CHILD DEVELOPMENT

Child Development, xxxx 2019, Volume 00, Number 0, Pages 1–20

Parents Influence the Visual Learning Environment Through Children's Manual Actions

Maureen E. McQuillan
Indiana University

Linda B. Smith
Indiana University and University of East Anglia

Chen Yu, and John E. Bates
Indiana University

The present research studied children in the second year of life ($N = 29$, $M_{age} = 21.14$ months, $SD = 2.64$ months) using experimental manipulations within and between subjects to show that responsive parental influence helps children have more frequent sustained object holds with fewer switches between objects compared to when parents are either not involved or over-involved. Regardless of parental involvement, sustained holds were visually rich, based on the size, centeredness, and dominance of the held object relative to other objects. These findings are important because they suggest not only that the child's body creates visually rich scenes across play contexts but also that a responsive parent can increase the frequency of these visually rich and informative moments.

Sustained attention is associated with advanced cognitive performance (e.g., Choudhury & Gorman, 2000) and is often measured as prolonged, intense focus on a toy during play, which importantly in early childhood involves manual engagement. A young child's manual engagement has been shown to reduce clutter in the visual scene by making the focal object the largest and most dominant in view. This process of visual selection via hand activity (Smith, Yu, & Pereira, 2011; Yoshida & Smith, 2008; Yu, Smith, Shen, Pereira, & Smith, 2009) has been shown to create optimal moments for learning (Pereira, Smith, & Yu, 2014; Yu & Smith, 2012). Child manual engagement also provides external cues to which parents can verbally respond (Bornstein, Tamis-Lemonda, Hahn, & Haynes, 2008; Karasik, Tamis-LeMonda, & Adolph, 2014; Tamis-LeMonda, Kuchirko, & Song, 2014; Tamis-LeMonda, Kuchirko, & Tafuro, 2013; Thurman & Corbetta, 2017). When a child leans toward a toy, grabs it with both hands, and brings it near to examine, a parent could infer that the child is interested in the toy and could label and teach the child about the toy. Once the child's hands have helped to select and stabilize the focus of attention, the child is likely well equipped to link an object's perceptual impressions with information about the object provided by the parent, such as its name.

How parents influence child attention is an important question. Experimentation provides a plausible way to address this question. The present study was designed to compare children's manual engagement, and corresponding visual scene properties, during toy play in conditions with experimentally manipulated degrees of parent involvement. Manual engagement with objects helps children discover the rich physical features of objects, including size, texture, and shape. Through exploration of these features, children learn to refine their actions and develop a sense of agency in dealing with their physical and social surroundings (Corbetta, Wiener, Thurman, & McMahon, 2018). Manual engagement also largely contributes to a child's first-person view of objects (Smith et al., 2011), which predicts object name learning during social play (Pereira et al., 2014). Not yet understood is how qualities of parent involvement might affect the dynamics of young children's manual

This work was funded by NIH grants R01 HD074601 and R21 EY017843. Maureen McQuillan was also supported by Institutional Training Grants from NICHD (HD007475) and NIMH (MH103213). We gratefully acknowledge the contribution of the research assistants, parents, and children involved in this research.

Correspondence concerning this article should be addressed to Maureen E. McQuillan, Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St. Bloomington, IN 47405. Electronic mail may be sent to memcquil@indiana.edu.

 0009-3920/2019/xxxx-xxxx
DOI: 10.1111/cdev.13274

engagement with objects and the resulting views they gain of the objects.

Sustained Attention During Free Play

Free play provides a context of richly varying and dynamically complex interactions between children, toys, and, often, their mature social partners (e.g., parents). The play context is an important setting for real-world, everyday learning about objects (Needham, 2000), their names, and other words (Hart & Risley, 1995; Hirsh-Pasek, Golinkoff, Berk, & Singer, 2009; Ruff & Rothbart, 2001). The free play context encourages exploratory behavior and higher order cognitive processing through examination, problem solving, and decision making with toys in view (Barkley, 1991). To study attentional processes during play, duration and intensity of attention on focal objects are commonly used as measures of sustained, focused attention, manifested by prolonged gaze, slowed heart rate, furrowed brow, serious facial expression, and children manually bringing the object of interest near (Richards, 1989; Ruff & Capozzoli, 2003).

Sustained, goal-directed attention shares surface similarities with, but is distinct from, long visual fixations on objects, often referred to as "obligatory attention" (Stechler & Latz, 1966) or "sticky fixation" in younger infants (Hood, 1995; Colombo, 2001). These obligatory or sticky visual fixations reflect an inability to disengage from a stimulus (Colombo, 2001). By 4 months, the ability to inhibit continued fixation develops, leading to shorter fixations as infants rapidly encode familiar stimuli and switch attention to novel stimuli (Colombo & Cheatham, 2006). This development adaptively enables more learning opportunities and increased exposure to new information. Cognitively advanced infants who show shorter fixations are thought to demonstrate faster attentional scanning and information processing (Colombo & Cheatham, 2006). However, normatively, fixation duration increases again around 6–8 months when children exert top-down control processes to intentionally examine objects, reflecting depth of processing rather than efficiency of processing (Colombo & Cheatham, 2006). Duration of sustained focused attention during free play increases from 7 to 42 months (Richards, 1989; Ruff & Capozzoli, 2003; Ruff, Capozzoli, & Weissberg, 1998; Ruff & Lawson, 1990). Higher levels of sustained focused attention are also positively associated with more advanced problem solving (Hunt & Lansman, 1986), decision making (Tattersall, 1998), and standardized test scores of intelligence and achievement (Aylward, Gordon, & Verhulst, 1997).

Manual Engagement and the First-Person View

When children sustain attention during toy play, they do not simply look passively at a jumble of toys on the floor or at a table, but rather, they use their whole body, posturing and positioning their hands, head, and eyes to visually isolate focal objects. Hand activity is crucial for language (Bates & Dick, 2002), social communication (Bakeman & Adamson, 1986; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998), tool use (Lockman, 2000), problem solving (Goldin-Meadow, 2000), and visual learning about objects (James, Jones, Smith, & Swain, 2014; James, Jones, Swain, Pereira, & Smith, 2014; Soska, Adolph, & Johnson, 2010). Manual engagement with objects during free play is associated with properties of the first-person view because the child's hands create and constrain the visual scene. This has been best demonstrated by research using head-mounted cameras (e.g., Pereira et al., 2014).

Head-camera research has shown that during mother–child free play with objects at a table, mothers tend to place four or more toys in front of the child, but only one or two of those toys are actually in the child's head-camera view at one time as the child moves, brings their body closer to the focal object(s), and visually isolates the object(s) using their own hand actions (Yoshida & Smith, 2008). Although additional objects could be seen in the periphery of the camera image, these objects tend to be occluded by objects that are dominant in the child's view as captured by head-camera images (Yu et al., 2009). This process of visually selecting an object and filtering out other objects through manual engagement creates more constrained and clean visual input, or one-object-at-a-time views, which support learning about the object (Smith et al., 2011; Yu et al., 2009). When parents named objects during moments with optimal views (i.e., when the child, through their own hand actions, had isolated an object so that it was large and dominant in view), the object names were more likely to be learned compared to naming events when the objects were not visually isolated (Yu & Smith, 2012).

Parental Responsiveness

A social partner who astutely notices the child's physical cues (i.e., child enclosing the object of interest by leaning their head down, moving their body close, and bringing the object near with their hands) may then name the object at this moment, when there is little referential ambiguity and greater likelihood of language learning (Yu & Smith, 2012). Parental responsiveness involves contingent and appropriate responses to the child's emotional, social, and cognitive needs (Ainsworth, 1973; Bornstein et al., 2008). The responsiveness of parents—in following and building on their children's interest—may encourage extended manual play that may also enrich the child's first-person view.

The larger literature on parental responsiveness suggests that parents play a direct a role in guiding and/or extending child attention. Here, we ask whether these effects of parent behavior emerge by influencing both child manual engagement and the child's first-person view of objects. An affirmative answer would suggest that the visual isolation of objects in children's first-person scenes is created and supported by the social context and is not strictly an intrinsic property due to children's short arms, for example (Spencer et al., 2009). Previous head-camera research showing that child manual engagement creates visually rich scenes that are optimal for learning has been primarily conducted in free-flowing play contexts with parents. These studies show correlations between parent behavior, child sustained attention, and child object name learning (Yu & Smith, 2016, 2017). However, these correlations do not show causal paths from what parents do, and what children do with their hands, to the creation of children's first-person scenes. The present study uses experimental manipulations of parent behavior to test a causal role for parent responsiveness.

The Present Study

The central question of this study pertains to the pathway from parent behavior—through the child's manual engagement with objects—to the properties of the child's first-person view of objects (which we refer to as "scenes"), especially to the properties previously shown to be key to learning. Two aspects of children's manual behaviors may be influenced by parent behavior and could consequentially influence the properties of the child's first-person scenes. First, parents could influence the duration of child holds and the rate at which objects are sampled. Parents may encourage children to have more prolonged bouts of examining a single object by maintaining child interest and avoiding restrictiveness and directiveness as shown in previous research (Landry, Smith, Swank, Assel, & Vellet, 2001). Alternatively, parents may interrupt sustained attention bouts with an object by introducing new objects and encouraging their children to switch between objects at a greater frequency. Second, parents may influence how closely children hold objects to their head and eyes, which aids visual isolation of the object in children's first-person views (Yu & Smith, 2012). Parents could either encourage close holds or distract children from holding objects close to themselves by introducing additional objects, leading to fewer child-perspective scenes in which a single object is visually dominant. Longer durations of holding and longer durations of uncluttered scenes observed during interactions with responsive parents would suggest a parent influence on child behavior similar to that reported with respect to sustained attention (Parrinello & Ruff, 1988; Suarez-Rivera, Smith, & Yu, 2019; Yu & Smith, 2016; Wass, Clackson, et al., 2018) but would show the effect on manual engagement and first-person scene structure rather than on gaze.

