

The title for this Special Section is **Contemporary Mobile Technology and Child and Adolescent Development**, edited by Zheng Yan and Lennart Hardell

Look At That! Video Chat and Joint Visual Attention Development Among Babies and Toddlers

Elisabeth R. McClure, Yulia E. Chentsova-Dutton, Steven J. Holochwost*, W. G. Parrott, and Rachel Barr
Georgetown University

Although many relatives use video chat to keep in touch with toddlers, key features of adult–toddler interaction like joint visual attention (JVA) may be compromised in this context. In this study, 25 families with a child between 6 and 24 months were observed using video chat at home with geographically separated grandparents. We define two types of screen-mediated JVA (across- and within-screen) and report age-related increases in the babies' across-screen JVA initiations, and that family JVA usage was positively related to babies' overall attention during video calls. Babies today are immersed in a digital world where formative relationships are often mediated by a screen. Implications for both infant social development and developmental research are discussed.

Many young children today are geographically separated from their immediate or extended families, and remote communication is an important resource for them (McClure & Barr, 2016). Telephones can be difficult to use for children under 7 years of age (Ballagas, Kaye, Ames, Go, & Raffle, 2009), however, and video chat has been suggested as a viable alternative because it offers opportunities for visual communication (Ballagas et al., 2009; Tarasuik, Galligan, & Kaufman, 2013). Mobile device usage is widespread, even among children under 1 year of age, with many children using a device daily by age 2 (Kabali et al., 2015). There is some evidence that young children are first being introduced to video chat at younger ages than in the past (Tarasuik & Kaufman, 2017), and even families that restrict screen time for their children under 2 have been shown to view video chat as an exception to their media restrictions

(McClure, Chentsova-Dutton, Barr, Holochwost, & Parrott, 2015). These trends are likely to continue, especially now that the American Academy of Pediatrics has revised its media recommendations for children under 2 to make an exception for moderate video chat usage (American Academy of Pediatrics: Council on Communications and Media, 2016). In fact, as many as 85% of U.S. families with children have used video chat—frequently on mobile devices—to help their babies and toddlers keep in touch with remote relatives, most of them with grandparents who live far away (30 or more miles; McClure et al., 2015). Despite the prevalence of usage, research examining video chat use by babies under 2 in their own homes is limited. Although recent lab research has demonstrated that toddlers can learn new words via video chat (Myers, LeWitt, Gallo, & Maselli, 2016; Roseberry, Hirsh-Pasek, & Golinkoff, 2014), it is unclear how families are using these technologies in their typical day-to-day lives.

In particular, it is unclear whether the limitations imposed by development and the technology itself prevent babies in this age range from engaging in successful joint visual attention (JVA) during video chat interactions. JVA, or “following the direction of attention of another person to the object of their attention”

*Current affiliation: Johns Hopkins University.

The authors thank Rachel Cohen, Jon McClure, Katerina Maylock, Laura Pevytoe, Danyang Han, Amritpal Singh, Maia Sekulski, Dorosi Valle Flores, Emilia Pickard, Theresa Gladstone, Aya Kellam, Vasiliki Dimopoulos, Marissa Menakaya, Alex Grevers, Nadia Al-Taie, Emily Fletcher, Amanda Grenell, Herietta Lee, and Tish Tyson for assisting with the data collection and collation. The authors also thank the parents and babies who made this research possible.

Correspondence concerning this article should be addressed to Elisabeth R. McClure, elisabethmclure@gmail.com, and Rachel Barr, rbb5@georgetown.edu. Paper mail may be sent to Department of Psychology, Georgetown University, 3700 O St. NW, Washington, DC 20057.

© 2017 The Authors

Child Development © 2017 Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2018/8901-0003
DOI: 10.1111/cdev.12833

(Butterworth, 2004; p. 213), is thought to be an important precursor to language and theory-of-mind (ToM) development (Baron-Cohen, 1991; Butterworth, 2004), so compromises in JVA may disrupt the interaction in critical ways for both the baby and the remote relative. JVA develops rapidly during the first 2 years of life. Starting around 9 months, infants begin to more frequently make bids for adult attention to engage around objects and also to follow an adult's gaze to objects of the adult's interest (Bakeman & Adamson, 1984). It is not until 18 months, however, that babies begin to follow others' attention to objects that are behind them (Butterworth & Cochran, 1980).

