

Understanding the impacts of video activity guides on parent-child interaction

Abstract

Early parenting practices play an important role in shaping the future outcomes of young children (Hart & Risley, 1995; Heckman, 2006). In particular, high-quality early interactions and language input appear to facilitate language learning and result in higher levels of school performance. The rise of phone- and tablet-based parenting applications (“apps”) holds the promise of delivering low-cost, positive interventions on parenting style to a wide variety of populations. Of special interest are the parents of very young children, who are often difficult to reach in other ways. Yet little is known about the effects of communicating to parents through app-based interventions. In a study of one commercial app offering a collection of age-appropriate activity videos, we find that the quality of parent-child interactions increases in some ways as a result of using the app. Specifically, the lexical diversity of parents’ child-directed speech increases, and measures of joint attention show...

Introduction

The quantity and quality of early language input has been found to be strongly associated with later language and academic outcomes (Cartmill et al., 2013; Hart & Risley, 1995; Hirsh-Pasek et al., 2015; Marchman & Fernald, 2008). Thus, because of the potential for large downstream effects (Heckman, 2006), there is tremendous interest in interventions that change children’s language environment. And because parents define a large portion of that environment, especially before the onset of formal schooling, parent behavior is a critical locus for such interventions. Many effective parenting interventions require large resource investments and require many hours of in-person contact (Gertler et al., 2014; Schweinhart et al., 2004), making implementation at scale a daunting proposition. For this reason, many researchers targeting early language are interested in delivering parenting interventions remotely – through texts, apps, and videos delivered on digital devices. But what do parents take away from these short messages about what to do or how to talk with their children?

The content provided by digital parenting interventions runs the gamut from general parenting messages and facts from child development research to specific advice and suggested activities. A growing body of evidence suggests that these digital interventions can be effective across a range of cultures, income levels, and children’s ages (for a review, see Breitenstein, Gross, & Christophersen, 2014). For example, in contrast to a face-to-face parent training intervention, a tablet-based version saw significantly higher session completion rates (51% attendance vs. 85% module completion) and comparable or larger effect sizes on parents’ and children’s (aged 2 to 5 years) behavior (Breitenstein, Fogg, Ocampo, Acosta, & Gross, 2016). Often, however, the theory of change presupposed by such interventions is relatively vague. Both within and outside the realm of academic interventions, messages to parents of young children often seek to provide knowledge about some aspect of development (e.g., early language), often in tandem with a suggestion regarding activities. Such messages are assumed to inform parents’ choice of behaviors, spurring them to engage in some target activity, which is assumed to be more stimulating than what parents would have done otherwise.

This theory of change is typically grounded in ideas about guided play and early language stimulation. Child-directed speech varies not only in quantity (i.e., the number of total tokens), but also in quality in terms of the diversity of the tokens (Malvern, Richards, Chipere, & Durán, 2004) or the context-appropriateness of the speech (Cartmill et al., 2013), both of which have been linked to children’s subsequent language development. Further, language learning – especially the acquisition of early vocabulary in the first years – appears to be supported preferentially by parents and children *jointly attending* to some object or activity (Baldwin, 1991; Bigelow, MacLean, & Proctor, 2004). Episodes of joint attention are frequent during guided play, when parents set goals and scaffold their child’s activities (Weisberg, Hirsh-Pasek, & Golinkoff, 2013; Wood, Bruner, & Ross, 1976). Thus, the current literature supports interventions that encourage parents to provide high-quality language and interaction through something like guided play – whether via reading books or playing with a shape-sorter at home, or via a conversation about categories in the supermarket.

But is this theory of change correct? That is, does the provision of knowledge and activities lead to higher-quality play? An alternative hypothesis is that this theory could be wrong in a number of ways. By focusing parents on a specific activity, they could over-focus on achieving the superficial goals of the activity. This problem might be especially likely with video messages, which could encourage parents to try to mimic a model. And these misinterpretations in turn could result in less high-quality talk, with less joint attention. Or, these messages could produce the desired effect, but only for those parents who already have a general orientation towards children’s early learning.

Our current experiments were designed to make a direct test of this question: How do parents change their interactions with young children on the basis of short video parenting messages? In both studies, we collected data from parent-child dyads in a local children’s museum. We showed parents in the experimental group a single short video modeling an interactive toy-based activity along with a scientific justification. Parents in the control group received either no video (Experiment 1) or a video of a recent finding in developmental psychology (Experiment 2). We then gave the toys from the video to all dyads and videotaped their interactions, coding for language quantity and quality as well as joint attention.

But does providing such information change parents’ behavior towards young children? And is this advice effective in producing change along the desired dimension? An alternative might be that asking parents to focus on a specific activity that may be novel to them may tax their cognitive resources, resulting in decreases in the quality of their child-directed speech, or in less flexible, lower-quality social interactions.