The present study used four experimental manipulations to encourage different parent–child interaction patterns during a free play task. The children were in the second half of their second year of postnatal life (16–25 months), a developmental period characterized by major development in attention (Ruff & Rothbart, 2001), language (Bloom, 2000), visual object recognition (Smith, 2003), and skilled manual behavior (Lockman, 2000). Head-mounted cameras were used to record, from the child's perspective, the visual scene properties created by handling objects. Each parent–child dyad was asked to follow two sets of instructions: one designed to encourage parents to play typically with their child and the other that was a perturbation of typical play. Previous research indicates that child sustained attention is greater during joint play (Lawson, Parrinello, & Ruff, 1992), and when the play is child-led and the social partner is responsive (Parrinello & Ruff, 1988). We used two typical play conditions that encouraged parents to interact in particular ways with their children: joint play in which one subset of parents was instructed to play with their child freely and child-led joint play in which the other subset of parents was instructed to be responsive by following their child's interest. We also used two perturbation conditions: one attempted to completely limit parent involvement (child plays alone without parent involvement for the first subset) and the other condition attempted to encourage more disruptive parent involvement (parent-led joint play when parent directs the play by redirecting the child and leading the interaction for the second subset).

We would expect to find differences in child manual engagement and visual properties of the child's first-person view at the subcondition level, specifically that children might show longer bouts of manual engagement and experience richer visual properties during joint play than during play alone and during child-led joint play than during parent-led joint play. Differences between the typical subconditions (joint play and child-led joint play) and the perturbed subconditions (play alone and parent-led joint play) might also be observed. For example, children may show longer manual engagement and experience richer visual properties in child-led joint play than in joint play because parents are specifically instructed to follow the child's lead in the child-led condition. As another example, children's bouts of attention and manual engagement may be more disrupted and shorter in the parent-led condition than in play alone because parents who purposefully interrupt and redirect their children likely disturb play dynamics more so than would be observed for a child playing independently. Despite these possibilities, the focus of primary analyses was on comparisons between the overarching typical and perturbed play conditions in an effort to maximize statistical power.

For comparisons between typical and perturbed play, we hypothesized four, mutually exclusive outcomes. First, parents could influence both the duration of child manual engagement with single objects as well as the proximity with which children hold objects near to themselves. A parent may help the child hold an object for a long duration to create a stable view while also encouraging the child to visually isolate the held object by holding it close to their face so the object is large, centered, and dominant in view. If this hypothesis were confirmed, we would find longer holds and the held objects to be larger, more centered, and more dominant in the typical conditions compared to the perturbed conditions. A second possibility is that parents may only influence the rate of manual actions, but the child's own body and action dynamics may constrain the visual properties—the size, centeredness, and dominance of the held object. If this hypothesis were confirmed, the visual learning environment would be unchanged across conditions, but the rate and duration of manual engagement would vary across conditions such that children would hold objects for longer when playing with a responsive parent. Our third hypothesis is a reverse of the second—that parents may instead only affect how closely the child holds objects without affecting the duration and frequency of holds. Although there may be little measurable change across conditions in the frequency or duration of child hand activity, parent responsiveness expected in the typical conditions may nonetheless help structure the visual properties of the child-perspective scenes by encouraging the child to hold objects close to themselves. A fourth possible outcome was that parents may have no effect on either the duration or the closeness of child manual engagement, regardless of level of involvement. If manual engagement is driven largely by the child's interests and the child's body, its rhythm, physical tendencies (e.g., for object proximity), and endogenous sensorimotor dynamics, the child would hold objects for similar durations and at similar distances from their face, with no observed differences across conditions. This possibility is not supported in several studies showing social effects on infant (10- and 12-month-olds') sustained engagement with toys (Parrinello & Ruff, 1988; Yu & Smith, 2016; Wass, Clackson, et al., 2018). However, the possibility that first-person object views during play are determined primarily by the sensorimotor dynamics of the child is an important possibility for older children. It would suggest a child role in observed dyad differences (Cohn & Tronick, 1988; Feldman, Greenbaum, & Yirmiya, 1999) and would highlight the critical importance of parents following the child's play patterns when naming and discussing objects (Mundy & Newell, 2007).

Method

Participants

Participants were recruited from a midsized, Midwestern city and surrounding rural communities. Recruitment was primarily through a database using county birth records and community outreach efforts. Advertisements, such as postcards and flyers, were also used. The final sample consisted of two groups of parent–child dyads (total $N = 29$) who engaged in free play with four sets of five unique, novel objects. Six additional children were recruited, but either did not tolerate the head camera or were excluded because of fussiness before the experiment started. In the final sample, children were 16–25 months of age ($M = 21.14$, $SD = 2.64$). Fifteen children were girls (52%), and 25 children were Caucasian (86%). Three children were African American, and one child was Asian American, which is relatively consistent with the general racial breakdown of the Midwestern city (approximately 4% African American and 9% Asian American according to recent US Census Data). The sample predominantly included middle class families, and 31% of the sample lived in school districts in which at least 40% of the total student population received free or reduced lunch. The two groups of subjects did not differ significantly in child age, $t(24) = -0.61$, $p = .55$, or gender, $t(24) = 1.54$, $p = .14$.

Procedure

The local institutional review board approved the study, and parents provided informed consent. Each parent–child dyad alternated between typical and perturbed play conditions for four total play trials, with randomized trial order. Dyads could begin the trial sequence with either typical or perturbed play, but they never experienced two similar trial types in a row. The first 13 subjects alternated between joint play (typical) and play alone (perturbed) conditions. For the second 16 subjects, the perturbation, parent-led joint play, involved parent over-involvement, as parents were instructed to lead the play interaction. The 16 dyads alternated between this and a contrasting, more typical condition, child-led joint play, in which parents were simply told to *follow* their child's lead. This instruction was given to ensure that parents changed their behavior and did not continue to direct the play interaction. Participants were told that the experimenters were interested in different styles of play and understanding how parents and children organize each other's attention. They were told that there were four sets of colorful, unique toy objects that were created in the lab. After first seeing a sample set of toy objects, the parent was told that they and their child would be invited to play with each toy set for one and a half minutes and that they would be asked to play with their child in two different ways (for a total of 6 min of play). The parents in the joint play versus play alone conditions were asked to alternate between playing with their child as they normally would and letting their child play alone as they read a research article provided to them. During play alone, when the child attempted to communicate by waving an object to show their parent, for example, the parent was asked to minimally respond by looking toward the child and the object with verbal acknowledgment (e.g., "ok," "yeah") and immediately return their attention to the article.

The second set of 16 parents, in the parent-led versus child-led joint play conditions, were asked to alternate between leading and following the child's attention. For leading, they were told: (a) to decide which object to focus on and talk about and (b) to direct the child's attention to a different object of the parent's choosing if the child is already holding another object. Parents were encouraged to play with all of the toys in any order. In contrast, for the two child-led joint play trials, the parents were told to

> Sit back, let your child lead the play and only play with the objects your child chooses. Child's choice can be a point, pick-up, or look. You should follow your child's lead, giving your attention to, commenting on, and playing with the toys that your child is interested in.

A nearby experimenter gave the parent a new set of toys and instructed the parent which play condition to implement in between each of the four trials.

As described in previous research (e.g., Yu & Smith, 2012), each novel object for toy play was a simple shape with a uniform color made from plastic, hardened clay, aggregated stones, or cloth. All objects were similar in size (on average 288 cm^3). Each set consisted of five objects with five different colors (blue, green, red, pink, and yellow). Parents and children sat across from each other at a small table (61 × 91 × 64 cm) that was painted white. The child's seat was 32.4 cm above the floor (the average distance from child's eye to the center of the table was 43.2 cm). Parents sat on the floor so that their eyes and heads were at approximately the same distance from the tabletop as those of the children (the average distance of eye to the table center for parents sitting on the floor was 44.5 cm). To aid in automatic image analysis, both participants wore white clothing. There were also white curtains from floor to ceiling and a white floor so that everything in the head-camera images was white with the exception of heads, faces, hands, and toys.

The toddler's head camera was embedded in a sports headband. The camera used was the Supercircuits (PC207XP; Supercircuits, Inc., Austin, TX) miniature color video camera and weighed approximately 20 g. The focal length of the lens was f3.6 mm. The number of effective pixels were 512 (horizontal) × 492 (vertical; NTSC). The resolution (horizontal) was 350 lines. The camera's visual field was 70° and provided a broad view of objects in the head-centered visual field that was less than the full visual field (approximately 170°). The recording rate was 10 frames per second. The direction of the camera lens when embedded in the sports band was adjustable. Input power and video output went through a camera cable connected to a wall socket, via a pulley, so as to not hinder movement. The head camera was connected via standard Radio Corporation of America cables to a digital video capture card in a computer in an adjacent room. The headband was tight enough that the camera did not move once set on the child. The multichannel video capture card in the recording computer adjacent to the experiment room simultaneously recorded the video signal from the camera. The head camera moved with head movements, but not with eye movements, and therefore provided a head-centered view of events that may be momentarily misaligned with the direction of eye gaze. However, head-mounted eye-tracking studies have shown that under active viewing conditions, human observers typically turn both heads and eyes in the same direction, align heads and eyes within 500 ms of a directional shift, and maintain head and eye alignment when sustaining attention (Bambach, Crandall, & Yu, 2013). The result is that the distribution of gaze in active viewing is centered in the head-camera image. Thus, in a large corpus of images recorded during active viewing, the likelihood of gaze falling within the head-camera image is over 97% (Bambach, Smith, Crandall, & Yu, 2016). A high-resolution camera (recording rate 30 frames per second) was mounted above the table, providing a bird's eye view. This camera recorded visual information independent of participants' movements and was used to resolve any ambiguities in the head-camera images. The parent's voice during the interaction was recorded with a standard headset with a noise reduction microphone.

Procedures involved a team of three experimenters. Upon entering the experiment room, the child was seated in the chair and a push-button, pop-up toy was placed on the table. One experimenter played with the child while the second experimenter placed the head-band low on the child's forehead when the child was engaged with the toy. The first experimenter then directed the child to push a button on the pop-up toy while the second experimenter adjusted the camera, such that the toy's button was near to the center of the head-camera image (as viewed by a third experimenter in the control room). During the head-camera calibration, the third experimenter in the control room confirmed that the object was at the center of the image, and if not, small adjustments were made to the camera.

Multimodal Measures in Free Play

The main dependent measures were holding activity and visual properties of the child head-camera images, namely the size, centeredness, and relative dominance of the held object compared to other objects in view in each frame. Parents' speech was also transcribed to measure verbal behavior in different interaction conditions.