Video chat may make it difficult for babies and toddlers to successfully participate in what we will call *across-screen JVA* during video chat interactions. First, video chat cameras are often mounted above the screen. For this reason, when a remote grandparent makes eye contact with the baby or points to something in the baby's environment as displayed on their own screen, the grandparent is actually referencing a point well below the camera's aperture. The baby's view of the grandparent, then, is of the grandparent focusing several inches below the eye level or below the item being referenced (McClure & Barr, 2016). This misalignment of gaze direction could make across-screen JVA more difficult for infants.

Furthermore, the shared visual space between babies and their video-mediated social partners is limited to what can be displayed through their screens. Because babies are usually fairly close to the camera during a video call, if the remote relative references something in the infant's environment the object is likely to be *behind* the infant. For this reason, across-screen JVA may be limited for children under 18 months without the assistance of a physically present adult (McClure & Barr, 2016). Moreover, given the constricted and disconnected "shared" visual environment during video chat, it is conceivable that across-screen JVA toward objects behind the toddler will be delayed relative to the face-to-face emergence of this ability. It is possible, then, that many children could continue to struggle with adult-initiated, across-screen JVA through the second year of life.

Babies and toddlers may still be able to participate unaided in successful JVA during video chat, however, if the adults involved adapt to their needs and to the limitations of the technology (McClure & Barr, 2016). A baby may be able to follow the gaze or point of a remote relative whose attention is directed toward an object or event within the relative's *own* environment (Hood, Willen, & Driver, 1998), as long as the object or event is within the view of the relative's camera. For example, if a

grandmother brings a teacup up to the camera to show her grandchild, she has bypassed the limitations of the camera's location, allowing the child to more easily follow the grandmother's attention to that teacup. Babies themselves may also be able to initiate this type of JVA, which we will call *within-screen JVA*, by directing their remote grandparent's attention to items nearby in the babies' own environment (e.g., the baby shows his grandmother his own snack or nearby toy). This is an encouraging possibility and one that may provide the opportunity for both generally successful social interaction and attempts at toy play across the screen. Within-screen JVA is quite similar to traditionally defined JVA, with few, if any, additional impediments added due to screen mediation, so one might expect this type of JVA to emerge among babies at the typical time developmentally, around 9 months of age.

This study utilized an observational design to investigate families' JVA usage when video chatting at home on both mobile devices and laptops, focusing on two sets of issues. First, we hypothesized that there would be age-related differences in the initiation of across- and within-screen JVA types: all babies were expected to initiate within-screen JVA more often than across-screen JVA, and given the hypothesized difficulty of using across-screen JVA, older babies were predicted to be more likely to begin initiating across-screen JVA than younger babies. Second, we assessed whether across-screen JVA predicted children's overall attention to the video calls. Given the early stage of these babies' attentional development (Betts, McKay, Maruff, & Anderson, 2006; Colombo, 2001), it was expected that across-screen JVA initiated by the adults (e.g., the parent directs the child's attention to something on the screen or the remote grandparent directs the child's attention to something in the child's environment) would support the child's attention during the video call period. Furthermore, given the well-studied relation between an adult's responsiveness and a child's attention during social interactions (e.g., Hains & Muir, 1996), we predicted that the rated sensitivity of the grandparent would be positively related to the child's attention during the video chat interaction.

Method

Participants

Families with a child between the ages of 6 and 24 months were recruited from the Washington, D.C. metro area via flyers, listservs, community events, and word of mouth. Recruitment was systematically

dispersed across the 19-month range and across genders. All potential families were sent an electronic media usage survey, and only families who reported using video chat with their babies at least once a month were selected for the study. Twenty-five video calls to remote grandparents were observed (56% both grandparents, 36% grandmothers alone, and 8% grandfathers alone) throughout the year in 2014. Babies had a mean age of 16.12 months ($SD = 5.27$), just under half were female (48%), and most were White (96%) and from higher income families (54% from families making \$150,000 or more per year). Mothers were always physically present with the baby during the video calls, and fathers joined them in half of the sessions (52%).

All families were regular video chat users, with 76% reporting that they used video chat at least once a week and 24% reporting that they used it several times a month with their babies. All babies in the sample had interacted with the remote grandparent both in-person and via video chat before. Most of the babies (68%) video chatted with this specific grandparent between 1 and 3 times per week, whereas 20% video chatted with them less often (1–3 times per month) and 12% video chatted with them more often (every other day to every day). About half the babies in the sample (56%) had seen their remote grandparent in-person within the last 6 weeks, whereas the other half had last seen the grandparent anywhere between 2 and 6 months before the video call.