What are digital interventions With the widespread use of smartphones and tablets worldwide, digitally-delivered interventions could address many of the logistical barriers that have limited scaling up face-to-face delivery methods. However, the parent and child outcomes assessed in the review (e.g. infant positive behaviors, satisfaction, emotional symptoms etc.) did not address the nature or quality of parent-infant interactions at a detailed level, for example whether the interventions lead to children paying more attention, or vocabulary changes in parents’ language usages. Although digitally-delivered activities are designed to promote learning and cognitive development, it is unclear how they might affect these dimensions of parent-child interactions. Thus, we want to conduct this experiment to explore if and how digital scaffolding of activities affect the social and linguistic characteristics of parent-child interactions. The quality of parent-child interactions can be measured by both the social engagement of parents (e.g., joint attention to objects in the environment) (Bigelow et al., 2004) and the quality of language (e.g., vocabulary diversity) (Malvern et al., 2004).

Communication and message content

In our study, we focus on videos that provide suggested play activities to parents. These videos constitute a short-term “guided play” intervention.

Young children spend a large portion of their waking time at play, variously manipulating objects, exploring their environment, and interacting with caregivers and peers. Playing with objects allows them to discover hidden object properties and relations, and to build a causal understanding of how objects interact (e.g., Schulz & Bonawitz, 2007). Meanwhile, play also gives children an opportunity to set and achieve goals (e.g., build a tower) and to practice a wide range of motor skills (e.g., stacking) that will help them navigate the world (Singer, Golinkoff, & Hirsh-Pasek, 2006). Social play can help children learn about human relationships, both through imitation of adult behaviors and by experiencing and learning to process emotional events such as failures (Singer et al., 2006). Of course, young children are rarely playing in isolation: caregivers often provide encouragement and guidance while scaffolding a child’s play (Kaye, 1970; Wood et al., 1976). The quality of interactions during such guided play has been shown to influence language learning: parents’ joint attention to objects that their child was focused on was positively correlated with the child’s subsequent vocabulary growth (M. Carpenter & Tomasello, 1998; Tomasello & Farrar, 1986). Episodes of joint attention during guided play have also been found to contain more age-appropriate advanced forms of play (Bigelow et al., 2004). More generally, parenting practices early in childhood have been shown to play an important role in shaping the future outcomes of young children (Hart & Risley, 1995; Heckman, 2006). While interventions often have trouble reaching many parents of very young children, the proliferation of mobile devices offers a good avenue for the digital delivery of parenting advice (Breitenstein et al., 2014). However, the efficacy of digital parenting advice has not been widely proven.

Guided Play Scaffolds Learning

Children’s early play behaviors are often assisted by more skilled and knowledgeable play partners such as their caregivers and older siblings (Kaye, 1970). Under such expert guidance, children are encouraged and motivated to engage in more advanced play, undertaking explorations that push the boundaries of what they would be able to do unaided (Vygotsky, 1980). These tutorial interactions have been shown to be important components of child development (Wood et al., 1976). Thus, with the knowledge of both play and tutorial interactions, guided play, which consists of both active and enjoyable activities as well as close guidance of adults (Hirsh-Pasek & Golinkoff, 2008) has drawn researchers’ interest. A study of preschoolers showed that guided play scaffolds the environment while still allowing children to maintain a large degree of control, and it outperforms direct-instruction approaches in encouraging a variety of positive academic outcomes (Weisberg et al., 2013). Another study found that guided play could facilitate children’s vocabulary and comprehensive language development and subsequent literacy skills (Massey, 2013).

Improving Parenting Practices and Language Use

Thus, including guided play in parenting practices from an early age may boost children’s language and educational outcomes. A recent home-visit parenting practices intervention targeting children of low socioeconomic status found that parents in the intervention group gained knowledge of language development, and that this effect sustained four months after the intervention (Suskind et al., 2015). However, although a number of interactive measures increased during the experiment, including the number of word tokens, conversational turns, and child vocalizations, these increases did not sustain after the intervention. That changes did not sustain could be due to the intervention itself, or could merely be that the methods of home-visiting is not sustainable enough for parents to easily and constantly get parenting advice. Thus, new methods of delivering parenting advice should be considered.

Effectiveness of Digital Delivery

Experiment 1

In Experiment 1, we invited parents of 6- to 24-month-old infants visiting the Children’s Discovery Museum in San Jose to complete activities from the Kinedu app. Parents were randomly assigned to the video group or the control group; parents in the video group watched a video from the Kinedu app (matched to their child’s age), and then performed the activity with their child using the props from the video. Parents in the control group did not watch a Kinedu video, instead they were given the same props and were told to play with their infants as they would at home.

Method

Participants.

60 infants (F = 42, M = 18) aged 6-24 months (20 6-11.9 month-olds, 20 12-17.9 month-olds, and 20 18-24 month-olds) and their parents participated in a museum in northern California. We included infants who were exposed to English at least 50 percent of the time (n = 58) or who were exposed less but whose participating parent reported that they primarily speak English with their child at home (n = 2). 61% of participants (n = 37) had been exposed to two or more languages as indicated by their parent. Parents identified their children as White (n = 25), Asian (n = 11), African American/Black (n = 2), Biracial (n = 12), other (n = 5), or declined to state (n = 5). Fifteen parents reported that their child was of Hispanic origin. Parents tended to be highly-educated, with reports of highest level of education ranging from completed high school (n = 5), some college (n = 6), four-year college (n = 14), some graduate school (n = 1), to completed graduate school (n = 28) or declined to state (n = 6).