Object Holding

Parent and child holding behaviors were coded manually, frame-by-frame, by trained coders who watched the entire session from multiple angles (head-cameras and third-person view cameras). Coders made frame-by-frame annotations when parents and their children touched each of the objects. Two coders independently coded the same randomly selected 25% of frames (checking head-camera images to resolve any ambiguities) with 100% agreement. Any in-hand moment that lasted 3 or more seconds was scored as a sustained hold. The threshold of 3 s corresponds to the average duration of concentrated attention for 1-year-olds reported in earlier work (Ruff & Lawson, 1990) and is consistent with approaches used in previous research to measure sustained attention (Suarez-Rivera et al., 2019; Yu & Smith, 2016). A threshold of 3 s ensures that the defined sustained holds are on the tail of the distribution and thus correspond to the upper limits of sustained attention abilities for children at this age (Ruff & Lawson, 1990; Ruff, Lawson, Parrinello, & Weissberg, 1990). This prolonged threshold is particularly relevant for research on the role of parent behavior in helping children sustain attention. The sustained in-hand moments were further screened to exclude moments in which a child may have left one hand on an object absent mindedly without actually interacting with the object. The frequency and duration of child sustained holds were used in all analyses. We also examined the rate of switching between objects—that is, we dummy coded each hold (for both in-hand moments and sustained holds) to indicate whether the same object was held (non-switch) or a different object was held from what was held previously (switch). We then looked at the number of switches relative to the number of holds.

Visual Properties of the Child's First-Person View

Properties of the visual scene from the child's head camera were automatically coded, frame-by-frame, via a machine vision program (Yu et al., 2009). The absolute size of the object in the child's hand was measured using the percentage of the field of view from the child's head-camera image that was occupied by the held object. Consistent with previous research (Pereira et al., 2014), centeredness of the object in the child's hand was measured by computing the average distance of all object pixels to the image center and expressing that average distance as a proportion of the head-camera image's half diagonal—that is, a fully centered object pixel corresponds to zero centering, and a head-camera image corner pixel has a centering value of one. Object dominance was measured using the relative size of an object—that is, the ratio of the largest object in view to the other objects in view. A ratio of 0.50 or higher characterized an object as "dominating" or not; a 0.50 ratio means that the larger object is at least larger than the combination of the other objects (Yu et al., 2009). The proportion of time that the object in the child's hand was dominant in view during a sustained hold was used in all analyses.

Parent Speech

Parents' speech was transcribed by research assistants and divided into utterances, which were defined as strings of speech between two periods of silence lasting at least 400 ms. We assessed reliability by having a second research assistant independently denote utterances for a random selection of 25% of the participants (Cohen's $\kappa = .84$). Parent utterances were then categorized into seven mutually exclusive codes: 1. Labeling an object (e.g., "That is a hammer." "That is a microphone."); 2. Referencing an object (e.g., "I like that one." "Mommy has all of these."); 3. Describing an object or asking descriptive questions about the object (e.g., "That one is like our vent at home." "That one feels spikey" "Is that blue? Is it heavy?"); 4. Mentioning an object function or action (e.g., "this can spin" or "we can bang this one"); 5. Giving directives (e.g., "Hey [child name], look!"); 6. Responding to something the child says about or does with the object(s) (e.g., "yes, good job!"); 7. Other (e.g., responding to the experimenter when it is time to switch trials or discussing the head-camera equipment). A second coder independently coded utterances for 20% of the cases to establish interrater reliability ($\kappa = .81$). The frequency of directive (Category 5) and responsive (Category 6) parent utterances, and the proportion of all utterances that were classified as either directive or responsive, were used in the present study.

Data Analysis Plan

Although we would expect to see differences in child manual engagement and visual properties of the child's first-person view at the subcondition level, these comparisons were statistically underpowered and nonsignificant. The focus of primary analyses, therefore, was on comparisons between the overarching typical and perturbed play conditions (see Figure 1). Our first set of analyses involved examining parent verbal and manual behavior to confirm that parents complied, on average, with task instructions and reliably changed their behavior when alternating between conditions. We also tested for similarities in parent behavior in the typical conditions, as parents were expected to be more responsive than directive based on previous research for naturalistic joint play and based on task instructions for the child-led condition.

Next, we examined children's manual engagement (i.e., duration, frequency, and switch rate of sustained holds) to determine whether change in parent behavior across the conditions altered children's handling of objects. We used overall nested analyses of variance (ANOVAs) to document any differences between the typical and perturbed play conditions within subjects. We also tested for differences across conditions in the frequency and duration of dominant events (i.e., when the held object was dominant in view relative to the other objects). Additionally, we examined histograms of the visual qualities of objects during sustained holds to examine the distributional properties and patterns that emerged across conditions. We also used nested ANOVAs and linear mixed effect models to assess whether the visual qualities of the held object (size, centeredness, and dominance) during sustained holds changed across conditions.

Results

Parent Behavior

In three of the subconditions (joint play, child-led joint play, and parent-led joint play), parents

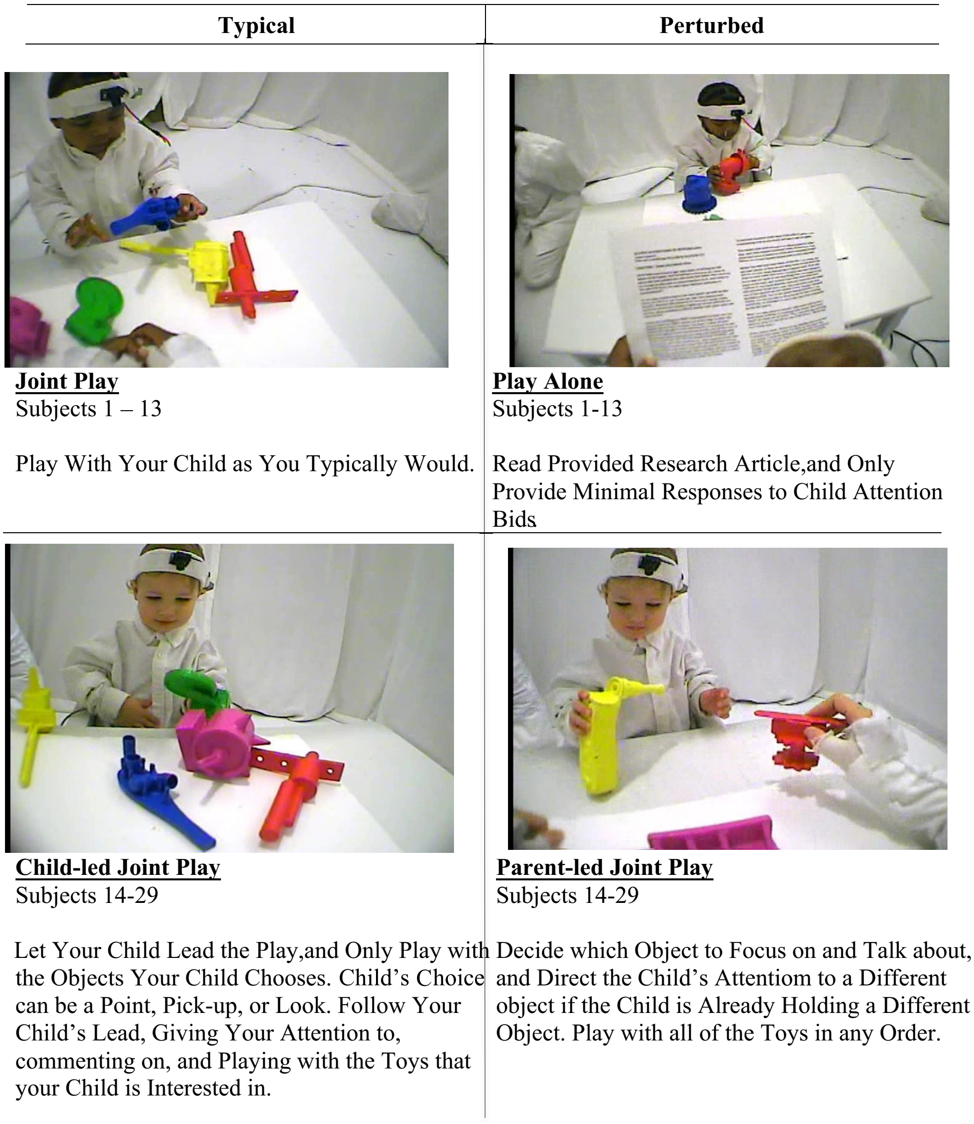

Figure 1. Summary of experimental conditions with images from parent head camera. All subjects alternated between two trials of typical play and two trials of perturbed play.

were instructed to engage their child in particular ways. In one subcondition (play alone), parents were instructed to avoid interaction with their child. Table 1 shows the statistics for parent utterances in the three conditions with parent involvement. We examined differences between the parent-led (perturbed) and child-led (typical) joint play conditions to confirm whether or not the instructions effectively manipulated parent behavior across these two conditions. We found that parents spoke more (M = 69.50, SD = 19.25 utterances), were more directive (M = 12.63, SD = 7.58 directive utterances), and held objects more (M = 35.43, SD = 21.41 parent in-hand moments) in the parent-led joint play condition than in the child-led joint play condition (M = 55.69, SD = 16.63 utterances, M = 7.06, SD = 7.03 directive utterances, M = 14.43, SD = 12.67 parent in-hand moments). Parents responded appropriately and as expected to the task instructions. In the parent-led joint play condition, the parent was also more likely to hold the focal object (i.e., the object subsequently held by the child) in the 3 s window prior to a child's sustained hold relative to the likelihood of the child touching or holding the object before a sustained hold. On average, parents held the focal object for 51% of the time in the 3 s before a child's sustained hold in the parent-led joint play condition (SD = 0.42), whereas

Table 1
Descriptives of Parent Verbal Behavior Across Conditions

		All parent utterances *M* (*SD*)	Directive utterances *M* (*SD*)	Responsive utterances *M* (*SD*)
Typical	Joint play *N* = 13	73.69 (19.51)	6.08 (5.39)	16.15 (10.65)
	Child-led joint play *N* = 16	55.69 (16.63)	7.06 (7.03)	12.38 (5.80)
Perturbed	Play alone *N* = 13	—	—	—
	Parent-led joint play *N* = 16	69.50 (19.25)	12.63 (7.58)	11.00 (5.05)

Note. All parent–child dyads alternated between two typical play trials (either two trials of joint play or two trials of child-led joint play) and two perturbed play trials (either two trials of play alone or two trials of parent-led joint play) for a total of four 1.5-min trials, or 6-min total of free play. The number of utterances (all, directive, and responsive) spoken by each parent across the two typical and two perturbed play trials was computed by condition, and then the average of these values across subjects was computed by condition. Those means are displayed here.

in the child-led joint play condition, parents only held the focal object for 15% of the time in the 3 s prior to a child's sustained hold (*SD* = 0.17). Similarly, parents held the focal object with a greater frequency in the 3 s prior to a child's sustained hold in the parent-led joint play condition compared to the child-led joint play condition (*M* = 0.71, *SD* = 0.56 parent in-hand moments with the focal object before a child's sustained hold in parent-led joint play versus *M* = 0.36 *SD* = 0.30 parent in-hand moments before a child's sustained hold in child-led joint play). This suggests that parents were indeed leading the interaction in the Parent-led Joint Play condition relative to the child-led joint play condition, perhaps by manipulating the object of interest before handing it to the child.

