address two different experiments performed by Sims that were designed to analyze differences in how children learn from screen media compared to in person interactions. My research will be a direct extension of Sims' work in toddlers' word and category learning.

In study 1a of Sims' dissertation thirty to thirty-six months old children were shown six novel exemplar objects which were then labeled with novel names. The children were taught these words either in person or through a pre-recorded video during a training session. Once the training session was completed, children were tested for word learning immediately and then again after a one week delay. Testing consisted of several different tasks that looked at how well the children could map the given labels to the novel objects, and how they generalized these labels to other objects that matched the trained objects in shape, material, or color. I will briefly describe how the testing sessions were conducted because the procedure used in my experiment was implemented in a similar manner.

The forced choice target identification task was designed to assess whether or not children could simply learn the novel labels given to the novel objects. During the task children were given a choice between two of the exemplar objects that they were originally taught and asked to pick the object with the correct label. From this experiment it was evident that

children of this age could learn simple word-to-referent mapping either from a person or a video, and they could retain this information following a one week delay.

The free choice generalization task was an unrestrained measure of generalization allowing the children to pick as many or as few items as they wanted. This task was designed to see how children generalized characteristic features of the exemplar objects to new unfamiliar objects. They were presented a set of four objects, one of which was the exemplar object and the others matched one characteristic of the exemplar object: shape, material, or color. This task revealed that children trained in person consistently chose objects that matched the trained item and then extended the words to objects that matched in shape, material, and then color respectively. However, children trained from the video did not differentiate between the different objects after the delay.

In study 1b of Sims' dissertation children ages thirty to thirty-six months were again taught novel objects either in person or through a video. The objects taught were slightly different than the previous experiment and were designed to focus on shape-based and material-based categories. The children were again tested immediately after training and after a one week delay using the same testing tasks as in the previous experiment. Results confirmed that children were able to learn simple word-to-referent mapping of the objects either from a person or a video, and retain this information after a one week delay. When directly comparing shape-based matches to material-based matches, testing indicated that children trained both in person and from a video preferred to make generalizations about objects based on shape rather than material. Another interesting finding was that, in both groups, at the delayed

testing children generalized labels to shape-based objects when given the opportunity to choose from multiple objects of various shape and material. However, the children trained by video chose both material-based objects and shape-based objects, indicating that children trained by video do not retain category generalizations as well as children trained in person.

Sims demonstrated that children thirty to thirty-six months old do not show a video deficit when it comes to novel word recognition, which is consistent with previous research (Anderson & Pempek, 2005). Both studies show that children were able to learn simple word-to-referent mapping of novel objects and retain this information over a delay. Furthermore, her research indicates that children prefer to make generalizations about objects based on shape rather than material or color. However, study 1b also indicates that children trained from a video rather than in person do not retain consistent lexical categories based on a single characteristic feature (i.e., shape) over a one week delay period. This indicates there may be a video deficit in this age group when it comes to word learning and retention of lexical categories. My research will delve further into this topic of word and category learning from screen media, focusing on videoconferencing instead of video as the source of information.

**Live Onscreen Interactive Learning**

The easy access to videoconferencing tools, such as Skype and Face Time, are a relatively new component to our society. Children are being exposed to these types of media at younger and younger ages. In a two year span between 2011 and 2013, the time that children under the age of eight spent on mobile devices, such as smartphones and tables, tripled (Common Sense Media, 2013). It is important to understand how effective these sources of

communication truly are for young children. Not only are these forms of communication being used in a home setting, but they are also being implemented as a cheap and easy method of providing early intervention therapies to young children for a wide array of speech and language impairments. It has been shown that children do not learn as well from videos as they do from equivalent in person interactions. The reason for this impairment is unknown; however, some have theorized that the video deficit is due to a lack of social interaction, which can easily be provided through videoconferencing programs (Roseberry, Hirsh-Pasek & Golinkoff, 2013). I will now address how the addition of a live onscreen interaction (i.e. videoconferencing) affects young children's screen learning.

Nielson, Simcock, and Jenkins (2008) conducted two experiments to assess children's learning from live onscreen interactions (videoconferencing) compared to their learning from video and in person sources. The study implemented an imitation task for twenty-four month old children that consisted of using a tool to open a box that contained a hidden toy. In the first experiment an in person condition was compared to a video condition and a control condition. An experimenter demonstrated three times how to open the box using the tool which revealed the hidden toy either in person or on a video, and then the child attempted to imitate the actions of the experimenter to retrieve the toy. Children were more likely to imitate the action of using the tool to open the box when the information was presented in person rather than from a video. The same procedure was used in the second condition, but an interactive condition using videoconferencing was compared to two different non-interactive video conditions. Children in the interactive condition were more likely to imitate the action of using the tool to open the box than the children in the non-interactive condition. Not only were

children more likely to imitate the action in the interactive condition, but the children in this group performed equally as well as the children in the in person condition from the first experiment. This suggests that the live interactive component of videoconferencing may help children two years of age overcome the video deficit. However, this may only be true for pure imitation and not more complicated tasks.