Materials.

Stimuli included videos from the Kinedu (Kinedu Inc.) commercial parenting application. The videos were designed to show activities to parents that they could perform with their child in order to foster cognitive and physical development, and were targeted to the child's age and level of development. In each video, an adult and child perform the activity while a narrator explains the activity and its purpose. We selected two videos for each of three age groups in our sample (6-11.9 months, 12-17.9 months, 18-23.94 months). More information about the specific videos is available in the Appendix. Participants were also given a set of toys corresponding to those in the video that they watched so that they could complete the activity. The toys associated with each video are listed in the Appendix.

Participants were randomly assigned to either the *Video* condition or the *Control* condition. Parents participating in the Video condition were assigned to watch one of the two activity videos available for their child's age group, while parents in the Control condition watched no activity video, and were simply asked to play with their child as they normally would. The Control condition was yoked to the Activity Video condition such that for every participant in the Video condition who saw a particular video and received the associated props, a participant in the Control condition received the same props but did not watch the activity video. Parents also completed the Early Parenting Attitudes Questionnaire (EPAQ; (Hembacher & Frank, 2018)). The EPAQ measures parents' attitudes about parenting and child development along three dimensions: rules and respect, early learning, and affection and attachment.

Procedure.

After providing informed consent, parents in the Video condition watched the assigned activity video on a laptop with headphones. To ensure that parents could give the video their full attention, the experimenter played with the infant with a set of toys (different from the experimental props used in the study) while the video was being played. Immediately following the video, each parent-child dyad was provided with the props to complete the activity the parent had viewed. The toys were placed on a large foam mat, and parents were instructed to sit on the mat with their child and re-create the activity they had viewed for a period of three minutes. In the Control condition, after consenting parents were told to play with their child as they would at home with the provided props for a period of three minutes. They were not given any additional instructions about how to use the props.

In both conditions, two video cameras were used to record the play session from different angles, and parents were fitted with a wireless microphone (Shure lavalier microphone) to record their child-directed speech. After three minutes of play had elapsed, parents were told they could stop playing and the cameras and microphone were turned off. Parents were then asked to complete the EPAQ before being debriefed.

Joint Attention Coding Procedure.

The video of each session was manually coded for episodes of joint attention (JA) using the Datavyu software (Team, 2014). The video taken at floor level was coded by default, but the other video was referred to if the participants were not visible or if there was technical difficulty with the first camera. Each session's video was coded for episodes of coordinated JA, episodes of passive JA, and parental bids for JA. Parental bids for JA were defined as any attempt to initiate joint attention (i.e. labeling, pointing, or otherwise drawing attention to an object) that did not result in passive or coordinated JA. If more than 3 seconds elapsed between bids, they were coded as separate attempts. An episode of joint attention was considered passive if both participants visually focused on an object for a minimum of 3 seconds but the child did not acknowledge the parent. If either participant looked away from the object for less than 3 seconds and then returned to the same object it was considered part of the same period of joint attention. A joint attention episode was considered coordinated if both participants visually focused on an object for a minimum of 3 seconds and at some point in the interaction the child indicated awareness of interaction with some overt behavior toward the parent such as looking at their face, gesturing, vocalizing, or turn-taking.

A second coder independently coded a third of the videos (i.e., 20 of the 60 videos, approximately equally distributed across ages) to establish reliability. The two coders had a reliability of $ICC = 0.80$ with 95% confident interval $(CI) = [0.57, 0.92]$ ($p < 0.05$) for number of parent bids for JA; $ICC = 0.20$ with 95% $CI =$

[-0.26,0.58] for number of passive JA episodes; ICC = 0.66 with 95% CI = [0.32,0.85] ($p < 0.05$) for number of coordinated JA episodes; ICC = 0.24 with 95% CI = [-0.21,0.61] for total duration of passive JA episodes, and ICC = 0.62 with 95% CI = [0.27,0.83] ($p < 0.05$) for total duration of coordinated JA episodes.

Results

The transcripts and hand-coded behavioral data was analyzed according to our preregistration¹. Below we first describe the lexical diversity results, followed by the joint attention results.

Lexical Diversity

Parents' child-directed speech during the play sessions was transcribed. For each transcript, the words were lemmatized using Honnibal (2017), and the word *types* (unique words) and *tokens* (total words) were then tallied and the type-token ratio (TTR) calculated as a measure of lexical diversity. Although TTR was our preregistered measure of lexical diversity, TTR is correlated with the length of a text, whereas the measure of textual lexical diversity (MTLD) is not (McCarthy & Jarvis, 2010). Thus, we also measure lexical diversity with MTLD, which is calculated as the mean length of sequential word strings in a text that maintain a given TTR value (here, .720).