Next, we tested for similarities in parent behavior in the two conditions classified as "typical." As depicted in Figure 2, parents were significantly more responsive than directive in both joint play (*M* = 0.21, *SD* = 0.10 proportion of utterances that were responsive compared to *M* = 0.08, *SD* = 0.06 proportion of utterances that were directive), $F(1, 24) = 15.78$, $p < .01$, and in child-led joint play (*M* = 0.23, *SD* = 0.09 proportion of utterances that were responsive compared to *M* = 0.13 *SD* = 0.13 proportion of utterances that were directive), $F(1, 30) = 6.53$, $p < .05$. These findings justify the classification of both of these play conditions as comparable to each other and reflective of typical, responsive play.

Child Holding Behavior

Figure 3 shows histograms of all child in-hand moments, including object touches and holds, for each of the four subconditions. The histograms reflect the non-normality and positive skew of child holding behavior, with many short instances of manual engagement and far fewer sustained holds, regardless of condition. The mean and median durations of child in-hand moments were longer in both joint play (*M* = 4.59, *SD* = 7.14 s, median = 2.3 s) and child-led joint play (*M* = 5.58, *SD* = 7.60 s, median = 3.39 s), the typical conditions, compared to their counterpart perturbed conditions with parental underinvolvement in play alone (*M* = 4.38, *SD* = 6.87 s, median = 2.2 s) and parental over-involvement in parent-led joint play (*M* = 5.23, *SD* = 7.87 s, median = 2.52 s). However, examining measures of central tendency for the duration of child holding behavior may be misleading given the skewed nature of the data. Instead, we collapsed these four subconditions into the overarching typical and perturbed conditions and examined the more normally distributed frequency of sustained holds (see Figure 3).

We then used overall nested ANOVAs to examine within subject differences in child holding (see Figures 4 and 5). We found that there were significantly more sustained holds in the typical conditions (*M* = 15.48, *SD* = 6.00 holds) compared to the perturbed conditions (*M* = 13.52, *SD* = 5.19 holds), which were characterized by either no parent involvement in play alone or parent overinvolvement (with more parental utterances, directiveness, and holds overall) in parent-led joint play, $F(1, 2) = 63.77$, $p < .05$. There was not significant variation between subjects within these overarching groups, $F(2, 54) = 0.03$, $p = .97$, meaning that the frequency of sustained holds in the two nested subconditions within the main conditions did not differ from each other. Because both of the typical subconditions were characterized by more responsive parent utterances, fewer directive utterances, and fewer parent utterances and object holds overall, this result suggests that parental responsiveness may be crucial for sustained manual engagement.

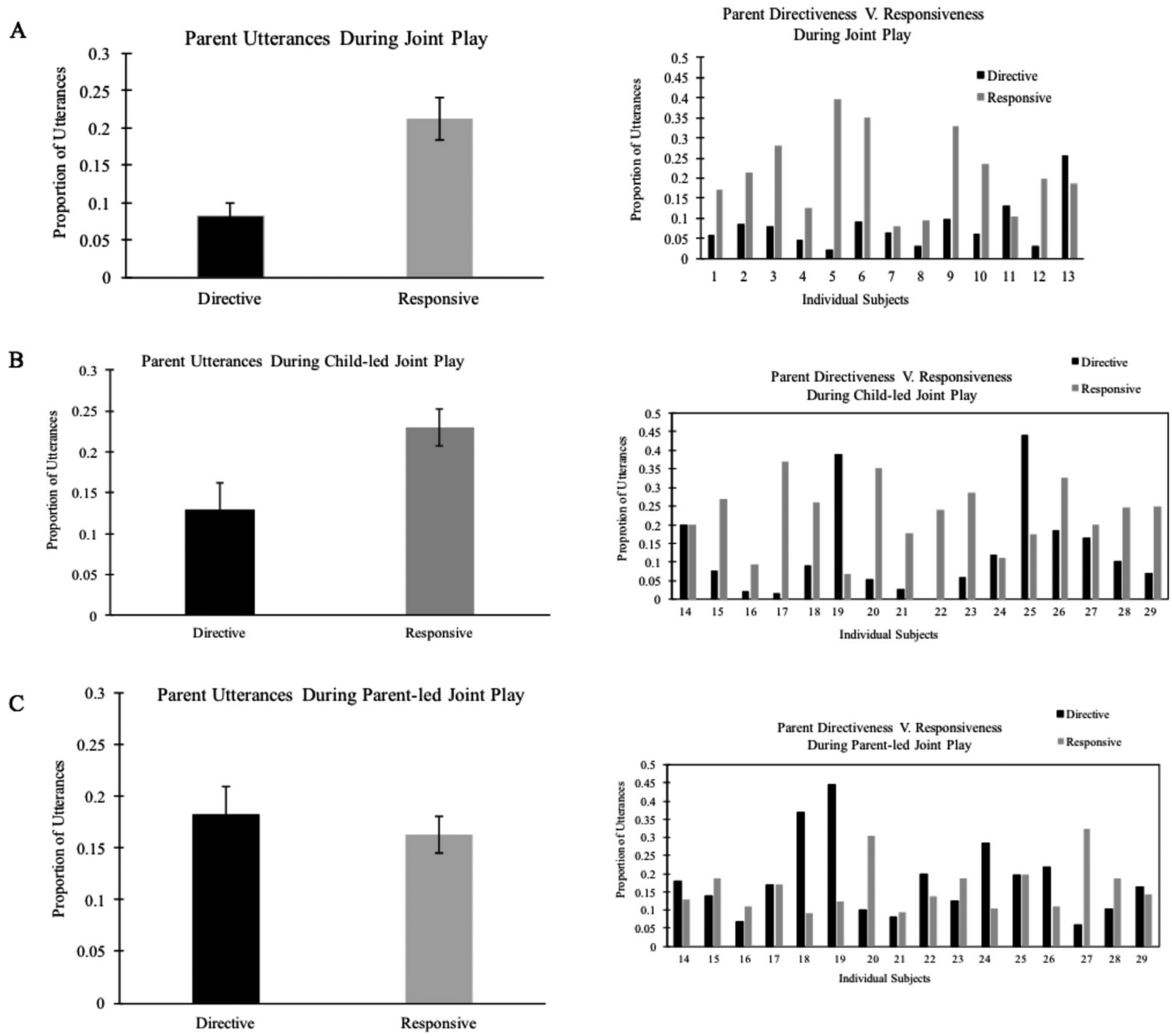

Figure 2. Parent utterances in conditions with parent involvement. (A) corresponds to joint play, (B) corresponds to child-led joint play, and (C) corresponds to parent-led joint play. Graphs on the left depict means across subjects, and graphs on the right show data for each subject. Error bars represent the standard error. Subjects 19, 20, 25, and 27 in B and C show poor compliance to task instructions, with Subjects 19 and 25 being more directive in child-led play and Subjects 20 and 27 being more responsive in parent-led play. Subjects 19 and 20 showed notable stability in their types of responses across both conditions.

Object manipulation was less likely to be sustained when the child played alone and when the parent tried to lead the child to engage with an object.

Consistent with this result, there was also less switching between objects of focus from one sustained hold to the next in the typical conditions compared to the perturbed conditions, $F(1, 2) = 28.69$, $p = .03$ (see Figure 5). Again, there was not significant variation between subjects within these groups, $F(2, 54) = 0.36$, $p = .70$, meaning that the rates of switching between objects in the two nested subconditions within the main conditions did not differ by subcondition. To summarize, a child was likely to show longer bouts with a given object in typical play conditions, characterized by more responsive parent utterances, less directive utterances, and fewer parent utterances and object holds overall, than in a perturbed condition, which either had no parent involvement (in play alone) or parent overinvolvement (in parent-led joint play). This suggests that parental responsiveness, rather than directiveness or lack of involvement, may