Another study performed to explore children's learning from videoconferencing used an object retrieval task. Twenty-four month old children participated in a study that involved finding a hidden toy after receiving information regarding its location from one of two different sources. The toy was hidden in one of several locations within the hiding room, which consisted of normal objects one might find in a standard living room. The children were then told the location of the toy by the researcher either via live onscreen interaction (videoconferencing) or directly in person. The results showed that children were less successful in finding the toy on the first try when the location of the toy was presented through the videoconferencing, 27% compared to 77% correct for the in person condition (Troseth, Saylor & Archer, 2006). These results indicate that the live onscreen social interaction provided by videoconferencing may not be enough to help child two years of age overcome the video deficit with more complicated tasks such as the item retrieval task.

In a different object retrieval task, children ages thirty to thirty-six months participated in a hide-and-seek object retrieval game where the source of information regarding the toy's location was presented in one of three ways: a simple interactive computer game, a video, or observing an adult through a one way mirror. Children in the interactive game condition and the children who watched the live demonstration performed significantly better on the object

search task than children who watched the video (Lauricella, et al., 2010). This study suggests that there is still a video deficit among children as old as two and one-half to three years old and that their learning may be able to be improved with the addition of an interactive component to the screen media source.

Roseberry, Hirsh-Pasek, and Golinkoff (2013) have addressed another line of inquiry by evaluating children's verb learning from live interactive onscreen sources (e.g., Skype). Children twenty-four to thirty months old were taught 4 novel verbs through an in person interaction, a live onscreen interaction (Skype), or through a non-interactive video of previous Skype verb learning sessions (i.e., no live interaction). After the children had been trained on the verbs, they were tested in several ways to gain different measures of how well the children actually learned the verbs. One testing trial assessed whether or not the children could generalize the verbs to a novel actor performing the action, while another testing trial used a disambiguation task to gain a more in depth evaluation of the children's verb learning. The results showed that children only learned the verbs from live interactive sources, such as Skype and in person, but not from a video. This study provides evidence to support the idea that the live interactive component of videoconferencing may be able to help children overcome the video deficit and be a valid method of learning for children this age.

**Telehealth**

A new emerging field utilizing videoconferencing for young children known as telehealth is gaining traction as a viable method of delivering early interventions to young children with a variety of disorders such as speech, language, or auditory impairments. These intervention techniques are most effective when they are delivered early, ideally as soon as the impairments

are recognized. However, these interventions can be costly and must be administered by professional speech and language pathologists. Not only that, but rural families that live away from major cities may not have access to these intervention therapies. Videoconferencing may provide a viable option for providing these types of intervention therapies to young children that may otherwise not have access to them. The validity of videoconferencing for young children must first be assessed before these types of treatments can be effectively administered.

Most studies regarding telehealth have focused on either intervention programs or telehealth evaluations for elementary school aged children. However, one study looked at the reliability of speech, language, and hearing screenings for children six years of age and younger. In total, 411 speech-language and/or auditory videoconferencing screenings were administered over a two year span with 151 of those children being under the age of three. The reliability of these screenings was determined by comparing pass/fail rates of the different screening components to screenings conducted at an on-site clinic. Reliability was 100% for pure tone hearing screening, distortion product otoacoustic emissions (DPOAE) screening, and speech-language screening, and 84% reliable for tympanometry screening (Ciccia, Whitford, Krumm & McNeal 2011). No comparison was reported for reliability differences based on the age of the children. Another study looking at the administration of childhood language assessments via an internet based telehealth system was administered by two speech language pathologists simultaneously through the internet and in person using the Clinical Evaluation of Language Fundamentals 4$^{th}$ edition. Twenty-five children ranging in age from five to nine years old were assessed, and the results indicated that there was no significant difference in the scores

achieved (Waite, Theodoros, Russel & Cahill, 2010). These studies suggest that screening processes via videoconferencing for these types of impairments in children, including children younger than three years of age, may be valid.

