UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Multimodal Behaviors from Children Elicit Parent Responses in Real-Time Social Interaction

Permalink

https://escholarship.org/uc/item/2qw0d216

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN

1069-7977

Authors

Yurkovic, Julia R
Kennedy, Daniel P
Yu, Chen

Publication Date

2021

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Multimodal Behaviors from Children Elicit Parent Responses in Real-Time Social Interaction

Julia R. Yurkovic[1], Daniel P. Kennedy[1+], and Chen Yu[1,2+]
[1] Indiana University, Department of Psychological and Brain Sciences, Bloomington, IN, 47401, USA
[2] Department of Psychological and Brain Sciences, University of Texas at Austin, Austin, TX 78712, USA
[+] These authors contributed equally to this work

Abstract

Social interactions serve as the primary training ground for many of a child's positive cognitive and developmental abilities. Parent responsiveness has been identified as one key mechanism through which children gain more mature skills, but the question of how children elicit responses from their parents remains to be fully investigated. In this study, we utilized head-mounted eye trackers to track moment-by-moment shifts in gaze, manual action, and speech during parent-child toy play. This allowed us to identify the moments preceding a parent response and the type and timing of the parent's response relative to the child's behaviors. We found that child events of attention and action – where they were both touching and looking at a toy – were more successful in eliciting parent responses overall and in eliciting multimodal parent responses than events of just child look or child touch. The parent's latency to respond to their child differed by event type and duration, suggesting that child behaviors influence parent responses. Implications and future directions are discussed.

Keywords: head-mounted eye tracking; dyadic toy play; parent responsiveness; child attention and action

Introduction

Coordinated social interactions between a parent and child serve as the primary training ground for many of the child's core cognitive abilities: sustained attention (Bornstein & Tamis-Lemonda, 1997; Suarez-Rivera et al., 2019; Yu & Smith, 2016), play abilities (Bornstein & Tamis-Lemonda, 1997), language development (Goldstein et al., 2003; Prime et al., 2020; Tamis-LeMonda et al., 2014), and general cognitive abilities (Landry et al., 2001). A necessary pre-requisite to understanding the mechanisms through which mature social partners influence a child's cognitive development is to understand if, when, and how the parent responds to their child. Formally, parent responsiveness is defined as a *prompt*, *contingent*, and *appropriate* behavioral response of the parent following a child's action (Ainsworth et al., 1974; Bornstein & Tamis-Lemonda, 1997; Bornstein et al., 2008). Positive development in child abilities following parental responses can be observed both in the moment and over developmental time, suggesting that parental responses to child behaviors may scaffold the continued development of more advanced skills (Goldstein et al., 2003; Landry et al., 2001; Tamis-LeMonda et al., 2014).

The study of parental responsiveness has traditionally adopted a unidirectional focus that highlights the role of the parent in responding to their child, but not the role of the child in *eliciting* responses from their parent. However, parent-child interactions can be viewed as a coupled system that builds upon both the parent's and the child's abilities and actions (Richardson & Dale, 2005; Richardson et al., 2007; Smith & Thelen, 2003; Thelen & Smith, 2007). It is likely, then, that children play a critical role in the rate and type of responses that they obtain from their parents. The results of a few studies offer support for this hypothesis by suggesting that the child's language abilities, attention span, and rate of manual actions predicted the rate of parent responsiveness (Bornstein & Tamis-Lemonda, 1997; Brigham et al., 2010; McDuffie & Yoder, 2010; Siller & Sigman, 2002). However, these studies tend to use global ratings of parent responsiveness that assume a consistent parenting style rather than one that adapts to the needs of the child (Bornstein & Tamis-LeMonda, 1990). Therefore, the moment-by-moment changes in child behaviors that elicit different types of parent responses remain to be fully investigated.

Two types of eliciting behavior that organize the parent's behavior are the child's gaze and manual actions towards toys. Studies utilizing head-mounted eye trackers to quantify the dynamics of parent-child interactions have shown that parents most often follow their child's gaze to objects, but are quicker to look at the object of the child's attention when they follow the child's hands (Yu & Smith, 2013, 2017a). The amount of time that children manually act on objects correlates with the amount of time that they spend jointly engaging with a toy with a parent (Elmlinger et al., 2019). The child's own manual actions also serve to extend their attention to toys (Yurkovic et al., 2020). Further, the child's coordinated visual attention and manual action creates more optimal learning opportunities for the child – the acted upon object becomes large and centered in the child's view and can reduce the complexity of mapping the parent's spoken words to objects (Bambach et al., 2016, 2017). Additionally, parents tend to coordinate their naming instances with the child's manual action or hand-eye coordination more than the child's visual attention (Chang et al., 2016; West & Iverson, 2017).