Procedure

A seminaturalistic observational method was employed for the study. Two researchers went to each baby's home when the family was already planning to use video chat with their child using either a mobile device or a laptop. Families were instructed to interact as they typically would during a video chat session with grandparents. The researchers recorded the interaction using small, hand-held video cameras while standing unobtrusively outside the interaction space. One researcher recorded the screen, while the other researcher recorded the physically present family members. Afterwards, the two video recordings of the video call were synced side by side and coded for behaviors of interest.

Behavioral Coding

Two coders were trained to identify behaviors of interest using pilot videos for the study. Both then coded 24% of the recordings for reliability purposes,

and inter-rater reliability using intraclass correlations (ICC) was calculated for each behavior. Any behavior that did not appear in at least six video calls was dropped from all further analyses. After reliability was reached, each coder was responsible for coding 50% of the recorded video calls.

Following established definitions (e.g., Butterworth, 2004), JVA was operationally defined as one individual directing another individual's attention to a third object, person, or event ($ICC = .70$). Each attempted instance of JVA was coded as an individual unit and was further labeled by subcategories, including the type of JVA (across-screen or within-screen), who was involved (and each participant's role as initiator or directee), and whether or not it was successful.

Screen-mediated JVA, however, has not been examined in previous research; therefore, we named and defined two primary types of screen-mediated JVA as follows (see Figure 1):

1. *Within-screen JVA*: The JVA initiator directs a partner's attention to an object or person on *his own side* of the screen (e.g., a baby brings a toy to the screen to show the grandparent; a grandmother shows the baby a cucumber growing in the grandmother's front garden). $ICC = .77$.
2. *Across-screen JVA*: The JVA initiator directs a partner's attention to an object or person on *the other side* of the screen (e.g., the baby points to her grandmother's dog; a grandfather asks about the snack the baby is eating). $ICC = .91$.

Two other types of screen-mediated JVA, in which users reference a person or object that is off screen, were also coded, but they were statistically rare and are not reported.

Successful JVA was coded when the person whose attention was being directed did indeed direct his or her attention to the object or person intended by the initiator ($ICC = .70$).

The *sensitivity of the target grandparent* ($ICC = .92$) was assessed using an adapted version of the *Parental Sensitivity Scale* from *The Emotional Availability Scales, abridged Infancy/Early Childhood Version* (Biringen, Robinson, & Emde, 2000). The scale incorporates a variety of indicators, including the adult's accuracy in reading infant signals, appropriate responsiveness to such signals, adult affect, awareness of timing, and flexibility. Coders continuously rated sensitivity on a 1–9 scale while viewing the video recorded interactions, using a joystick, an input device made of a pivoting stick atop a fixed



Figure 1. Two types of video-mediated JVA. Left panel: a grandmother shows the baby an apple, initiating *within-screen JVA*. Right panel: a baby points at an object on the laptop screen, initiating *across-screen JVA*.

base, and joystick monitoring software, which continuously recorded the angle of the joystick. The average value for each adult in each video call was used for analysis. In cases where there was more than one grandparent participating in the video call, coders selected one target grandparent on video chat to score on this measure, choosing the adult in each case who was rated as the most involved in the video call.

The *child's attention* ($ICC = .91$) to the screen was defined as the amount of time the child spent in screen-directed behaviors (e.g., gazing, gesticulating, vocalizing, showing things, or participating in activities facing toward the device) in each video call. The number of times the child turned his or her attention away from the video call and then returned his or her attention back to it was also noted.

The total frequency of each JVA behavior per video call was calculated. Then variables were created to demonstrate the occurrence of specific subcategories within each code, as a ratio between the total number of subcategory instances (e.g., successful JVA instances) and the total number of instances of the general code per call (e.g., total JVA instances). These ratios were converted to percentages for ease of interpretation, and standard deviations are provided in percentage points as the unit.

Results

General Video Call Characteristics

The average video call length was 20 min 18 s ($SD = 9$ min 28 s). When two outliers in call length (both longer than 40 min) were removed, the average call length was 18 min 16 s. Calls that were longer than the average call length were coded for approximately 20 min of their duration, starting

from the call's initiation. Babies paid attention for 41% of the video call length on average. There was no relationship between the babies' ages and the call lengths ($r = .25$, $p = .23$) or the percentage of the video call during which they paid attention ($r = -.11$, $p = .62$). Across the calls, grandparents displayed highly sensitive behavioral styles when interacting with the babies ($M = 7.68$, $SD = 0.80$).