No studies have actually implemented speech and language intervention therapies for children under the age of three, but this next study addresses the validity of a telehealth intervention program via videoconferencing in elementary school aged children. Thirteen children with speech sound impairments (aged 6 to 11) from a rural Ohio school district participated in the study. The intervention was administered either in person or through videoconferencing software by a speech-language pathologist. Goldman Fristoe Test of Articulation-2 was used as an assessment to measure pre and post intervention scores. The students in each group showed significant gains and there was no significant difference between the groups with regard to their pre-test and post-test scores (Grogan-Johnson et al., 2011), suggesting that telehealth may be a viable option for implementing speech sound intervention for elementary school aged children. Much more research will be required to fully understand the implications of telehealth intervention programs, but first assessing the validity of young children's learning via videoconferencing will provide an important starting point.

**The Current Study**

I will address how children learn from live onscreen interactive media (i.e. Skype) using novel noun word learning as the foundation for the evaluation. Word learning through Skype will be compared to an in person control group to assess whether or not there is a difference in the initial learning or retention of the material. I am interested in not only simple word-to-referent mapping of the novel words, but also more subtle differences in learning such as the

generalizations children make about the meaning of the words. Based on previous research I expect that children will be able to learn the simple word-to-referent mapping through Skype equally as well as they do from in person interactions. However, if Skype is not as valid of a medium for learning as an in person interaction, there should be differences in how children generalize and/or retain the words.

# Methods

### Participants

Thirty-six children (Mean age = 33.45 months, SD = 1.58 months, 14 boys, 22 girls) from the Boulder, Colorado area participated in the study. All children completed both sessions, and are included in the data analysis.

### Materials

Children were taught six novel words for six novel objects for the experimental task. The six exemplar objects were given one of six names: Zeb, Gub, Ife, Lug, Nork, and Elg (Figure 1a). For each exemplar object there was a set of five objects that consisted of two shape matches, two texture matches, and one non-match/distractor item (Figure 1b). The novel words were chosen to be simple, single-syllable words that are phonologically allowable in the English language.

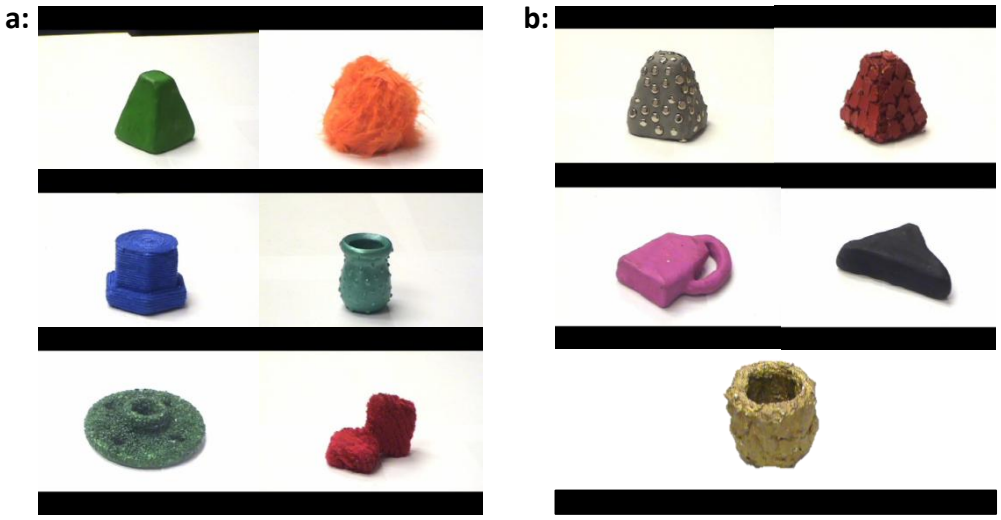


Figure 1: a) Novel exemplar objects used for the experiment during training (e.g. Zeb, Gub, etc.) b) A set of five objects that represent one of the exemplar objects, consisting of two shape matches (top), two texture matches (middle), and one non-match/ distractor item (bottom).

**Design**

Children completed two sessions one week apart (M = 7.69 days, SD = 2.2 days). They were randomly assigned to one of two training conditions: in person or through Skype. The two conditions represent how the novel words and objects were presented during the training portion of the first session. The second session was completed in person in both conditions.

**Procedure**

**Training:** Children were trained on the novel words and novel objects either in person or through the program Skype. Each of the six exemplar objects was presented and labeled three times, and then repeated so that the children saw each object twice and each object was labeled a total of six times. The labeling of each exemplar object was presented as follows: "This is an *ife*. Do you see the *ife*? Can you say *ife*?" During training parents were asked not to participate or help their child learn the words in any way.