Additionally, the type of parent response (look, touch, or naming the object) as well as the amount of parent response (uni- or multimodal cues) have been identified as important factors for the child's development. Children extend their attention to toys when their parent looks to the same toy (Yu & Smith, 2016), especially when the parent is also talking or is holding the toy (Suarez-Rivera et al., 2019; Yurkovic et al., 2020). Further, parent responses that use multimodal cues have been linked to positive scaffolding of child language development (Tamis-LeMonda et al., 2014).

In this study, we examine the coordination of parent and child behaviors, specifically focusing on moments where the

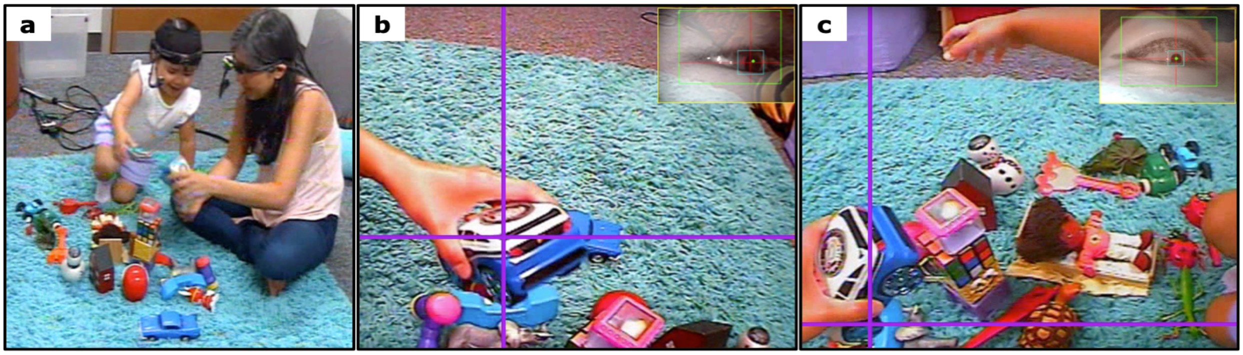

Figure 1: *Experimental Set-Up.* (A) The parent and child were both equipped with a head-mounted eye tracker. Parents and children were given 24 toys to play with. Scene cameras, located on the center of the forehead, capture the (B) child and (C) parent scene views. Eye cameras directed towards the right eye captured eye movements. Images of the eye view are visible superimposed in the upper right corners of the (B) child and (C) parent views. (B-C) Calibration procedures result in a cross-hair on the scene view that represents the location of the eye gaze at each moment of play.

child's actions prompt parent responses. Using head-mounted eye trackers, we identify parent responses that follow the child's attention and action and are time-locked to the child's behavior. We investigate how child look, touch, and look-touch (hand-eye coordination) differentially relate to parent responses. We will then examine the *promptness* and *contingency* of the parent response to the child's behavior, and how these features may differ in relation to the child's moment-by-moment behaviors. Specifically, we will analyze the rate of response, the *promptness* of the response (measured as latency from child attention/action onset to parent response), and *contingency* of the response (measured as the type and modality of the parent's response).

Methods

Participants

Thirty-nine dyads (15 female) contributed data to the current study. Participants ranged from 12 to 48 months old. Five participants were excluded from analyses because we were unable to collect or transcribe parent speech data. An additional two children were excluded because they were not tolerant of either wearing the eye tracker or allowing research personnel to make the necessary adjustments. One additional participant was excluded for vision concerns.

Stimuli and Procedure

Figure 1 shows the experimental set-up. Parents and children were seated on the floor in a room resembling a playroom. Both dyad members were equipped with head-mounted eye trackers that allowed us to measure moment-by-moment gaze shifts during play. The parent eye tracker resembled a pair of glasses and the child eye tracker was affixed to a soft, adjustable hat. The eye trackers contained two cameras: an infrared camera pointed towards the right eye to detect eye movements and a camera in the center of the forehead to record the scene in front of the dyad member. Camera positions were adjusted until a clear eye image and scene view were attained. A laser pointer was then used to direct both the parent and child's gaze to toys across both a vertical and horizontal axis in the play space. This procedure provided a starting point for later eye gaze calibration. Additional cameras spread throughout the room captured third-person camera views of the interaction.

Following head-mounted eye tracker set-up, parents and children were given 24 everyday toys to play with, including blocks, a doll, cars, and animals. The toys were spread out on the carpet so that both the parent and child could see all of them. Dyads were told to play together like they might at home for 5-10 minutes. The interaction was only interrupted if the eye trackers were bumped and no longer provided a clear eye or scene view, in which case the experimenters would pause the session and adjust the eye tracker.