We fit a mixed-effects linear regression predicting TTR as a function of condition, age (scaled and 0-centered), gender, and parent's education level with a random intercept per video using lme4 (Bates, Mächler, Bolker, & Walker, 2015). There was significantly lower TTR in the Video condition (mean: 0.32) than in the Control condition (mean: 0.43, $\beta = -.12$, $t(39.9) = 4.25$, $p < .001$). There were no significant effects of age, gender, or parent's level of education. A similar mixed-effects linear regression instead predicting MTLD also found significantly lower lexical diversity in the Video condition (mean: 15.7) than in the Control condition (mean: 21.9, $\beta = -11.67$, $t(43) = 3.08$, $p < .01$), with no other significant effects. Figure 1 shows the mean of each lexical diversity measure (TTR and MTLD) by condition.

We also conducted similar regressions predicting the number of word tokens and types, finding only a significant effect of condition on the number of word tokens ($\beta = 57.23$, $t(34.4) = 2.19$, $p < .05$), with parents using more words in the Video condition (mean: 225, 95% CI: [197,252]) than in the Control condition (mean: 165, 95% CI: [139,193]).

(Table with mean and SD of tokens, types, and TTR)

¹Preregistration: <https://osf.io/2bpdf/>

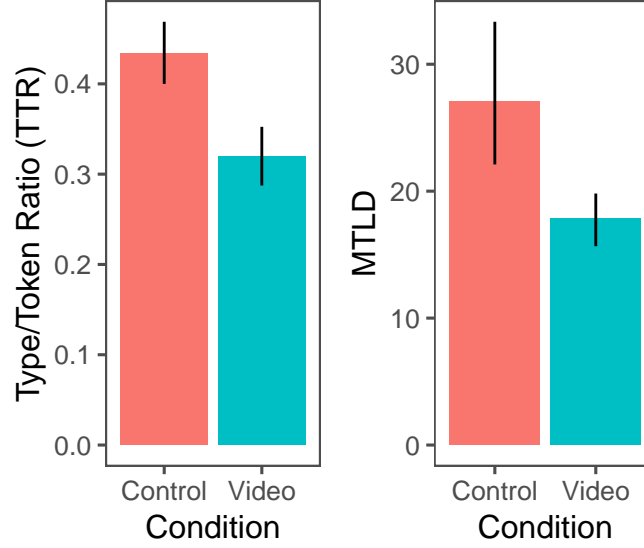


Figure 1: Mean lexical diversity scores by condition (left: Type/Token ratio, right: MTLD) in Experiment 1. Error bars show bootstrapped 95 percent confidence intervals (CIs).

Word tokens and word types

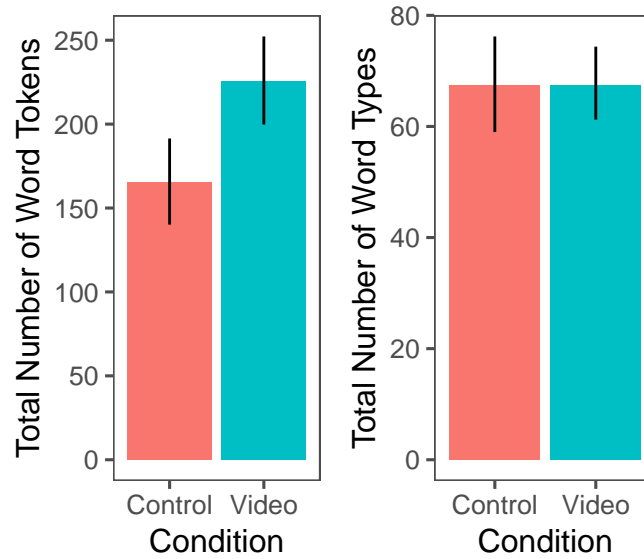


Figure 2: Mean number of word types and word tokens by condition in Experiment 1.

Joint Attention

We fit a mixed-effects linear regression predicting the number of bids for joint attention (JA) as a function of fixed effects of condition, age (scaled and 0-centered), gender, parent's education level, and the subscales of the EPAQ: Early Learning (EL), Affection and Attachment (AA), and Rules and Respect (RR), along with interactions of condition and EL, AA, and RR. This lme4 model included random intercepts per video. There were significantly more bids for JA in the Video condition (mean: 6.24, sd: 2.79) than in the Control condition (mean: 3.56, sd: 2.50, $\beta = 3.51$, $t(40.3) = 2.95$, $p < .01$). There were no other significant effects. Mixed-effects regressions with the same structure were performed predicting the number of episodes of coordinated and passive JA, and the total duration of time spent in coordinated and passive JA. There were no significant

effects on the number or total duration of coordinated JA episodes, nor on the total duration of passive JA episodes. For the regression predicting the number of passive JA episodes, the only significant effect was an interaction of condition and RR ($\beta = 1.83$, $t(41.5) = 2.22$, $p < .05$), showing that for parents in the Video condition, those with higher Rules and Respect subscores engaged in more passive JA episodes. Figure X shows ... by condition.