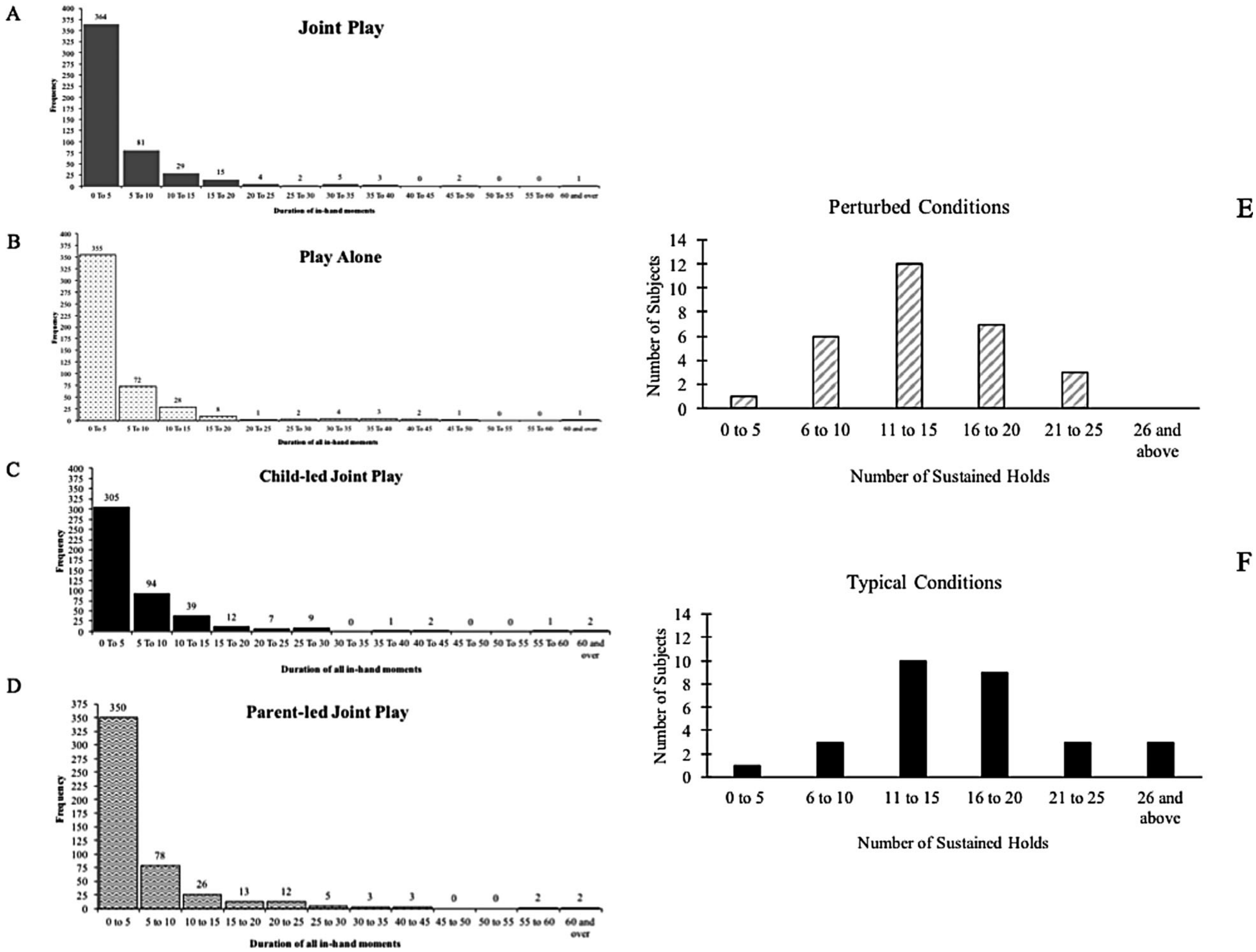

Figure 3. Histograms of holding behavior. A–D correspond to joint play, play alone, child-led joint play, and parent-led joint play, respectively, and depict the duration distribution of all in-hand moments. All parent–child dyads alternated between two typical play trials (1.5 min each in either joint play [N = 13] or child-led joint play [N = 16]) and two perturbed play trials (1.5 min each in either play alone [N = 13] or parent-led joint play [N = 16]) for a total of 6 min of free play. All in-hand moments observed during either of the two trials by subcondition are shown here. Total in-hand instances across the two trials was 506 for A, 477 for B, 472 for C, and 494 for D. E and F collapse across perturbed and typical conditions, respectively, and depict the distribution across subjects of the frequency of sustained holds.

promote continued child interest on a focal object. Even when all in-hand moments, not just sustained holds (>3 s), were examined, there was still relatively more switching between objects from hold to hold in play alone (M = 0.45, SD = 0.11) compared to joint play (M = 0.38, SD = 0.09), at a level trending toward significance, $F(1, 24) = 3.01$, $p = .09$.

Visual Properties

Next, the properties of the child's first-person view, namely the dominance, size, and centeredness of held objects relative to other objects, were examined. Figure 6 shows the histograms for the durations of time in which the held object was dominant in view for each of the four subconditions. All four histograms show similar skews with many short instances of dominance, and there were no significant differences between the typical play and perturbed play conditions for either the duration or frequency of dominant events. Figure 7 shows histograms by subcondition for the proportion of time the held object was dominant in view during a sustained hold. Across conditions, held objects tended to be dominant in view for up to 10% of the time during the majority of sustained holds.

Figure 7 also shows histograms for the size and centeredness of held objects in each of the four subconditions during sustained holds. The histograms reflect a high degree of consistency in how objects appeared in the child's visual learning environment

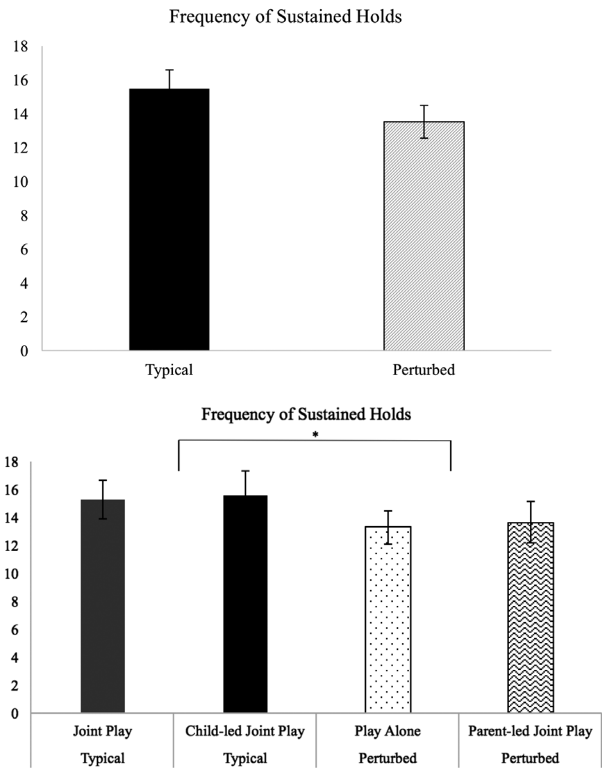

Figure 4. Overall nested analysis of variance comparing frequency of sustained holds in typical and perturbed play conditions. There was significant variation within subjects between typical and perturbed play, $F(1, 2) = 63.77$, $p = 0.02$, but there was not significant variation between subjects within these groups, $F(2, 54) = 0.03$, $p = 0.97$. Error bars represent the standard error. The lower graph displays the means for the subconditions, which are collapsed together into typical and perturbed play in the higher graph. Subjects 1–13 alternated between joint play (for typical) and play alone (for perturbed). Subjects 14–19 alternated between child-led (for typical) and parent-led (for perturbed).

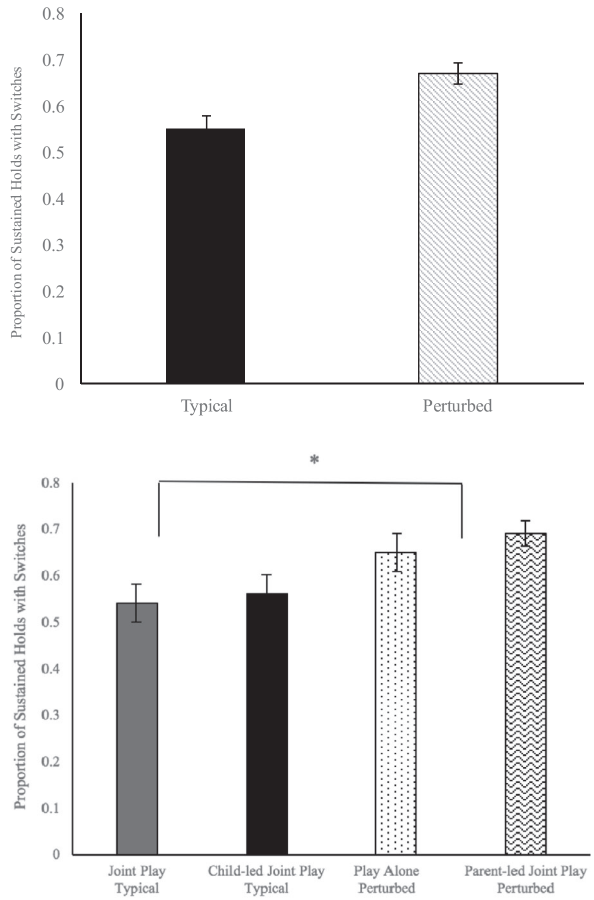

Figure 5. Overall nested analysis of variance comparing switch rate in typical and perturbed play conditions. There was significant variation within subjects between typical and perturbed play, $F(1, 2) = 28.69$, $p = 0.03$, but there was not significant variation between subjects within these groups, $F(2, 54) = 0.36$, $p = 0.70$. Error bars represent the standard error. The lower graph displays the means for the subconditions, which are collapsed together into typical and perturbed play in the higher graph. Subjects 1–13 alternated between joint play (for typical) and play alone (for perturbed). Subjects 14–19 alternated between child led (for typical) and parent led (for perturbed).

across the four play conditions. For example, regardless of condition, held objects generally occupied up to 14% of the visual field during sustained holds. The descriptives of these visual scene properties are also summarized for the overarching typical and perturbed play conditions in Table 2. There were no significant differences across conditions in the size, $F(1, 2) = 0.07$, $p = .82$, centeredness, $F(1, 2) = 1.23$, $p = .38$, or dominance of the held object in the child's view during sustained holds, $F(1, 2) = 0.08$, $p = .80$.

Given the skewed nature of the duration, dominance, size, and centeredness data, we also used linear mixed effect modeling to examine each sustained hold at the instance level across all subjects ($N = 879$ sustained holds). We tested whether condition (either typical or perturbed) predicted differences in the duration of sustained holds or differences in object size, centeredness, or dominance during sustained holds. None of these models were statistically significant ($\beta = -.15$, $p = .78$ for duration, $\beta = -.04$, $p = .86$ for size, $\beta = -.02$, $p = .68$ for centeredness, and $\beta = -.02$, $p = .39$ for dominance, such that holds tended to be shorter and objects tended to be smaller, less centered, and less dominant in view during perturbed play compared to typical Play but not to significant levels). Thus, even though parent behavior influenced the rate of child manual actions, the child's own body and actions appeared to determine the visual properties of what children see, independent of parent behavior.