Data Processing

Eye Gaze Calibration. After the session, Yarbus software (Positive Science, LLC) was used to map the child's eye movements from the eye camera to the objects in their first-person view. Trained experimenters moved through the synchronized videos frame-by-frame, indicating points on the scene view where the child was looking. Points were distributed throughout the x- and y-coordinates of the video and across time to ensure an accurate calibration. This process was repeated until the intra-point correlation was greater than 95%. Following current best-practices (Hayhoe et al., 2012), we repeated the calibration procedure 2-3 times to reach a final calibration. The procedure was repeated for the parent. A crosshair was generated on top of the scene videos to indicate the point of gaze (Figure 1b-c).

Region of Interest and Speech Coding. Following eye-gaze calibration, trained coders indicated the region-of-interest (ROI) of the participant's gaze. Using a custom program, the continuous gaze stream was broken into individual looks based on the velocity of eye movements. Coders then used the crosshair generated after calibration to determine the ROI of each individual look. Coders were able to see all camera views, including the eye camera, during ROI coding to ensure

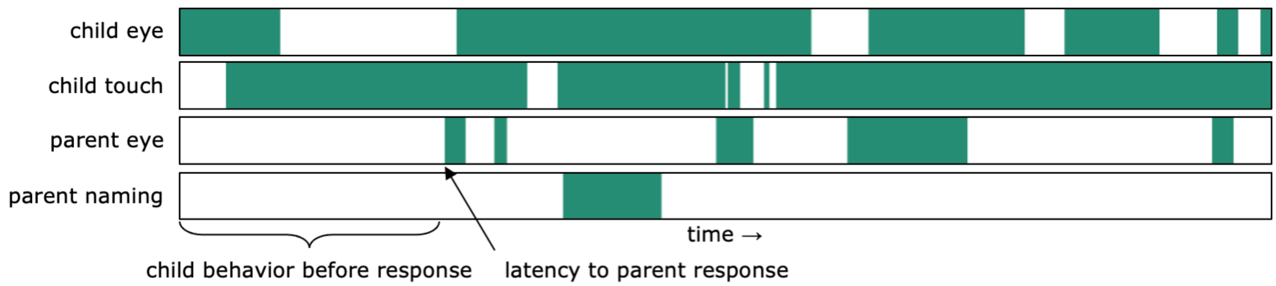

Figure 2: *Child Event Identification and Parent Responses.* The child's behaviors (look, touch) and parent's behaviors (look, touch, naming) are visualized as data streams over time. Child events were defined as an unbroken sequence of attention and action to a single toy. The current event is categorized as a Look-Touch event because of the presence of both child look and touch preceding parent response. The latency to parent response is identified as the time from the onset of the child event to the time of parent first action towards the toy of the child's interest.

that they were correctly identifying each ROI. Possible ROIs included the 24 toys or the other dyad member's face.

Additionally, ROI coding was completed for objects that came into contact with each dyad member's hands. Trained coders annotated a data stream for each hand individually, indicating which of the 24 toys was being touched at any moment during play. A toy was marked as being touched if the participant's hand came into contact with the toy. A second coder independently coded a randomly selected subset of 5 participants and the inter-coder reliability ranged from 93% to 99% (Cohen's kappa=0.96). The child touch data streams were later merged to find events where one or both hands were in contact with a toy.

Finally, parent speech was coded using Audacity. Trained coders indicated when the parent was speaking during the play interaction, and then transcribed the parent's utterances. Speech was segmented into separate utterances if there was a silence of 400ms between speech sounds. Utterances containing a label to one of the 24 toys were considered naming utterances and are used in subsequent analyses.

Data Analysis

Child Event Identification & Classification. We first identified all *events* of child interest in a toy, defined as an unbroken series of looks to and touches of a toy. A single event may consist of several overlapping looks to and touches of the toy (Figure 2). Importantly, only events where the child initiated the interest in the toy were considered – events where the child interacted with a toy following a parent cue towards that toy were excluded. Events were categorized based on the behaviors occurring *before* the parent response or, in the cases of no parent response, as the behavior throughout the entire event. This approach ensured that only the child behaviors that parents had the opportunity to respond to were considered. Look events are those where the child only looked at the toy before parent response or for the entire event without parent response. Touch events are those where the child only touched the toy. Look-Touch events are those where the child both looked at and touched the toy.

Parent Response. We defined parent responses as any behavior from the parent – look, touch, or naming – directed towards the object of the child's interest. Responses began after the onset of the child event (i.e., did not lead the child's attention) and began before the offset of the child event. The latency of the parent to respond (Figure 2) and types of parent behaviors used in the response were recorded.

Permutation Testing. We ran a permutation test to assess chance levels of parent response by child event type. On each of 1000 permutations, we assigned a parent response type to each event that was sampled from the overall probability distribution. We calculated chance levels of overall response rate, type of parent response, and count of modalities in the parent response by computing the 95% confidence interval (CI) of the permuted distribution.