In the Skype condition the child and parent were escorted into the first experiment room which contained a laptop with Skype open. The experimenter then moved to the second experiment room, which contained an iPad connected to the laptop online through Skype. The experimenter then began training by placing the first exemplar object in front of the camera so that the child could see the object on their screen. While the object was in front of the camera it was rotated and labeled with the three phrases. The object was then placed out of view of the camera and the same procedure was repeated with the remaining five exemplar objects. The entire process was then repeated. Once the training session was complete the experimenter reentered the first experiment room to put the laptop away.

In the in person condition the experimenter was in the same room as the child and parent. The experimenter then began training by placing the first exemplar object in front of the child rotating the object as it was labeled. This was done for the remaining five exemplar objects and then the entire process repeated. In both conditions, only one object was in the child's view at any one time, and once training was completed they immediately moved on to testing.

**Testing Session 1:** Children were tested in person for both conditions. Each testing session began with a practice free choice generalization task. A tray with two different balls and two non-ball shaped items were presented to the child. The experimenter had a third different ball and labeled it as a ball to the child, then asked the child to pick the other balls on the tray. If the child picked a non-ball item he/she was told why it is incorrect and asked to try again. The practice trial was to establish that the child should only pick the items asked for.

*Free Choice Generalization Task:* The first task completed was a free choice novel noun generalization test. The children were presented with a tray of five objects (e.g., Figure 1b) that consisted of two shaped-based matches, two material-based matches, and one distractor item. These objects matched one of the six exemplar object in one distinguishing feature (i.e. shape or material), and the child was asked to choose an item by a trained label (e.g., "Can you show me an *ife*?"). Once an item was selected they were asked if they saw another item with the same label until they either picked all the items or confirmed that there were no more items with that label (e.g., "Do you see another *ife*?"). The order in which each item was picked was recorded.

*Forced Choice Target Identification Task:* Next, the children were given a choice between two of the original exemplar objects and asked to pick the object with the correct label. This was repeated for a total of six times so that each exemplar object was asked for once. (e.g., "Which one of these is an *Ife*?") The experimenter recorded the responses as either correct or incorrect. The forced choice task was the last task given for the first testing session.

**Testing Session 2:** When the children came in for their second visit there was not another training session. They began with the practice ball trial and then moved immediately into testing. This testing session began with the free choice generalization task and the forced choice target identification task, just like the first testing session. Then the children participated in two additional testing tasks which tested in a more controlled fashion whether or not children would generalize the trained labels to objects that matched in shape and material. While the children were completing the second testing session the parents were asked to fill out a survey concerning the child's screen media usage.

*Forced Choice Shape vs. Shape Generalization Task:* Children were presented with two shape-based objects matches from different item sets and asked to pick an object based off of a trained label. This task gives an extra measure of whether or not the children will extend a trained label to objects that match the original trained object in shape.

*Forced Choice Shape vs. Material Generalization Task:* In the final task children were presented with a shape match and a material match from the same set and asked to pick the correct object based off of a trained label. This task looked directly at whether children preferred to extend the trained label to objects based on either shape or material.

## Results

**Target Identification Task**: The first analysis examined is a simple measure of learning, looking at whether or not the children were able to learn the word-to-referent mapping. In other words, were the children able to learn the novel names given to the novel objects? The forced choice target identification task which was administered at both visits will be analyzed to give the most accurate answer to this question. Analysis showed that children were initially able to learn the word-to-referent mapping above chance levels in both conditions immediately (Skype: $M$ =.73, $SD$ = .21, $t(34) = 4.56$, $p < .001$; In Person: $M$ = .69, $SD$ = .22, $t(34) = 3.69$, $p < .001$), and after the one week delay (Skype: $M$ = .64, $SD$ = .21, $t(34) = 3.06$, $p < .001$; In Person: $M$ = .73, $SD$ = .21, $t(34) = 4.56$, $p < .001$). The next question to address is whether or not the learning of the word-to-referent mapping was better in one training condition compared to the other. Two tailed t-tests revealed that there was no significant difference between conditions at either visit ($t < 1$, $p > .05$ for both visits).

Not only were children in both conditions (Skype and In Person) able to learn the novel names of the novel objects and retain this information over a delay period of one week, but performance did not differ significantly depending on how the words were initially learned for this task. These results indicate that there is no video deficit for simple word-to-referent mapping in children two and a half to three years old which is consistent with previous research.

**Generalization Tasks**: The remaining tasks were used to gain a measure of how children generalized the novel names to other objects that matched the exemplar object in either shape or material.