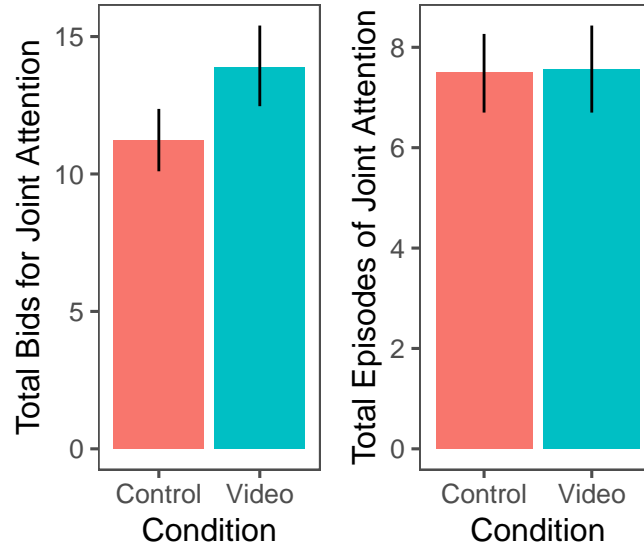


Figure 3: Mean number of bids and episodes of Joint Attention in Experiment 1.

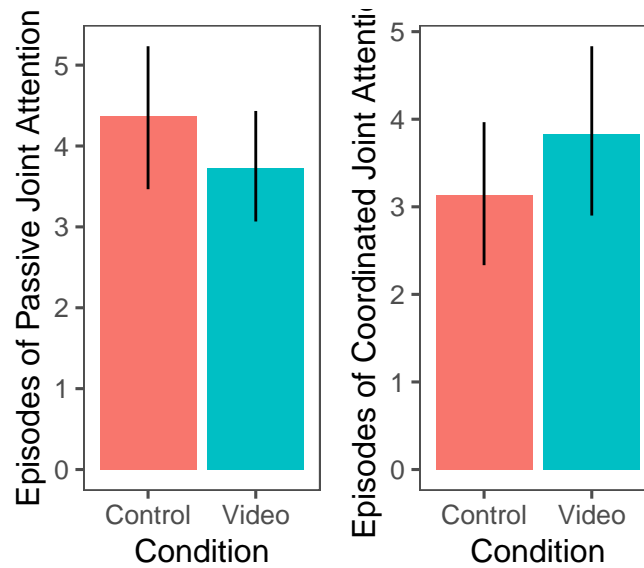


Figure 4: Average number of passive and coordinated episodes of JA in Experiment 1.

Discussion

In summary, while parents produced more word types and tokens after viewing the activity video, lexical diversity (Type/Token ratio and MTLT) was higher when parents were just asked to play as they normally would. This may suggest that parents in the Video condition are being more repetitive in their attempt to stick to the task prescribed in the video. Demographics and EPAQ do not interact with condition, but

there is a marginal effect of RR score on lexical diversity (lower diversity for higher RR scores), and marginal effects of parent education on word types and tokens (more types and tokens for higher parent education).

There was a main effect of condition on total bids for joint attention. Parents in the Video condition, after seeing a video demonstrating an activity, made a greater number of bids for joint attention with their child. There was no effect of condition on the number of episodes of either passive or coordinated joint attention, or the duration of these episodes. There was a marginal effect of gender on bids for joint attention, with parents of males producing more bids. There was a marginal interaction between RR scores and condition on passive joint attention, such that the experimental condition increased the number of episodes of PJA to a greater extent for people with high RR scores. While the electronically-delivered parenting advice increased the number of bids for joint attention by parents, it did not significantly effect the number or duration of episodes of joint attention. It may be that child variables had a larger relative impact on the attainment of joint attention.

Experiment 2

Experiment 1 found that parents who watched a Kinedu video spoke more words overall, but had lower lexical diversity compared to parents who played with their children as they normally would at home. Parents who watched a Kinedu video also made more bids for joint attention, although these bids did not result in more episodes of joint attention compared to the control group. Experiment 2 attempts to replicate these findings from with a restricted number of preregistered predictions ([link](#)).

Method

Participants.

83 infants ($F = 37$, $M = 47$) aged 12-24 months (42 12-17.9 month-olds, 42 18-24 month-olds) and their parents participated in the same museum as Experiment 1. We included infants who were exposed to English at least 75 percent of the time or who were exposed less but whose participating parent reported that they primarily speak English with their child at home. Forty nine% of participants ($n = 41$) had been exposed to two or more languages as indicated by their parent. Parents identified their children as White ($n = 39$), Asian ($n = 20$), African American/Black ($n = 1$), Biracial ($n = 9$), other ($n = 7$), or declined to state ($n = 8$). Sixteen parents reported their child was of Hispanic origin. Parents tended to be highly- educated, with reports of highest level of education ranging from completed high school ($n = 0$), some college ($n = 5$), four-year college ($n = 28$), some graduate school ($n = 2$), to completed graduate school ($n = 35$) or declined to state ($n = 14$).

Materials.