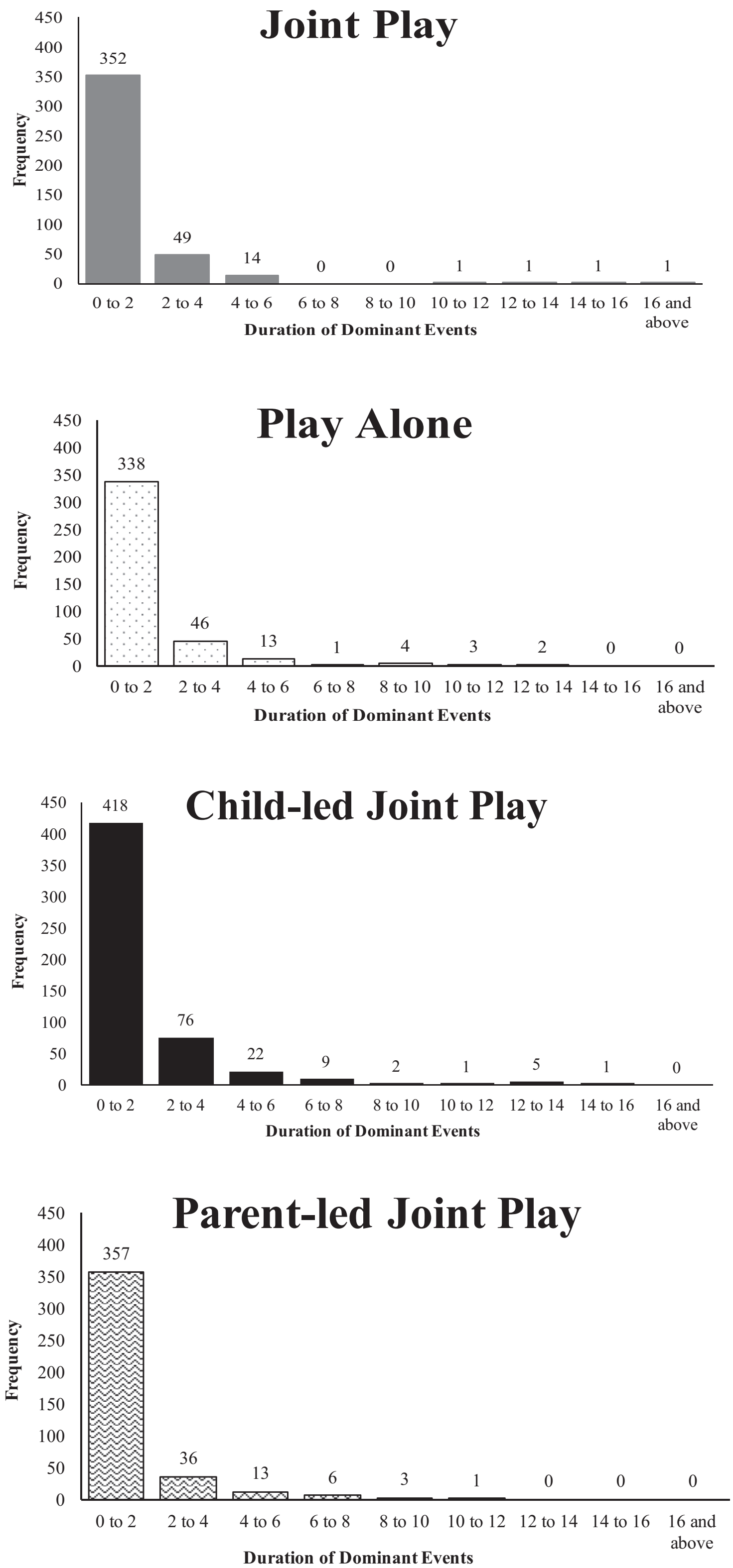

Figure 6. Histograms for the durations of time (in seconds) in which the held object is dominant in view.

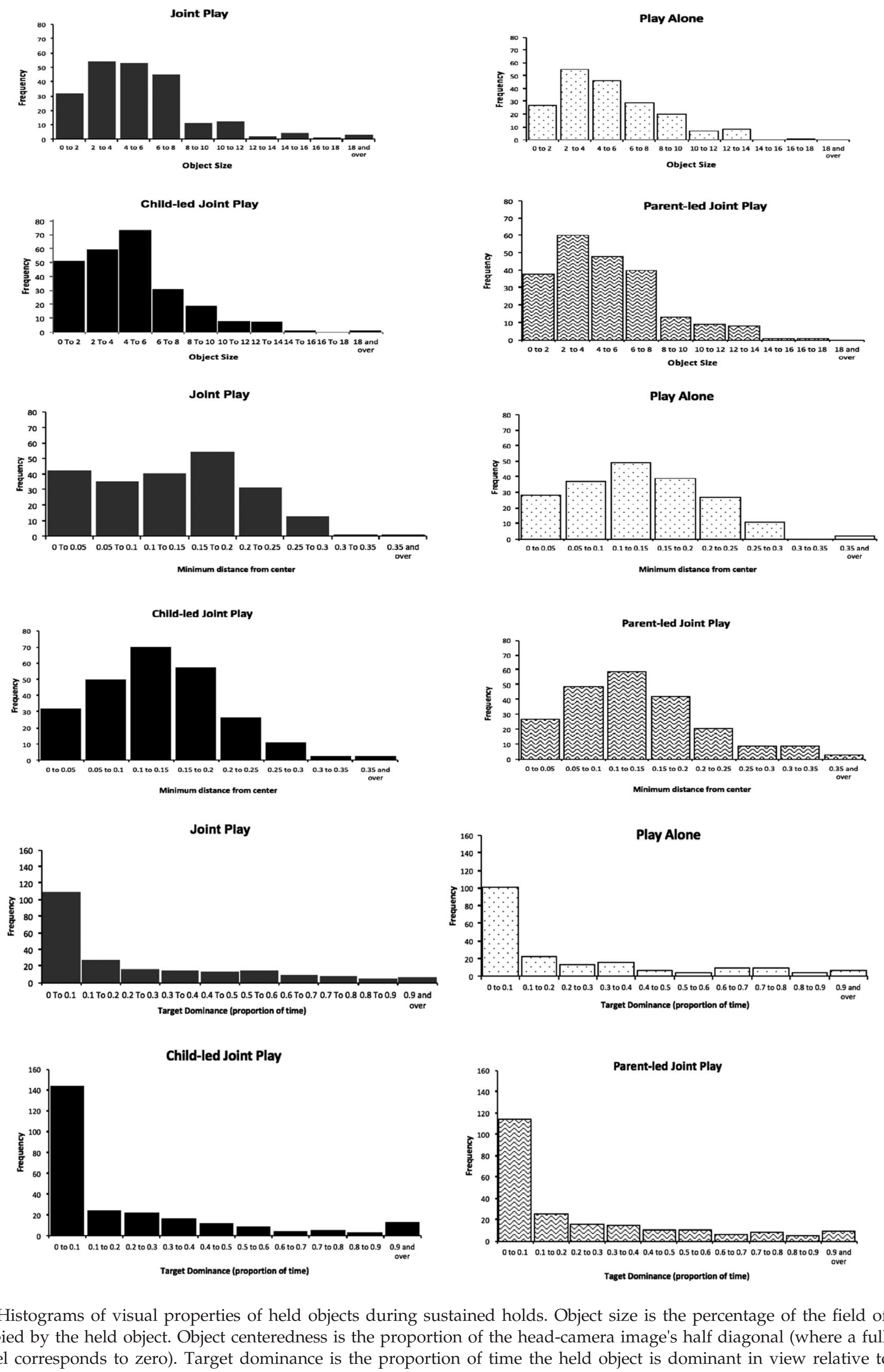

Figure 7. Histograms of visual properties of held objects during sustained holds. Object size is the percentage of the field of view that was occupied by the held object. Object centeredness is the proportion of the head-camera image's half diagonal (where a fully centered object pixel corresponds to zero). Target dominance is the proportion of time the held object is dominant in view relative to the other objects.

Table 2
Descriptives for Visual Properties of Sustained Holds Between Typical and Perturbed Conditions in the Child's Head Camera

	Typical play conditions			Perturbed play conditions		
	M	*SD*	Range	*M*	*SD*	Range
Object size	5.20	1.64	2.84–8.59	5.07	1.41	2.12–8.49
Object centeredness	0.13	0.04	0.06–0.19	0.14	0.04	0.07–0.20
Object dominance	0.21	0.13	0.05–0.55	0.22	0.11	0.06–0.55

Note. The typical play conditions include joint play (for Subjects 1–13) and child-led joint play (for Subjects 14–29). The perturbed play conditions include conditions with parent under-involvement (play alone) for Subjects 1–13 and parent overinvolvement (parent-led joint play) for Subjects 14–29. Sustained holds included any moment the child held an object for 3 or more seconds.

Discussion

Previous research has suggested that children may be more attentive with a responsive social partner than when playing alone or with a less responsive partner. Lawson et al. (1992) found higher rates of focused exploration of objects and less inattention during joint play than independent play. Similarly, infants' look durations toward play objects were longer, and moments of inattentiveness were fewer and shorter, during joint play compared to independent play (Wass, Clackson, et al., 2018). Yu and Smith (2016) also found children's visual attention to an object was extended when parents joined their gaze by looking at the same object. However, previous to the present study, it was unknown whether there would be observable differences in child hand activity and visual scene properties at differing levels of parent involvement.

After first confirming that experimental manipulations significantly altered parent behavior, this study yielded two main findings. First, changes in the level of parent involvement resulted in changes in child manual engagement with objects. In the typical play conditions, which were characterized by more responsive parent utterances, less directive utterances, and fewer parent utterances and object holds overall, children showed more frequent sustained holds and less switching between objects compared to the perturbed play conditions, which were characterized by either no parental involvement (in play alone) or parent overinvolvement (in parent-led joint play). This result is consistent with the finding from Wass, Clackson, et al. (2018) showing that infants looked longer at play objects during joint play than they did during solo play. The present study extends this previous research by also including a second form of perturbation—parent-led joint play—for one subset of participants. Across perturbed play trials when parents were either not involved or overly involved by being intentionally directive, children showed fewer sustained holds and flitted between objects at a faster rate. Hearing parents of deaf children tend to be more directive and controlling in their interactions with their children (Vaccari & Marschark, 1997), which may have downstream negative consequences for their children's manual engagement and attention development. Of course, this notion requires further research as the present study cannot definitively conclude that parent directiveness causes disruptions in sustained manual engagement.

Second, we found no changes across conditions in how held objects appeared in the child's head camera during sustained holds. That is, on occasions when there were sustained holds, independent of whether there were many or few for a child, the visual size, centeredness, and dominance of the held object was not influenced by the experimental condition. Thus, parent behavior affected the frequency of manual engagement but not the visual effects of manual engagement itself. Parents in the typical play conditions appeared to affect the quantity of sustained holds but not the quality of them. Regardless of parental involvement, sustained holds were visually rich, based on the size, centeredness, and dominance in view of the held object. With more statistical power, it is possible that parental responsiveness would be discovered to be associated with better visual qualities of sustained holds. An alternative possibility, consistent with the present findings, is that visual quality is determined by the child's sensorimotor dynamics. This interpretation parallels earlier work showing that the child's body and sensorimotor dynamics constrain the child's first-person view across both joint play and independent play contexts (Xu, Chen, & Smith, 2011). This is also reflective of the well-established, closed loop of perception and action in which initially perceiving an object precedes acting upon that object, which in turn constrains how the object is further perceived (e.g., Metta & Fitzpatrick, 2003).