Mobile devices were used most often to conduct the video calls (56% of calls: 32% touchscreen tablets, 24% smartphones), and laptops were used for all others (44%; two families projected the image from a laptop onto a large television screen, with the webcam placed on top of the television). There was no difference in the average age of children using mobile versus nonmobile devices, $t(23) = .25$, $p = .81$. Families mostly used Skype (60%) or FaceTime (36%) to conduct the video calls, and a single family used Google Hangouts. All families who used a mobile phone used FaceTime, whereas most families who used a laptop (82%) and most families who used a tablet (75%) used Skype.

Joint Visual Attention

JVA Characteristics During Video Chat

All 25 video calls included instances of JVA, and the mean number of JVA attempts per minute was 2.4 ($SD = 0.97$). On average, JVA attempts were fairly successful (84%; $SD = 11$), and the minimum success rate in any call was 63%. The average success rate of JVA did not differ between mobile ($M = 84\%$, $SD = 11$) and nonmobile ($M = 84\%$, $SD = 11$) device users, $t(23) = .02$, $p = .98$, or across families in which the grandparent and baby had seen one another in the last 6 weeks (56% of the sample; $M = 84\%$, $SD = 13$) and those who had not seen one in more than 6 weeks ($M = 83\%$, $SD = 8$),

$t(23) = .21, p = .84$. Surprisingly, age was related neither to the amount of total JVA per minute used across all members of the family ($r = -.13, p = .53$) nor its success rate ($r = .20, p = .34$).

As predicted, within-screen JVA was the most frequently used JVA type at 64% of total JVA instances per video call ($SD = 14$) and was the type used most often in 88% of the calls. Across-screen JVA was less frequent, occurred in 33% of total JVA instances per video call on average ($SD = 15$) and was the type of JVA used most often in only 12% of the calls. Within- and across-screen JVA usage (as percentages of total JVA) were ultimately near-complements of one another and were, thus, highly negatively correlated ($r = -.98, p < .001$). For this reason, individual inferential analyses examining JVA types are conducted using only across-screen JVA.

The degree to which the family members were involved in the JVA instances was surveyed, and the average level of participation of each individual is presented in Table 1. Babies were involved in nearly all instances of JVA (87%), mostly by having their attention directed (69%), whereas mothers and remote grandmothers were the primary initiators of JVA (initiating 35% and 30% of JVA instances, respectively, on average across the calls).

Tests of JVA Hypotheses

First, it was predicted that babies would initiate JVA of any kind more often as their ages increased. This hypothesis was partially supported: There was a marginally significant, positive association between

baby initiations of JVA and the babies' ages ($r = .37, p = 0.07$), such that babies initiated more of the total JVA instances as they got older.

Second, it was predicted that when babies of any age initiated JVA, they would be more likely to initiate within-screen JVA than across-screen JVA. This hypothesis was confirmed: A McNemar's test revealed that the percentage of babies that initiated at least one instance of within-screen JVA (80%) was significantly greater than the percentage who initiated at least one instance of across-screen JVA (28%, $p < .001$).

Third, it was hypothesized that older babies would be more likely to initiate at least one instance of across-screen JVA than younger babies. Using a median split (15 months), a chi-square test confirmed that older babies (16–24 months) were more likely to initiate across-screen JVA (46% initiated it at least once) than younger babies (6–15 months; 8% initiated it at least once), $\chi^2(1) = 4.42, p = .04, \phi = .42$. Due to the very small cell sizes (only 1 of the 12 younger babies initiated an instance of across-screen JVA), this result should be interpreted with caution. The number of babies in each age group who initiated each type of JVA is presented in Figure 2, and the distribution of ages across the types of JVA initiations is presented in Figure 3.

Exploratory Tests

The remaining two analyses in this section regarding JVA were exploratory in nature, so a limited number of questions were pursued, and statistical corrections were used where appropriate. Given the important supporting role of the physically present parent during video chat with babies (McClure & Barr, 2016), we explored whether mothers themselves demonstrated a tendency toward within- or across-screen JVA initiations and how this would relate to their babies' JVA behaviors. First, we examined whether mothers initiated across-screen JVA or within-screen JVA more to direct their babies' attention. A chi-square test revealed no difference in the number of mothers who initiated across-screen JVA more (56%) or within-screen JVA more (44%) when directing their babies' attention, $\chi^2 = 0.36, p = .55, \phi = .12$.

Second, the present data offered a unique opportunity to examine patterns of mother-initiated and baby-initiated JVA types. Mothers initiated within-screen (51%) and across-screen (48%) JVA at equal rates, whereas babies, as demonstrated above, tended to initiate within-screen JVA more often.