*Forced Choice Shape vs. Material Generalization Task:* This task was designed to measure whether children preferred to make generalizations about the trained exemplar objects based off its shape or its material. This task was administered only at the second visit so it provides no measure of retention. Children trained in person chose shape-based matches more often than material-based matches, which was significant when compared to chance ($M$ = .74, $SD$ = .19, $t(34) = 5.33$, $p < .001$). The same result was obtained for children trained through Skype ($M$ = .64, $SD$ = .24, $t(34) = 2.59$, $p < .05$), and a t-test revealed there was no significant difference between conditions ($t < 0$, $p > .05$). This result provides evidence to support the prediction that children will generalize about a new label based on shape rather than material (Figure 2). However, this test offers only a restrained measure of generalization because children were forced to choose between two objects.

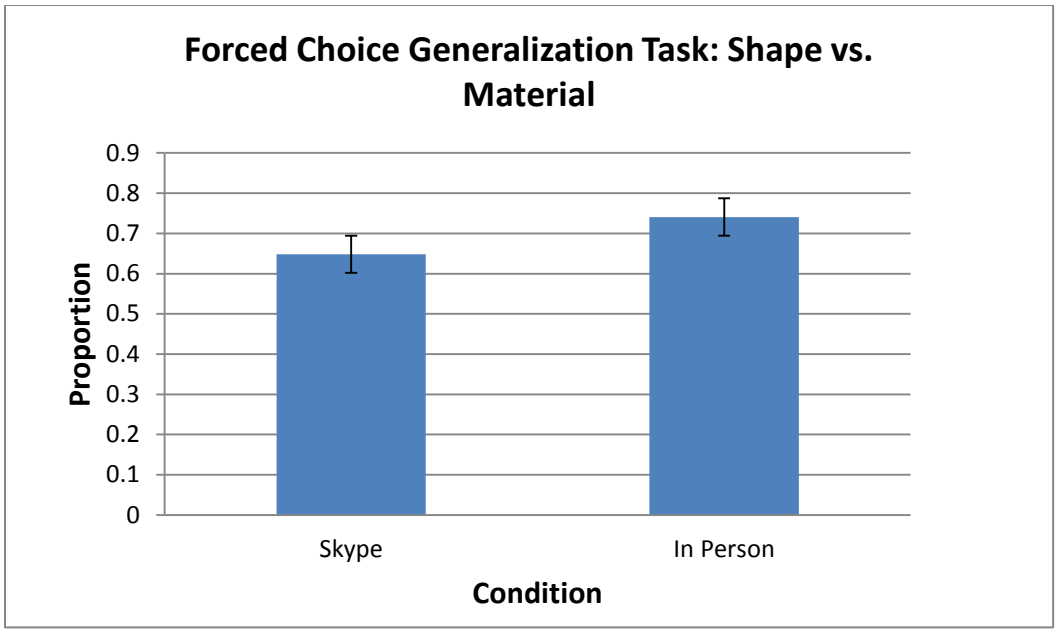


Figure 2. The proportion of shape-based objects chosen out of two choices: a shape-based match or a material-based match. For this task children who chose the shape-based match scored a 1 and children who chose the material-based match scored a 0; therefore, a mean score of 1 indicates that the children chose the shape match every time, a mean score of .5 indicates that the children had no preference, and a mean score of 0 indicates that children chose the material-based match every time. Children from both conditions chose the shape-based object significantly above chance.

*Forced Choice Shape vs. Shape Generalization Task:* This task asked whether or not the children would extend a trained label to objects that matched the original trained object in shape. This task was also administered only at the second visit so there is no measure of retention, and it provides a restrained measure of generalization like the previous task. Children trained in person succeeded above chance levels ($M$ = .69, $SD$ = .14, $t(34) = 5.77$, $p < .001$) and so did children trained through Skype ($M$ = .62, $SD$ = .15, $t(34) = 3.5$, $p < .01$), indicating that they could extend the trained label to other objects matching in shape. When the conditions were compared to each other there was no significant difference ($t < 0$, $p > .05$), suggesting that this generalization about shape was not dependent on how the children were originally trained on the labels.

*Free Choice Generalization Task:* This task was an unrestrained measure of generalization allowing the children to pick as many or as few items as they wanted. For each set, they were presented with a set of five objects that matched the exemplar objects in shape, material, or with no matching features as a control, which was referred to as the distractor. This task was designed to see how children generalized characteristic features of the exemplar objects to new unfamiliar objects (i.e., Did they prefer to extend the label given to the exemplar to new objects based on shape or material?). To gain a broad overview of the data, it was subjected to a mixed models analysis of variance (ANOVA) including all three independent variables 2 (Condition: Skype or in person) x 2 (Visit: first or second) x 3 (Item Type: shape match, material match, or distractor). This revealed that there was no significant interaction among all three variables, but that there was a main effect among the different item types ($F(2,68) = 19.06$, $p < .001$) and a main effect between visits ($F(1,34) = 4.43$, $p < .05$). No other effects were found.