The design of Experiment 2 was similar to that of Experiment 1, except that instead of a No-Video control condition, parents instead watched a video that was generally related to child development research, but did not give any specific instructions about how to interact with infants or children. This was to control for the possibility that differences in language output and joint attention in Experiment 1 could be due to simply cuing parents to think about infants' learning and cognitive development. The videos presented in the Control Video condition were media clips (available on YouTube) of developmental psychologists explaining their research interleaved with footage of infants or toddlers engaged in developmental research studies. Thus, the content of the videos superficially matched those in the Activity Video condition, but did not suggest any particular activities. The videos were trimmed to approximately match the average video length in the Activity Video condition (close to 90 s).

Procedure.

The procedure for Experiment 2 matched that of Experiment 1, except that parents in the Control Video condition watched a control video before the play session. Consistent with the No-Video control condition in Experiment 1, parents in the Control Video condition were told to play with their child as they would at

home, and were not given additional instructions. The coding procedure also matched that of Experiment 1. A second coder independently coded a third of the videos (i.e., 26 of the 84 videos, approximately equally distributed across ages) to establish reliability. The two coders had a reliability of ICC = 0.80 with 95% confident interval (CI) = [0.60,0.90] ($p < 0.05$) for number of parent bids for JA; ICC = 0.74 with 95% CI = [0.59,0.87] ($p < 0.05$) for number of passive JA episodes; ICC = 0.78 with 95% CI = [0.58,0.90] ($p < 0.05$) for number of coordinated JA episodes; ICC = 0.72 with 95% CI = [0.46,0.86] ($p < 0.05$) for total duration of passive JA episodes, and ICC = 0.88 with 95% CI = [0.75,0.94] ($p < 0.05$) for total duration of coordinated JA episodes.

Results

Parents' child-directed speech was transcribed and processed according to the same procedure used in Experiment 1.

Lexical Diversity

We fit a mixed-effects linear regression predicting TTR and MTLD as a function of age (scaled and 0-centered) and condition with an interaction term, and with random intercepts per video using lme4 (Bates et al., 2015). There was significantly lower TTR in the Video condition (mean: 0.38) than in the Control condition (mean: 0.47, $\beta = -.09$, $t(8.7) = 3.06$, $p = .01$). There was no significant effect of age. A similar mixed-effects linear regression instead predicting MTLD found no significant effects of age or condition. Figure X shows the mean of each lexical diversity measure (TTR and MTLD) by condition. Regressions with the same structure predicting the number of word tokens and types found no significant effects of age or condition.

(Table with mean and SD of tokens, types, TTR, MTLD)

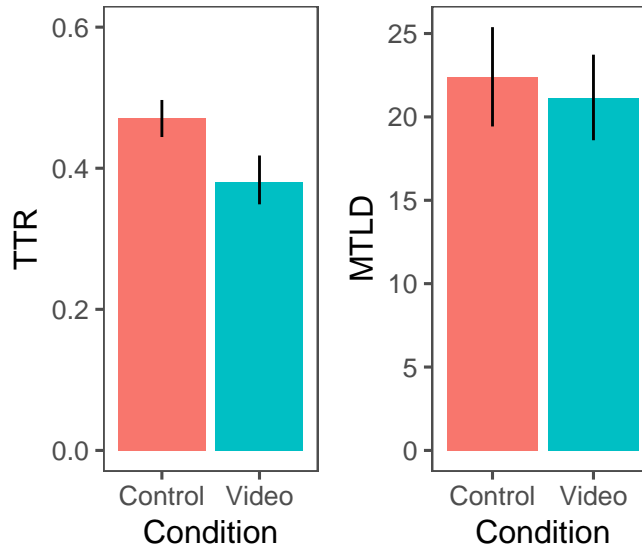


Figure 5: Mean lexical diversity scores by condition (left: Type/Token ratio, right: MTLD) in Experiment 2.

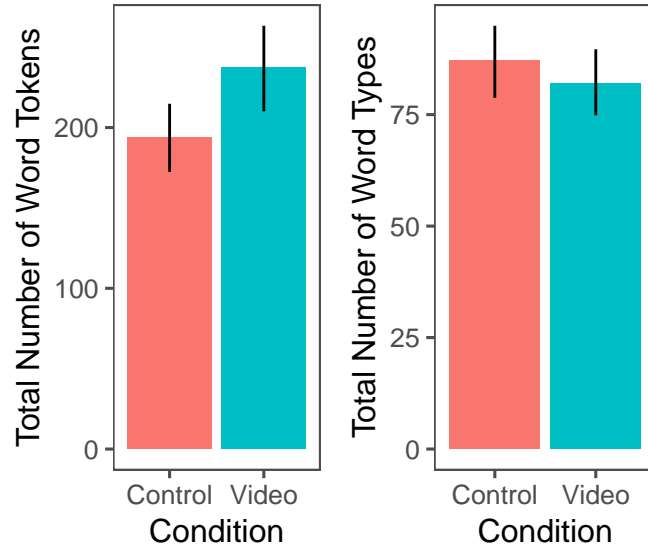


Figure 6: Mean number of word types and word tokens by condition in Experiment 2.