Table 1
Average Participation and Role Type in Video Mediated JVA

	Total participation % (SD)	Initiations % (SD)	Being directed % (SD)
Babies	87 (17)	17 (14)	69 (18)
Mothers	55 (19)	35 (15)	11 (9)
Fathers	11 (17)	7 (11)	2 (4)
Remote grandmothers	43 (20)	30 (17)	9 (7)
Remote grandfathers	17 (18)	11 (12)	4 (5)

Note. Total participation does not add up to 100% for each individual for two reasons: (a) It also includes instances in which the individual assisted in JVA instances, and (b) Not all individuals (especially dads) were involved in all the JVA instances. Furthermore, whereas babies were involved in nearly all the JVA instances, sometimes JVA was exchanged between the adults only or, occasionally, with the baby's sibling. Standard deviations are reported in percentage points.

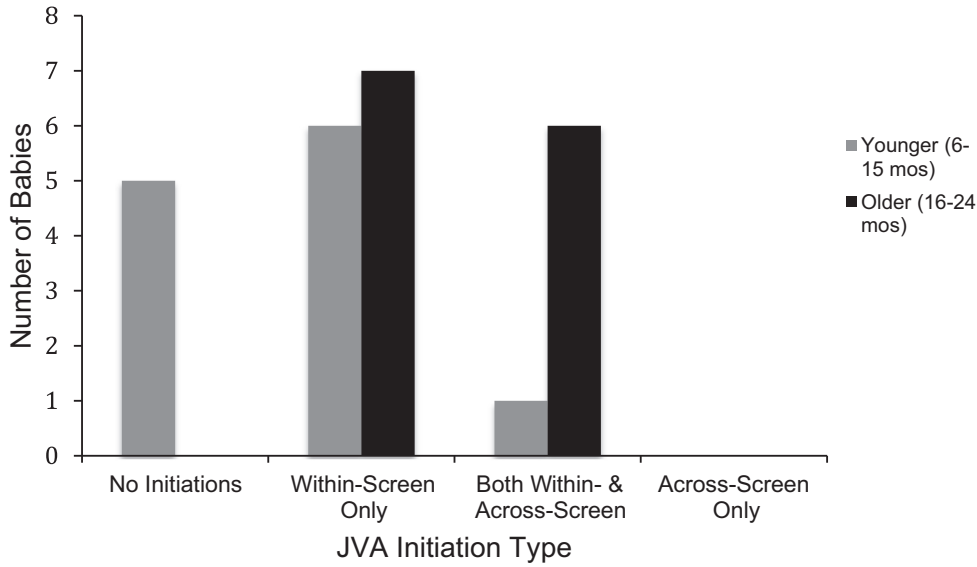


Figure 2. Number of babies who initiated at least one instance of within-screen and/or across-screen JVA, by age group.

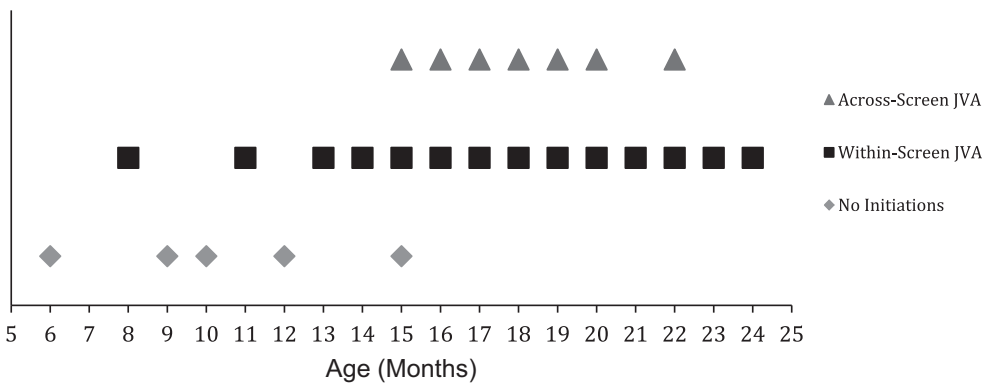


Figure 3. Distribution of baby ages for those who did not initiate any JVA and those who initiated at least one instance of within-screen and across-screen JVA.

Baby initiations of across-screen JVA, however, were positively correlated with maternal initiations of across-screen JVA ($r = .50, p = .02$). In other words, although mothers as a group appeared to have no overall preference toward one or the other JVA type, a mother’s and baby’s initiation types reflected one another: Babies whose mothers initiated more instances of across-screen JVA tended to do so more often also, and similarly for within-screen JVA.