The main effect between visits indicates that during the free choice generalization task children chose more total object matches at the delayed testing ($M = .80$, $SD = .30$) compared to the initial testing session ($M = .73$, $SD = .32$) ($t(107) = 3.33$, $p < .001$), suggesting that children became less conservative in their generalizations after a week delay.

Further analysis of item type was conducted to determine the exact differences among item type. Overall, children in both conditions preferred to make generalizations about the exemplar objects based on shape rather than material or the control (Figure 3). Post hoc t-tests revealed that at the first visit children in the in person condition chose shape-based objects significantly more often than material-based objects ($t(34) = 2.98$, $p < .01$) or the distractor

objects ($t(34) = 3.49$, $p < .01$). However, there was no difference ($t(34) = 2.01$, $p > .05$) between the number of material-based objects and distractor objects chosen, indicating no preference or distinction among these items. Children in the Skype condition also displayed the same pattern of generalization at the first visit. They chose shape-based objects more often than material-based objects ($t(34) = 3.49$, $p < .01$) or distractor items ($t(34) = 4.26$, $p < .01$), and they did not distinguish between material-based objects and distractor objects ($t(34) = 1.55$, $p > .05$). This data indicates that the initial learning in both conditions did not differ significantly regarding the generalizations children made about the exemplar objects.

At the second visit, children demonstrated the same pattern of object discrimination. In both the in person and Skype conditions children chose shape-based objects more often than material-based objects (In Person: $t(34) = 2.87$, $p < .05$; Skype: $t(34) = 3.12$, $p < .05$) or distractor objects (In person: $t(34) = 2.67$, $p < .05$; Skype: $t(34) = 2.87$, $p < .05$), and did not distinguish between material-based objects and distractor objects (In person: $t(34) = .84$, $p > .05$; Skype: $t(34) = -0.14$, $p > .05$). These results suggest that children preferred to make generalizations about the characteristic features of objects based on shape rather than material, and that this discrimination was consistently retained over a one week delay period, regardless of the source of the information.

| | Immediate Testing (Visit 1) | | Delayed Testing (visit 2) | |
|---|---|---|---|---|
| | In Person | Skype | In Person | Skype |
| Shape | .85 | .88 | .95 | .89 |
| Material | .64 | .74 | .72 | .78 |
| Distractor | .59 | .68 | .70 | .78 |

Table 1. The average proportion of choices made for each item type at each testing session and for each condition.