Joint Attention

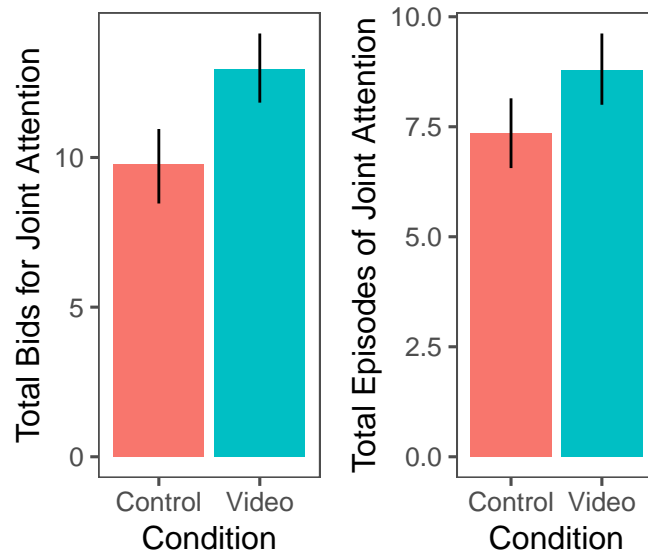


Figure 7: Mean number of bids and episodes of Joint Attention in Experiment 2.

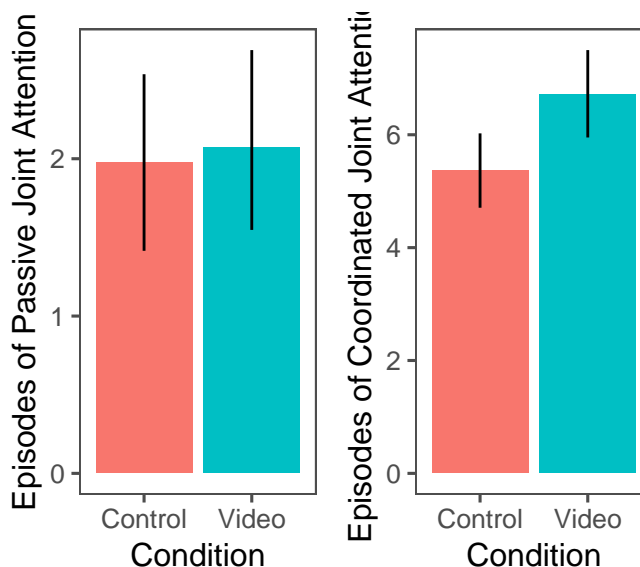


Figure 8: Average number of passive and coordinated episodes of JA in Experiment 2.

There was a marginal effect of condition on total bids for joint attention. Parents in the experimental condition (i.e., those who saw a video demonstrating an activity) made a greater number of bids for joint attention with their child. There was no effect of condition on the number of episodes of passive or coordinated JA, nor on the duration of these episodes. For passive joint attention, there was a main effect of age on both the number and duration of episodes, with older children having fewer episodes of passive JA and episodes of shorter duration. Moreover, older children had longer duration episodes of coordinated JA. These results suggest that as children age, they become more socially engaged in interactions with their caregivers.

Discussion

In both experiments, the number of tokens was higher in the experimental condition, while the number of types and lexical diversity (Type/Token ratio) were higher in the control condition. Parents may be relatively more repetitive in the experimental condition since they are attempting to stick to a specific prescribed task, but they talk more overall.

Acknowledgements

This work was supported by a gift from Kinedu, Inc. Thanks to members of the Language and Cognition Lab at Stanford for helpful discussion.

References

Appendix

Experiment 1

Video A (6-11.9 months) “Pick it up”

Parents are told to encourage their child to pick up and drop individual objects. They are also encouraged to place toys on a small cloth and show the child that they can drag the cloth towards them to reach the toys.

Props: cloth, plastic horse, plastic sheep, plastic elephant, toy car

Video B (6-11.9 months) “Animal sounds”

Parents are told to call different animals and imitate different sounds the animals make. They are also encouraged to observe which animal the child prefers.

Props: plastic sheep, plastic horse, plastic frog, plastic cow, bowls

Video C (12-17.9 months) “Give me the toy”

Parents are told to ask their child to hand over individual toys. They are also encouraged to praise the child after they give them the toys, and repeat the process until the child could follow the verbal instructions.

Props: toy boat, plastic frog, plastic elephant, toy bus

Video D (12-17.9 months) “Classifying my toys”

Parents are told to place toys of different sizes (big or small) in two hoops. They are also encouraged to ask their child to distinguish between two objects and identify which one is larger.

Props: two yellow and green rings, big car, small car, big horse, small horse

Video E (18-23.9 months) “My toys”

Parents are told to show the child toys of the same shape but different sizes, to place one of the objects in a basket and to ask the child to take out the object. They are also encouraged to ask their child if the object is bigger or smaller compared to its pair.

Props: two buckets, big car, small car, big horse, small horse

Video F (18-23.9 months) “The Orchestra”

Parents are told to give their child a musical instrument to play. They are also encouraged to play a song and see if the child follows the rhythm.