We had proposed four possible outcomes from our experimental conditions: (a) parents would influence both the duration and closeness of child holds, (b) parents would influence only the duration of holds but not the visual properties experienced during these holds, (c) parents would influence only the visual scene properties without affecting the rate of child manual activity, or (d) parents would not influence either aspect of child hand activity. Our results confirmed the second hypothesis. Parents appear to influence the child's visual learning environment during toy play through their influence on how long children examine an object and how quickly they sample, or switch between, objects. A child was more likely to hold an object for a sustained period of time and switch less between objects when they were playing with a responsive parent compared to when they played alone or with a parent who was directive. When long holds with sustained attention on a focal object occurred, the held object was large, centered, and dominant in the child's first-person view, regardless of how involved or responsive the parent was. This is consistent with work from Corbetta et al. (2018) revealing the importance of holding leading to rich visual input.

The behavior of bringing objects close to create scenes in which a single object is visually isolated may reflect the fundamental operating characteristics of visual and manual engagement with objects for children in their second year of life. This aspect of child holding behavior for visual isolation of an object in the child-perspective scene has been shown to predict learning (Pereira et al., 2014; Yu & Smith, 2012). Based on the results of the present study, this process of visual isolation does not appear to be directly affected by parenting variations. Therefore, the role of parent behavior in supporting sustained attention (Yu & Smith, 2016) must lie elsewhere. The present study begins to answer this question of the role of parent involvement by showing that responsive parents help children experience more frequent, prolonged, high-quality, visually rich moments that previous research has shown are optimal for learning.

Limitations and Future Directions

The present findings suggest that responsive parenting contributed to more frequent sustained holds and less switching between objects. This is important because objects held during sustained holds consistently appeared large, centered, and dominant relative to other objects, creating good opportunities for learning object names (Pereira et al., 2014), and perhaps serving fine motor and spatial-mechanical development, too, through additional experience with the perception-action loop (Gibson, 1988). Children's perception of objects guides their actions with them, and these actions—grasping an object and bringing it close to their body for examination, exploration, and manipulation—result in improved perception of the object.

The visual scenes experienced during sustained holds are thought to constrain children's visual attention because the child can only attend to what is in view. Child eye gaze may be influenced by these visual scene constraints, child interest, and by parent behavior. Parents can operate on what is visually available to the child using their own gaze, utterances, and manual engagement to respond to the child's external signs of interest (Karasik et al., 2014; Tamis-LeMonda et al., 2013; Thurman & Corbetta, 2017). These parent behaviors may influence child eye gaze within visual scenes, likely promoting sustained attention and object name learning. To test this notion, differences in children's visual gaze patterns in lead-lag relationships with parent behavior could be examined using head-mounted eye trackers (e.g., Corbetta, Guan, & Williams, 2012; Franchak, Kretch, Soska, & Adolph, 2011; Yu & Smith, 2016). The child's learning experience likely arises from eye gaze during moments of manual engagement with objects through sensorimotor coupling and from parental input (i.e., what the parent highlights as important for learning through their own manual activity, vocalizations, and eye gaze). Because head-camera views and eye gaze are only aligned, not equivalent (Yoshida & Smith, 2008), future research with larger and more diverse samples and added methods (e.g., head-mounted eye trackers) could replicate and extend the present findings, such as by determining whether moments of parental responsiveness indeed extend children's sustained attention and whether moments of parental directiveness impede or disrupt sustained attention bouts. Observation of more diverse samples may capture more variance in parent–child interaction behaviors, leading to stronger inferences about the role of parental responsiveness in promoting manual engagement, sustained attention, and positive cognitive and social outcomes.

Three additional directions for future research involve clinical applications of these findings, examination of the precise timing of these effects over the course of development, and explicit tests of the effects of parental responsiveness and manual engagement on word learning. First, responsive parenting can be targeted in parenting interventions to improve parent–child interaction styles and child adjustment outcomes. Children born prematurely may particularly benefit from such interventions, because they tend to have problems with shifting attention and contingency learning, both of which can be targeted by specific responsive parenting behaviors, such as efforts to maintain child attention (Landry, Smith, Miller-Loncar, & Swank, 1997). The present findings suggest that parents can effectively change their behavior in response to explicit instructions to be more responsive, and this change has positive effects on child manual engagement. Future research could test formal interventions with these instructions to determine how generalizable and lasting these effects are for real-world parent–child interactions.

Second, previous research has shown that the cognitive and social benefits of parental responsiveness are greatest when responsiveness is experienced consistently across early and later childhood (Landry et al., 2001). This intriguing finding of potentially important individual differences in developmental experiences calls for further longitudinal research on how parental responsive play changes with older children, its role, if any, in older children's manual engagement, and if it is involved in socially important outcomes, such as school readiness. Third, the findings from the present study could be linked with word learning in future research, either using forced choice tests of learning novel object names (e.g., Yu & Smith, 2012) or using measures of real-world word learning, such as the MacArthur Communicative Development Inventory for American English (Fenson et al., 1994). It will be useful to learn more about how parental responsiveness promotes sustained holds and optimal moments for word learning.

The results from this study should also be interpreted in light of two important limitations. First, this study collected dense visual data from the child's perspective but used a small sample so statistical tests to detect differences between subconditions (either joint play/play alone or child-led/parent-led) were underpowered. The validity of the decision to collapse the data across participants and subconditions to form the overarching comparison conditions of typical and perturbed play and the replicability of the findings should be tested in future research with larger samples. However, a similar design and sample size used by Wass, Clackson, et al. (2018) with 38 infants (12-month-olds) and their parents who alternated between joint play and solo play yielded similar results. In their study, longer visual attention durations were observed in joint play relative to solo play, and there were consistent interindividual differences in child attention across the joint play and solo play conditions. In a subsequent study using the same sample and design, dual electroencephalography was recorded for the infants and their mothers (N = 20 and 22 in joint play and solo play, respectively; Wass, Noreika, et al., 2018). Parents showed neural responsivity to changes in their infants' attention, and when parents showed greater neural responsivity, infants sustained attention for longer durations. Taken together, these findings provide additional support for the three key findings from the present study: (a) longer durations of attention were observed in joint play relative to solo play, (b) consistent interindividual differences in child attention were observed across play conditions, and (c) parent responsiveness was associated with more sustained attention. The present findings are also consistent with earlier research showing that 9-month-olds (N = 22) showed more sustained attention when interacting with a sensitive, unfamiliar adult than when interacting with a redirective, unfamiliar adult (Miller, Ables, King, & West, 2009). Additionally, Belsky and colleagues (among others) have shown that when maternal sensitivity is experimentally manipulated through several intervention sessions, 12-month-olds (N = 16) tend to show more focused object exploration during play (1980).

Further research with larger and broader samples is necessary to understand the role of parent behavior in direct comparisons of three distinct conditions (e.g., no parental involvement, "optimal parental involvement," and parental over-involvement). The recent work from Wass and colleagues suggests that attention may be more endogenously controlled during solo play and more exogenously controlled during joint play (Wass, Clackson, et al. 2018; Wass, Noreika, et al., 2018). The exogenously controlled nature of child attention during joint play may be more detrimental when parents are directive. Gaertner, Spinrad, and Eisenberg (2008) previously found that maternal control, such as intrusiveness and directiveness during play, was associated with less focused attention concurrently and longitudinally, especially for children with lower attentional abilities. This result is consistent with the present study in which children in perturbed play conditions, including parent-led play, showed fewer sustained bouts of manual engagement, but Gaertner et al.'s work also suggests that not all children may be affected by parent behavior in similar ways. A second notable limitation of the present study is that participants were not randomly assigned to the subconditions of typical and perturbed play, and thus, nonobvious carry-over effects between play trials were possible. Although randomized trial order was used, the possibility that a child's exposure to directive parenting in a Parent-led trial, for example, could have interfered with the child's natural play tendencies in the subsequent child-led trial cannot be ruled out.

Conclusions

The study's limitations are offset by several strengths, namely that parental involvement was experimentally manipulated to examine its influence on children's manual engagement and visual scene structure, which were measured via head cameras providing rich moment-to-moment data about children's sensorimotor dynamics during play. Parental responsiveness appeared to promote longer bouts of sustained manual engagement compared to when parents were either not involved or intentionally overinvolved. Regardless of parental involvement, sustained holds were visually rich, as held objects appeared large, centered, and dominant in view. These findings are important because they suggest not only that the child's body creates visually rich scenes across play contexts but also that parental responsiveness can increase the frequency of these visually rich moments.

References

Ainsworth, M. D. (1973). The development of infant–mother attachment. In B. M. Caldwell & H. N. Ricciuti (Eds.), *Review of child development research* (Vol. 3, pp. 1–94). Chicago, IL: University of Chicago Press.

Aylward, G. P., Gordon, M., & Verhulst, S. J. (1997). Relationships between continuous performance task scores and other cognitive measures: Causality or commonality. *Assessment*, *4*, 325–336. https://doi.org/10.1177/107319119700400403

Bakeman, R., & Adamson, L. B. (1986). Infants' conventionalized acts: Gestures and words with mothers and peers. *Infant Behavior and Development*, *9*, 215–230. https://doi.org/10.1016/0163-6383(86)90030-5

Bambach, S., Crandall, D. J., & Yu, C. (2013). Understanding embodied visual attention in child–parent interaction. In *Development and Learning and Epigenetic Robotics* (ICDL), 2013 IEEE Third Joint International Conference on (pp. 1–6). New York, NY: IEEE.

Bambach, S., Smith, L. B., Crandall, D. J., & Yu, C. (2016). Objects in the center: How the infant's body constrains infant scenes. In *Development and Learning and Epigenetic Robotics* (ICDL-EpiRob), 2016 Joint IEEE International Conference on (pp. 132–137). New York, NY: IEEE.