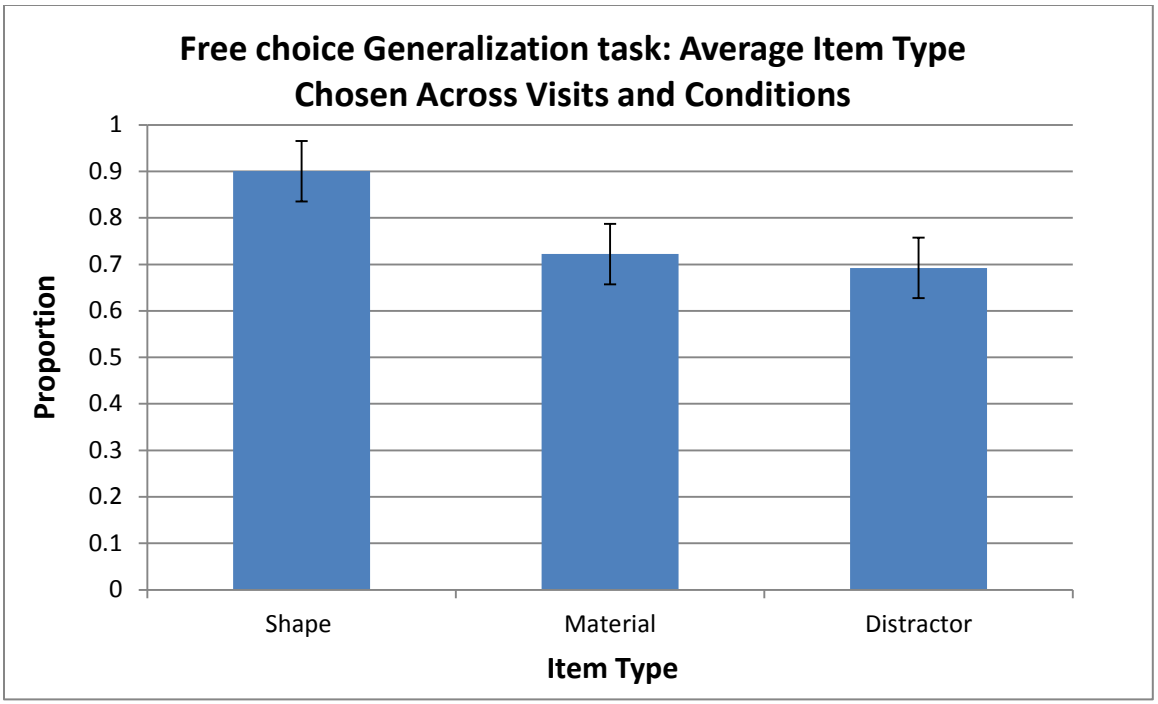


Figure 3. The average of children's choices during the free choice generalization task for both conditions and both visits.

The results of the free choice generalization task differ from the results obtained by Sims (2013) in a very similar task with children of the same age comparing children's video learning to in person learning. Sims concluded that children who learned the words from a video did not retain the discrimination between shape-based and material-based objects over a one week delay, while the children who learned the words in person did. From these conclusions it would appear that learning through Skype is more equivalent to in person learning than a video.

**Video Comparison:** In order to more fully examine the video deficit it would be interesting to compare my results to a video condition (i.e. no live interaction), which my study does not provide. However, I was able to obtain unpublished data from the CU Language Project lab of a highly analogous experiment that used a video condition and taught the novel

objects from a video rather than through Skype. The objects used in both experiments were also the same, except that my study contained a distractor object during the free choice generalization task which was not included in the video condition. Therefore, comparison of the free choice generalization tasks would not provide accurate conclusions. The forced choice target identification task, however, used the same objects and was run identically in both studies. In addition, procedure, training, testing, age, and length between visits were all fundamentally the same. The main difference is that children were trained on the words from a video recording rather than through the program Skype. While these are not the ideal circumstances in which to make this comparison they will still provide a rough analysis of how children's word learning differs between a video and Skype. With that said, all conclusions drawn from the following analysis are not scientifically valid because they are not exactly the same, and should only be used to give a rough estimate of how these two sources of learning might be different.

The in person and Skype conditions of my study will be compared to the video condition of the unpublished data obtained from the CU Language Project lab for the forced choice target identification task only. This task was administered at both visits, which were one week apart, and will provide a measure of whether or not the children were able to learn the novel labels given to the novel objects (simple word-to-referent mapping). T-tests revealed that children in the video condition did not perform above chance levels at the first visit ($M$ = .54, $SD$ = .14, $t(45) = 1.44$, $p > .05$), but that they did perform above chance levels at the second visit ($M$ = .59, $SD$ = .18, $t(45) = 2.50$, $p < .05$). The forced choice target identification task for the in person and

Skype conditions, analyzed earlier, revealed that children in both groups were able to perform above chance levels at both visits.

Further analysis revealed that at the first visit children performed significantly better on the task when trained in person compared to the video condition ($t(40) = 2.71$, $p < .01$), and children in the Skype condition performed significantly better then children in the video condition ($t(40) = 3.45$, $p < .01$). At the second visit, children trained in person still performed better than children in the video condition ($t(40) = 2.14$, $p < .05$), however, there was no difference in performance between the Skype and video conditions ($t(40) = .83$, $p > .05$). These results suggest that initially word learning may be more effective in person than from a video, and that retention is better when the words are originally learned in person. Word learning through Skype, is also initially more effective than from a non-interactive video, but that over a delay these benefits may not be retained any better when taught through Skype.

## Discussion

The current study begins to address how young children learn from screen media with the addition of a live interactive component (i.e., Skype) and using word learning as the medium. There were some interesting trends in the data, but overall there was no significant difference in the performance on any of the various tasks between the two conditions. These results indicate that word learning through Skype, for children between the ages of thirty and thirty-six months of age, is very similar to their word learning in person.

Performance on the target identification task revealed that children were able to map the novel labels given to the novel objects. This was based on their ability to select the correct objects above chance levels after the training session. Furthermore, there was no significant

difference in performance between either of the condition, indicating a similar degree of learning with regards to simple word-to-referent mapping. The generalization tasks provide a more in depth look into the children's word learning by gaining a measure of how the children extended the novel labels to other objects. The forced choice shape vs. material generalization task provided an answer to this question with the caveat that the task only allowed the children to pick between two objects. However, from this task it was evident that the children preferred to make generalizations about the exemplar objects based on their shape rather than their material which is consistent with previous data (Sims, 2013). Additionally, there was no difference between the two groups in the degree to which the children made these generalizations.