Props: maracas, drum, tambourine, clapper

Experiment 2

Video A (12-17.9 months) “Give me the toy”

Parents are told to ask their child to hand over individual toys. They are also encouraged to praise the child after they give them the toys, and repeat the process until the child could follow the verbal instructions.

Props: plastic pig, plastic horse, plastic dog, plastic cat, plastic cow

Video B (12-17.9 months) “Classifying my toys”

Parents are told to place toys of different sizes (big or small) in two hoops. They are also encouraged to ask their child to distinguish between two objects and identify which one is larger.

Props: two yellow and green rings, big car, small car, big horse, small horse

Video C (12-17.9 months) “Geometric shapes jigsaw puzzle”

Parents are told to encourage their child to name different shapes on a jigsaw puzzle. Then they are told to undo the puzzle and invite the child to complete the puzzle.

Props: A jigsaw puzzle of geometric shapes

Video D (18-23.9 months) “My toys”

Parents are told to show the child toys of the same shape but different sizes, to place one of the objects in a basket and to ask the child to take out the object. They are also encouraged to ask their child if the object is bigger or smaller compared to its pair.

Props: two buckets, big car, small car, big horse, small horse

Video E (18-23.9 months) “The Orchestra”

Parents are told to give their child a musical instrument to play. They are also encouraged to play a song and see if the child follows the rhythm.

Props: maracas, drum, tambourine, clapper

Video F (18-23.9 months) “My Yellow Toys”

Parents are told to show their child yellow toys and to ask “what color are they”. They are also told to give the child toys of different colors, ask them to only play with the yellow ones and praise the child after they do so.

Props: blue car, yellow car, yellow block, red block, blue block, green block

Baldwin, D. (1991). Infants’ contribution to the achievement of joint reference. *Child Development*, 62, 875–890.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <http://doi.org/10.18637/jss.v067.i01>

Bigelow, A. E., MacLean, K., & Proctor, J. (2004). The role of joint attention in the development of infants’ play with objects. *Developmental Science*, 7, 518–526.

Breitenstein, S. M., Fogg, L., Ocampo, E. V., Acosta, D. I., & Gross, D. (2016). Parent use and efficacy of a self-administered, tablet-based parent training intervention: A randomized controlled trial. *JMIR mHealth and uHealth*, 4(2), e36. <http://doi.org/10.2196/mhealth.5202>

Breitenstein, S. M., Gross, D., & Christophersen, R. (2014). Digital delivery methods of parenting training interventions: A systematic review. *Worldviews on Evidence-Based Nursing*, 11, 168–176.

Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences*, 110(28), 11278–11283. <http://doi.org/10.1073/pnas.1309518110>

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., . . . Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in jamaica. *Science*, 344(6187), 998–1001. <http://doi.org/10.1126/science.1251178>

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young american children*. Baltimore, MD: Brookes.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900–1902. <http://doi.org/10.1126/science.1128898>

Hembacher, E., & Frank, M. C. (2018). The early parenting attitudes questionnaire: Measuring intuitive theories of parenting and child development. *PsyArXiv*. <http://doi.org/10.31234/osf.io/hxk3d>

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., . . . Suma, K. (2015). The contribution of early communication quality to low-income children’s language success. *Psychological Science*, 26(7), 1071–1083. <http://doi.org/10.1177/0956797615581493>

Hirsh-Pasek, K., & Golinkoff, R. M. (2008). Why play= learning. *Encyclopedia on Early Childhood Development*, 1–7.

- Honnibal, I., Matthew AND Montani. (2017). SpaCy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. *To Appear*.
- Kaye, K. (1970). Mother-child instructional interaction. *Unpublished Doctoral Thesis, Department of Psychology, Harvard University*.
- Malvern, D., Richards, B. J., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development*. Palgrave Macmillan.
- Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11, F9–F16.
- Massey, S. L. (2013). From the reading rug to the play center: Enhancing vocabulary and comprehensive language skills by connecting storybook reading and guided play. *Early Childhood Education Journal*, 41, 125–131.
- M. Carpenter, K. Nagell, & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), i–vi, 1–143.
- McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-d, and hd-d: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392.
- Schulz, L., & Bonawitz, E. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, 43(4), 1045–1050.
- Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2004). *Lifetime effects: The highscope perry preschool study through age 40*. Ypsilanti, MI: HighScope Press.
- Singer, D. G., Golinkoff, R. M., & Hirsh-Pasek, K. (Eds.). (2006). *Play = learning: How play motivates and enhances children's cognitive and social-emotional growth*. New York, NY: Oxford University Press.
- Suskind, D. L., Leffel, K. R., Graf, E., Hernandez, M. W., Gunderson, E. A., Sapolich, S. G., . . . Levine, S. C. (2015). A parent-directed language intervention for children of low socioeconomic status: A randomized controlled pilot study. *Journal of Child Language*. <http://doi.org/10.1017/S0305000915000033>
- Team, D. (2014). Datavyu: A video coding tool. *Databrary Project*. Retrieved from <http://datavyu.org>
- Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 57(6), 1454–1463.
- Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*.
- Weisberg, D. S., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Guided play: Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education*, 7, 104–112.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17, 89–100.