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of Abnormal Child Psychology*, *19*, 149–178. https://doi.org/10.1007/BF00909976

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology*, *40*, 293–310. https://doi.org/10.1002/(ISSN)1098-2302

Belsky, J., Goode, M. K., & Most, R. K. (1980). Maternal stimulation and infant exploratory competence: Cross-sectional, correlational, and experimental analyses. *Child Development*, *51*, 1168–1178. https://doi.org/10.2307/1129558

Bloom, P. (2000). *How children learn the meanings of words* (Vol. 377). Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/3577.001.0001

Bornstein, M. H., Tamis-Lemonda, C. S., Hahn, C. S., & Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, *44*, 867–874. https://doi.org/10.1037/0012-1649.44.3.867

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (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*(Serial No. 255), i–174. https://doi.org/10.2307/1166214

Choudhury, N., & Gorman, K. S. (2000). The relationship between sustained attention and cognitive performance in 17–24-month old toddlers. *Infant and Child Development*, *9*, 127–146. https://doi.org/10.1002/(ISSN)1522-7219

Cohn, J. F., & Tronick, E. Z. (1988). Mother–infant face-to-face interaction: Influence is bidirectional and unrelated to periodic cycles in either partner's behavior. *Developmental Psychology*, *24*, 386–392. https://doi.org/10.1037/0012-1649.24.3.386

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367. https://doi.org/10.1146/52.1.337

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior*, *34*, 283–322. https://doi.org/10.1016/S0065-2407(06)80010-8

Corbetta, D., Guan, Y., & Williams, J. L. (2012). Infant eye-tracking in the context of goal-directed actions. *Infancy*, *17*, 102–125. https://doi.org/10.1111/j.1532-7078.2011.00093.x

Corbetta, D., Wiener, R. F., Thurman, S. L., & McMahon, E. (2018). The embodied origins of infant reaching:

Implications for the emergence of eye-hand coordination. *Kinesiology Review*, *7*, 10–17. https://doi.org/10.1123/kr.2017-0052

Feldman, R., Greenbaum, C. W., & Yirmiya, N. (1999). Mother–infant affect synchrony as an antecedent of the emergence of self-control. *Developmental Psychology*, *35*, 223–231. https://doi.org/10.1037/0012-1649.35.1.223

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(Serial No. 242), i–185. https://doi.org/10.2307/1166093

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-mounted eye tracking: A new method to describe infant looking. *Child Development*, *82*, 1738–1750. https://doi.org/10.1111/j.1467-8624.2011.01670.x

Gaertner, B. M., Spinrad, T. L., & Eisenberg, N. (2008). Focused attention in toddlers: Measurement, stability, and relations to negative emotion and parenting. *Infant and Child Development*, *17*, 339–363. https://doi.org/10.1002/ICD.580

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, *39*, 1–42. https://doi.org/10.1146/annurev.ps.39.020188.000245

Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child Development*, *71*, 231–239. https://doi.org/10.1111/1467-8624.00138

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

Hirsh-Pasek, K., Golinkoff, R., Berk, L., & Singer, D. (2009). *A mandate for playful learning in preschool: Presenting the evidence*. New York, NY: Oxford University Press.

Hood, B. M. (1995). Shifts of visual attention in the infant: A neuroscientific approach. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 9, pp. 163–216). Norwood, NJ: Ablex.

Hunt, E., & Lansman, M. (1986). Unified model of attention and problem solving. *Psychological Review*, *93*, 446–461. https://doi.org/10.1037/0033-295X.93.4.446

James, K. H., Jones, S. S., Smith, L. B., & Swain, S. N. (2014). Young children's self-generated object views and object recognition. *Journal of Cognition and Development*, *15*, 393–401. https://doi.org/10.1080/15248372.2012.749481

James, K. H., Jones, S. S., Swain, S., Pereira, A., & Smith, L. B. (2014). Some views are better than others: Evidence for a visual bias in object views self-generated by toddlers. *Developmental Science*, *17*, 338–351. https://doi.org/10.1111/desc.12124

Karasik, L. B., Tamis-LeMonda, C. S., & Adolph, K. E. (2014). Crawling and walking infants elicit different verbal responses from mothers. *Developmental Science*, *17*, 388–395. https://doi.org/10.1111/desc.12129

Landry, S. H., Smith, K. E., Miller-Loncar, C. L., & Swank, P. R. (1997). Predicting cognitive-linguistic and social growth curves from early maternal behaviors in children at varying degrees of biologic risk. *Developmental Psychology*, *33*, 1040–1053. https://doi.org/10.1037/0012-1649.33.6.1040

Landry, S. H., Smith, K. E., Swank, P. R., Assel, M. A., & Vellet, S. (2001). Does early responsive parenting have a special importance for children's development or is consistency across early childhood necessary? *Developmental Psychology*, *37*, 387–403. https://doi.org/10.1037//0012-1649.37.3.387

Lawson, K. R., Parrinello, R., & Ruff, H. A. (1992). Maternal behavior and infant attention. *Infant Behavior and Development*, *15*, 209–229. https://doi.org/10.1016/0163-6383(92)80024-O

Lockman, J. (2000). A perception-action perspective on tool use development. *Child Development*, *71*, 137–144. https://doi.org/10.1111/1467-8624.00127

Metta, G., & Fitzpatrick, P. (2003). Better vision through manipulation. *Adaptive Behavior*, *11*, 109–128. https://doi.org/10.1177/10597123030112004

Miller, J. L., Ables, E. M., King, A. P., & West, M. J. (2009). Different patterns of contingent stimulation differentially affect attention span in prelinguistic infants. *Infant Behavior and Development*, *32*, 254–261. https://doi.org/10.1016/j.infbeh.2009.02.003

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*, 269–274. https://doi.org/10.1111/j.1467-8721.2007.00518.x

Needham, A. (2000). Improvements in object exploration skills may facilitate the development of object segregation in early infancy. *Journal of Cognition and Development*, *1*, 131–156. https://doi.org/10.1207/S15327647JCD010201

Parrinello, R. M., & Ruff, H. A. (1988). The influence of adult intervention on infants' level of attention. *Child Development*, *59*, 1125–1135. https://doi.org/10.2307/1130279

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review*, *21*, 178–185. https://doi.org/10.3758/s13423-013-0466-4

Richards, J. E. (1989). Development and stability in visual sustained attention in 14, 20, and 26-week-old infants. *Psychophysiology*, *26*, 422–430. https://doi.org/10.1111/j.1469-8986.1989.tb01944.x

Ruff, H. A., & Capozzoli, M. C. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology*, *39*, 877. https://doi.org/10.1037/0012-1649.39.5.877

Ruff, H. A., Capozzoli, M., & Weissberg, R. (1998). Age, individuality, and context as factors in sustained visual attention during the preschool years. *Developmental Psychology*, *34*, 454–464. https://doi.org/10.1037/0012-1649.34.3.454

Ruff, H. A., & Lawson, K. R. (1990). Development of sustained, focused attention in young children during free play. *Developmental Psychology*, *26*, 85–93. https://doi.org/10.1037/0012-1649.26.1.85

Ruff, H. A., Lawson, K. R., Parrinello, R., & Weissberg, R. (1990). Long-term stability of individual differences in sustained attention in the early years. *Child Development*, *61*, 60–75. https://doi.org/10.2307/1131047

Ruff, H. A., & Rothbart, M. K. (2001). *Attention in early development: Themes and variations*. New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195136326.001.0001

Smith, L. B. (2003). Learning to recognize objects. *Psychological Science*, *14*, 244–250. https://doi.org/10.1111/1467-9280.03439

Smith, L. B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: The dynamics of toddler visual experience. *Developmental Science*, *14*, 9–17. https://doi.org/10.1111/j.1467-7687.2009.00947.x

Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: Motor skill acquisition facilitates three-dimensional object completion. *Developmental Psychology*, *46*, 129–138. https://doi.org/10.1037/a0014618

Spencer, J. P., Blumberg, M. S., McMurray, B., Robinson, S. R., Samuelson, L. K., & Tomblin, J. B. (2009). Short arms and talking eggs: Why we should no longer abide the nativist–empiricist debate. *Child Development Perspectives*, *3*, 79–87. https://doi.org/10.1111/j.1750-8606.2009.00081.x

Stechler, G., & Latz, E. (1966). Some observations on attention and arousal in the human infant. *Journal of the American Academy of Child Psychiatry*, *5*, 517–525. https://doi.org/10.1016/S0002-7138(09)62098-7

Suarez-Rivera, C., Smith, L. B., & Yu, C. (2019). Multimodal parent behaviors within joint attention support sustained attention in infants. *Developmental Psychology*, *55*, 96–109. https://doi.org/10.1037/dev0000628

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science*, *23*, 121–126. https://doi.org/10.1177/0963721414522813

Tamis-LeMonda, C. S., Kuchirko, Y., & Tafuro, L. (2013). From action to interaction: Infant object exploration and mothers' contingent responsiveness. *IEEE Transactions on Autonomous Mental Development*, *5*, 202–209. https://doi.org/10.1109/TAMD.2013.2269905

Tattersall, A. J. (1998). Individual differences in performance. In M. W. Smolensky & E. S. Stein (Eds.), *Human factors in air traffic control* (pp. 185–213). San Diego, CA: Academic Press.

Thurman, S. L., & Corbetta, D. (2017). Spatial exploration and changes in infant–mother dyads around transitions in infant locomotion. *Developmental Psychology*, *53*, 1207–1221. https://doi.org/10.1037/dev0000328

Vaccari, C., & Marschark, M. (1997). Communication between parents and deaf children: Implications for social–emotional development. *Journal of Child Psychology and Psychiatry*, *38*, 793–801. https://doi.org/10.1111/j.1469-7610.1997. tb01597.x

Wass, S. V., Clackson, K., Georgieva, S. D., Brightman, L., Nutbrown, R., & Leong, V. (2018). Infants' visual sustained attention is higher during joint play than solo play: Is this due to increased endogenous attention control or exogenous stimulus capture? *Developmental Science*, *21*, e12667. https://doi.org/10.1111/desc.12667

Wass, S. V., Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R., . . . Leong, V. (2018). Parental neural responsivity to infants' visual attention: How mature brains influence immature brains during social interaction. *PLoS Biology*, *16*, e2006328. https://doi.org/10.1371/journal.pbio.2006328

Xu, T., Chen, Y., & Smith, L. (2011). It's the child's body: The role of toddler and parent in selecting toddler's visual experience. In *Development and Learning (ICDL), 2011 IEEE International Conference* (Vol. 2, pp. 1–6). New York, NY: IEEE.

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head-camera to study visual experience. *Infancy*, *13*, 229–248. https://doi.org/10.1080/15250000802004437

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*, 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Yu, C., & Smith, L. B. (2016). The social origins of sustained attention in one-year-old human infants. *Current Biology*, *26*, 1235–1240. https://doi.org/10.1016/j.cub.2016.03.026

Yu, C., & Smith, L. B. (2017). Hand–eye coordination predicts joint attention. *Child Development*, *88*, 2060–2078. https://doi.org/10.1111/cdev.12730

Yu, C., Smith, L. B., Shen, H., Pereira, A., & Smith, T. (2009). Active information selection: Visual attention through the hands. *IEEE Transactions on Autonomous Mental Development*, *2*, 141–151. https://doi.org/10.1109/TAMD.2009.2031513