Overall Attention

The studied behaviors were tested for their predictive power of the child’s overall attention to the video call, where attention was defined as the total duration of the baby’s attention to the call, divided

by the total coded duration of the call. Given the small sample size, only two predictors were chosen for inclusion in a multiple regression. The first predictor chosen was the type of screen-mediated JVA used—entered as the ratio of across-screen JVA to within-screen JVA (its complement) usage during the video call, with the caveat that the baby’s initiations of across-screen JVA were not included because when babies initiate across-screen JVA they are already attending to the screen. The second predictor chosen was the remote grandparent’s sensitivity, because the contingency of adults while interacting with babies has been shown to relate to infant attention (e.g., Hains & Muir, 1996), and our measure of sensitivity includes contingency as an important indicator. Both were expected to be significant predictors of the child’s attention. The

results of the regression are presented in Table 2. The regression model had a good fit to the data, $F(2, 24) = 6.37, p < .01$, and accounted for 37% of the variance in babies' attention rates. Both the usage of across-screen JVA ($\beta = .40, p = .03$) and the remote grandparent's sensitivity ($\beta = .36, p = .05$) were significant predictors of the babies' attention. A second regression with the same predictors was performed, this time using the babies' distractibility—the number of times the baby's attention was lost and then returned to the video call—as the dependent variable. This regression was not a good fit to the data.

Discussion

Prior to this study, there was no existing research examining natural video chat use by babies under two in their own homes. These video chat calls between remote grandparents, parents, and their babies involved rich, child-focused social interactions that were facilitated by JVA. Furthermore, *across-screen JVA*, a term coined in this paper, had not to our knowledge been reported regarding toddlers until now. It was therefore unclear whether babies would be capable of using JVA successfully in the video chat context. This study intended to bridge this gap by examining how babies and families responded to the technical and developmental challenges of sharing attention using video chat at home. Overall, JVA attempts among these families were remarkably successful, notwithstanding the babies' ages and the challenges that video chat presents. This surprising result suggests that the adults are modulating JVA attempts well and responding appropriately to their babies' developmental needs. This explanation is supported by the high level of rated sensitivity of the grandparents ($M = 7.68, SD = 0.80$) in this sample. For example, a sensitive grandparent may use across-screen JVA as a way to

demonstrate interest in objects that are nearby or that the baby is already engaged with in some form—like the snack on the baby's tray or the toy that is already in the baby's hands. This avoids the kinds of challenging, and potentially unsuccessful, scenarios that would occur if grandparents were referencing objects behind the babies. Future research should examine the features of the object of attention more closely.

This study contributes three major findings to the existing literature. First, it situates the development of screen-mediated JVA within the timeline of general JVA development. These cross-sectional changes are suggestive of a potential age-related developmental trend: The youngest babies rarely if ever initiated JVA of any kind during the video calls, the majority of babies initiated only within-screen JVA, and older babies were almost exclusively the ones who added initiations of across-screen JVA to their repertoire. This trend is in agreement with the hypothesized difficulty level of these two types of video-mediated JVA, based on the existing literature on JVA development in babies.

Second, although mothers in general initiated within-screen and across-screen JVA equally often, mothers of babies who initiated more across-screen JVA also tended to initiate it more often themselves. This finding is correlational, so the direction of this relation is unclear. Mothers may be modeling JVA initiations for their babies, who imitate their behavior, or mothers may be imitating their babies' initiations of JVA to encourage them. It is also possible that the relation may be bidirectional, with babies and mothers cyclically responding to the initiations of the other.

Third, the amount of across-screen JVA used and the grandparent's sensitivity were both significant predictors of the baby's attention rate during the video calls. Again, this result cannot be interpreted causally, so the relation between across-screen JVA and the baby's attention remains unclear. More use of across-screen JVA may lead to more infant attention, more engaged babies may provide more opportunities for using across-screen JVA, or there may be a bidirectional relation between them. These variables do not predict the babies' distractibility, however, suggesting that across-screen JVA and grandparent sensitivity are not simply being used as tools to bring highly distracted toddlers' attention back to the calls. Future research will be able to untangle these explanations by systematically manipulating the behaviors of remote partners.