The free choice generalization task provided another look at the question of how the children preferred to extend the novel labels to other novel objects based on a matching characteristic feature. This task was different because it was an unrestrained measure of generalization allowing the children to pick as many or as few objects that they thought matched the exemplar. The results supported the findings of the forced choice shape vs. material generalization task in that children preferred to make generalizations about the exemplar objects based of shape rather than material or the control. There was also no difference between the conditions in how the children made these generalizations, further indicating a similarity in how children learn words from these two sources. The results from the experiment all support the conclusion that word learning in children thirty and thirty-six months old through Skype and in person is relatively comparable.

Comparisons to the video condition indicated that in person learning was better both initially and after the delay. When word learning through Skype was compared to the video condition it revealed that initially learning through Skype was more robust. However, at delayed testing Skype word learning was not significantly better than the video condition, but it was also not significantly worse than in person word learning. These results indicate that word learning through Skype may fall somewhere in between learning in person and from a video. With that said, these conclusions can only be taken as speculation because the conditions come from two different experiments.

The reason for this increase in quality of learning through Skype compared to a video may be due to the social interaction that it provides (Roseberry, Hirsh-Pasek & Golinkoff, 2013). Learning in person provides a social interaction which may be crucial for learning, and Skype provides a more equivalent socially interactive learning experience. My experiment provides support for the claim that learning in person and through Skype are comparable, which appears to be due to the social interaction that Skype offers.

With telehealth emerging as a new and likely candidate for administering early speech and language intervention for young children, validating the technology being implemented becomes even more significant. Most telehealth studies have focused on children above the age of three; however, it will be important to establish an age range in which this type of therapy is most effective, particularly since these intervention therapies produce the most beneficial results when administered as early as possible. My results may help validate the implementation of telehealth programs for children even younger than those currently being targeted.

## Conclusion

Overall, the results of my experiment point toward the conclusion that there is no video deficit when it comes to novel noun word learning through Skype for children thirty to thirty-six months of age. It would appear that these children can learn new words through Skype equally as well as they can learn them in person. In addition to simple word-to-referent mapping, the source of the information does not seem to impair the generalizations that children make about the words. These results may have implications in both a professional and home setting as children's exposure to Skype continues to increase.

Works Cited

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

Anderson, Daniel R., and Tiffany A. Pempek. "Television and very young children." *American Behavioral Scientist* 48.5 (2005): 505-522.

Barr, Rachel, and Harlene Hayne. "Developmental changes in imitation from television during infancy." *Child development* 70.5 (1999): 1067-1081.

Ciccia, Angela Hein, et al. "Improving the access of young urban children to speech, language and hearing screening via telehealth." *Journal of telemedicine and telecare* 17.5 (2011): 240-244.

Common Sense Media, and Victoria Rideout. *Zero to eight: Children's media use in America*. Common Sense Media, 2013

Grogan-Johnson, Sue, et al. "A pilot investigation of speech sound disorder intervention delivered by telehealth to school-age children." *International Journal of Telerehabilitation* 3.1 (2011).

Krcmar, Marina, Bernard Grela, and Kirsten Lin. "Can toddlers learn vocabulary from television? An experimental approach." *Media Psychology* 10.1 (2007): 41-63.

Lauricella, Alexis R., et al. "Contingent computer interactions for young children's object retrieval success." *Journal of Applied Developmental Psychology* 31.5 (2010): 362-369.

Merriman, William E., Laura L. Bowman, and Brian MacWhinney. "The mutual exclusivity bias in children's word learning." *Monographs of the society for research in child development* (1989): i-129.

Nielsen, Mark, Gabrielle Simcock, and Linda Jenkins. "The effect of social engagement on 24-month-olds' imitation from live and televised models."*Developmental science* 11.5 (2008): 722-731.

Roseberry, Sarah, Kathy Hirsh-Pasek, and Roberta M. Golinkoff. "Skype me! Socially contingent interactions help toddlers learn language." *Child development* 85.3 (2013): 956-970.

Sims, Clare Elizabeth. *Toddlers' Word and Category Learning from Video: Links Between Learning and the Screen Mediated Environment*. Diss. UNIVERSITY OF COLORADO AT BOULDER, 2013.

Troseth, Georgette L., and Judy S. DeLoache. "The medium can obscure the message: Young children's understanding of video." *Child development* 69.4 (1998): 950-965.

Troseth, Georgene L., Megan M. Saylor, and Allison H. Archer. "Young children's use of video as a source of socially relevant information." *Child Development* 77.3 (2006): 786-799.

Waite, Monique C., et al. "Internet-based telehealth assessment of language using the CELF–4." *Language, Speech, and Hearing Services in Schools* 41.4 (2010): 445-458.