This study had several limitations. First, the sample was small and consisted of a fairly

Table 2
Grandparent's Sensitivity and Across-Screen JVA as Predictors of Baby Attention

	B	SE	β
Intercept	− 0.46	.33	
Across-screen JVA	0.55*	.24	.40*
Grandparent's sensitivity	0.09*	.05	.36*
R^2	0.37		
$F(2, 24)$	6.37**		

* $p \leq .05$. ** $p \leq .01$.

homogeneous group of educated families, so the results may not be widely generalizable. Second, it is likely that patterns of face-to-face interactions among parents, grandparents, and children will influence the success of video chat interactions. Future research should test this empirically and help determine the role that video chat plays in maintaining long-distance relationships. Third, although every attempt was made by the researchers to remain unobtrusive during the video calls, it is still possible that observer effects modified the behavior of the participants. Fourth, the naturalistic methods employed led to several coding challenges: Many parents and children moved around during the call, and those with mobile devices also moved their devices frequently; both of these issues made the interactions difficult to film and some behaviors could therefore not be coded. Future research should systematically examine how video chat differs as a function of smaller mobile devices versus larger more stable devices. More generally, future studies may add to this one by examining screen-mediated JVA behaviors under more controlled conditions, either in the laboratory or at home with greater restrictions placed on families' natural movements (e.g., request that the baby video chat from a high chair). This would also allow researchers to control other important factors related to screen-mediated JVA, like the number of child-friendly or interesting items visible to the child on each side of the screen, and help disentangle factors like screen-size and device mobility.

The present study has important implications for research on child development, and future studies should consider the role that children's video chat

exposure may play in the outcomes of their studies (Troseth, Russo, & Strouse, 2016). JVA is part of a constellation of sociocognitive abilities that include language, pretend play, perspective taking, and ToM (Baron-Cohen, 1991; Butterworth, 2001), commonly defined as the ability to attribute mental states to oneself and others (Premack & Woodruff, 1978). More advanced early JVA predicts subsequent vocabulary and cognitive development (Brooks & Meltzoff, 2008; Strid, Tjus, Smith, Meltzoff, & Heimann, 2006), and most recently Brooks and Meltzoff (2015) have demonstrated a longitudinal link between JVA, language development, and ToM. Studies have shown that practice with perspective taking, which video chat may offer, may speed the developmental trajectory of ToM (Lieberman, Woodward, & Kinzler, 2017) and that media exposure can affect children's symbolic understanding (Troseth, 2003). It is possible, then, that video chat exposure during the early years of a child's life may affect the development of these abilities.

This raises several important questions for future examination. In the present study, for example, a 23-month-old boy was observed several times pointing a smartphone screen toward objects and toys to show them to his remote grandmother (see Figure 4). His actions demonstrated a burgeoning ability to understand what his remote grandmother was capable of seeing through the built-in camera. Future studies should examine how understanding what another individual can see relates to understanding what can be seen by that individual through a camera and whether their developmental trajectories are similar or distinct. Furthermore, this child's mother frequently used mental state



Figure 4. Baby turns phone screen to show his grandmother a toy airplane (left panel), then turns the phone to show her a block house he is working on (right panel).

language to help the child understand the grandmother's perspective (e.g., "She can't see it there. Turn her this way so she can see."), thus creating an environment that supports the development of ToM (Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991). Video chat may provide a unique opportunity to hone early JVA and sociocognitive abilities, and subsequent research should investigate whether additional opportunities using mind-minded language and practicing perspective taking during video chat alter developmental trajectories of children who frequently participate in these interactions.

Furthermore, future studies should also explore whether using video chat as a delivery method of JVA-associated (e.g., language, ToM) stimuli affects their outcomes. In at least one seminal ToM study, a false belief task (in which children viewed an object being hidden) was presented to 3-year-olds via noncontingent video (Moses & Flavell, 1990). The 3-year-olds in this study did poorly on the test, but it is not clear whether these results would have been different if the hiding scenario were demonstrated via video chat. Research has demonstrated that toddlers succeed more often in retrieval tasks when the object's hiding location is delivered via live video feed than via noncontingent video (Troseth, Saylor, & Archer, 2006), which may suggest that the results of Moses and Flavell's (1990) study may have been influenced by their delivery method. A replication with a video chat condition would be informative.

Babies today are immersed in a digital world in which formative relationships with important family members like grandparents are often mediated by a screen. For some children, well-established developmental processes, including the development of JVA, ToM, and other social-emotional communication skills, are now being influenced by frequent, lengthy exposure to video chat interactions. The results of this study indicate that babies' media exposure and developmental processes can no longer be studied in isolation from one another and suggest a critical new lens through which to examine development.

References

- American Academy of Pediatrics: Council on Communications and Media. (2016). Media and young minds. *Pediatrics*, *138*, e20162591. doi: 10.1542/peds.2016-2591
- Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, *55*, 1278–1289. doi:10.2307/1129997
- Ballagas, R., Kaye, J. J., Ames, M., Go, J., & Raffle, H. (2009). Family communication: Phone conversations with children. *Proceedings of the 8th international conference on interaction design and children* (pp. 321–324). ACM Press. doi: 10.1145/1551788.1551874
- Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 233–251). Oxford, UK: Basil Blackwell.
- Betts, J., Mckay, J., Maruff, P., & Anderson, V. (2006). The development of sustained attention in children: The effect of age and task load. *Child Neuropsychology*, *12*, 205–221. doi:10.1080/09297040500488522
- Biringen, Z., Robinson, J. L., & Emde, R. N. (2000). The emotional availability scales (3rd ed.; an abridged Infancy/Early Childhood Version). *Attachment & Human Development*, *2*, 256–270.
- Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, *35*, 207–220. doi:10.1017/s030500090700829x
- Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology*, *130*, 67–78. doi:10.1016/j.jecp.2014.09.010
- Butterworth, G. (2001). Joint visual attention in infancy. In J. G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 213–240). Malden, MA: Blackwell.
- Butterworth, G. E., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioural Development*, *3*, 253–272. doi:10.1177/016502548000300303
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367. doi:10.1146/annurev.psych.52.1.337
- Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development*, *62*, 1352–1366. doi:10.2307/1130811
- Hains, S. M. J., & Muir, D. W. (1996). Effects of stimulus contingency in infant-adult interactions. *Infant Behavior and Development*, *19*, 49–61. doi:10.1016/S0163-6383(96)90043-0
- Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*, 131–134. doi:10.1111/1467-9280.00024
- Kabali, H. K., Irigoyen, M. M., Nunez-Davis, R., Budacki, J. G., Mohanty, S. H., Leister, K. P., & Bonner, R. L. (2015). Exposure and use of mobile media devices by young children. *Pediatrics*, *136*, 1044–1050. doi:10.1542/peds.2015-2151

- Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). Exposure to multiple languages enhances communication skills in infancy. *Developmental Science*, 20(1). Advance online publication. doi: 10.1111/desc.12420
- McClure, E. R., & Barr, R. (2016). Building family relationships from a distance: The role of communication technology. In R. F. Barr & D. Linebarger (Eds.), *Media exposure during infancy and early childhood: The effect of content and context on learning and development* (pp. 227–248). Cham, Switzerland: Springer.
- McClure, E. R., Chentsova-Dutton, Y. E., Barr, R. F., Holochwost, S., & Parrott, W. G. (2015). "Facetime doesn't count": Video chat as an exception to media restrictions for infants and toddlers. *International Journal of Child-Computer Interaction*, 6, 1–6. doi:10.1016/j.ijcci.2016.02.002
- Moses, L. J., & Flavell, H. (1990). Inferring false beliefs from actions and reactions. *Child Development*, 61, 929–945. doi:10.2307/1130866
- Myers, L. J., LeWitt, R. B., Gallo, R. E., & Maselli, N. M. (2016). Baby FaceTime: Can toddlers learn from online video chat? *Developmental Science*. Advance online publication doi:10.1111/desc.12430
- Premack, D. G., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1, 515–526. doi:10.1017/S0140525X00076512
- Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, 85, 956–970. doi:10.1111/cdev.12166
- Strid, K., Tjus, T., Smith, L., Meltzoff, A. N., & Heimann, M. (2006). Infant recall memory and communication predicts later cognitive development. *Infant Behavior and Development*, 29, 545–553. doi:10.1016/j.infbeh.2006.07.002
- Tarasuik, J., Galligan, R., & Kaufman, J. (2013). Seeing is believing but is hearing? Comparing audio and video communication for young children. *Frontiers in Psychology*, 4, 1–6. doi:10.3389/fpsyg.2013.00064
- Tarasuik, J., & Kaufman, J. (2017). When and why parents involve young children in video communication. *Journal of Children and Media*, 11(1), 88–106. doi: 10.1080/17482798.2016.1233124
- Troseth, G. L. (2003). TV guide: Two-year-old children learn to use video as a source of information. *Developmental Psychology*, 39, 140–150. doi:10.1037/0012-1649.39.1.140
- Troseth, G. L., Russo, C. E., & Strouse, G. A. (2016). What's next for research on young children and interactive media? *Journal of Children and Media*, 10(1), 54–62. doi:10.1080/17482798.2015.1123166
- 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, 786–799. doi:10.1111/j.1467-8624.2006.00903.x