Results

Rate and Duration of Child Event Types

We first aimed to determine what types of behaviors children generate for their parents to respond to. Children produced an average of 25.87 events per minute of play (SD=7.89, range=10.11-43.31). All child events were entered into a corpus analysis, resulting in 5,839 events. Of the 5,839 events generated by all children, 4363 (75%) were characterized as Look only events, 877 (15%) were Touch only events, and 599 (10%) were Look-Touch events.

An analysis of variance (ANOVA) on a linear mixed effects model (LME) of duration with fixed effects of event type and random effects for participant revealed a main effect of event type, $F(2,5836)=359.92$, $p<0.01$. The mean duration of Look events was 1.26s (SD=3.96s), of Touch events was 3.86s (9.48s), and of Look-Touch events was 12.31s (25.21s). Consistent with past research demonstrating that looks during naturalistic toy play unfold on a quicker timescale than touches (Yurkovic et al., *In press*), post-hoc analyses revealed that Look events tended to be shorter than Touch events, $t(5238)=-12.93$, $p<0.01$. Look-Touch events were longer than both Look events, $t(4960)=-26.49$, $p<0.01$, and Touch events, $t(1474)=-9.09$, $p<0.01$, reflecting that child attention is related to child action (Yurkovic et al., 2020).

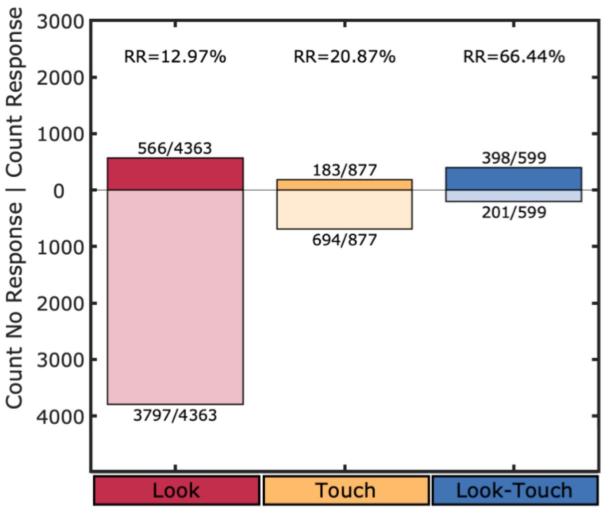

Figure 3: *Response Rate to Child Event Types.* The y-axis shows the count of events that did not receive a response (lightly colored regions) compared to the count of events that did receive a response (darker colored regions) overall and for each event category. The response rate (RR) is shown above each event type. Look events and Touch events were overall less successful at eliciting parent responses compared to Look-Touch events.

Parent Response Rate

We next determined which child behaviors elicited parent responses (i.e., look, touch, and/or naming) (Figure 3). Parents responded to 19.64% of child events overall. A χ^2 test revealed that different event types related to greater parent responsiveness, $\chi^2(1,N=5839)=954.99$, $p<0.001$. Despite making up the largest percentage of events, Look events received a parent response only 12.97% of the time, significant less than the 20.87% response rate to Touch events, $\chi^2(1,N=5240)=37.14$, $p<0.001$, and the 66.44% response rate to Look-Touch events, $\chi^2(1,N=4962)=962.04$, $p<0.001$. Additionally, Touch events received fewer responses than Look-Touch events, $\chi^2(1,N=1476)=309.75$, $p<0.001$. This result is consistent with past literature that the child's manual actions, specifically hand-eye coordination, signal the opportunity for parents to respond.

We next compared actual response rates to those expected by chance. The parent response rate to child Look events is lower than expected by chance, 95% CI=19-20%, Cohen's $d=34.92$, $p<0.001$. The response rate to child Touch events is within chance levels, 95% CI=19-22%, Cohen's $d=0.93$, $p=0.15$. Finally, the response rate to child Look-Touch events is above that expected by chance, 95% CI=19-23%, Cohen's $d=45.87$, $p<0.001$.

The child's gaze alone elicits low levels of parent responsiveness relative to both other child event types and chance. Child touch elicits more parent responsiveness than the child's gaze but does not elicit parent responses above what we would expect by chance throughout the interaction. Finally, the child's hand-eye coordination elicits a parent response rate that is greater than both other event types and is well above what is expected by chance. The child's own coordination of attention with action appears to signal the opportunity to respond to the parents.

Child Duration and Parent Responsiveness

We ran a generalized linear model of parent response as a function of child event duration and child event type. There was a significant effect of duration ($\beta=0.83$, $p<0.001$) such that parent response became more likely as duration increased. Additionally, there was a significant duration by child event type interaction ($\beta s>|0.95|$, $ps<0.001$). In other words, while duration does predict parent responsiveness, child event type predicts parent responsiveness above and beyond the role of duration. It should be noted that parent response may extend the child's attention (Yu & Smith, 2016) thereby contributing to the relationship between child event duration and parent responsiveness.

Parent Latency to Response

Considering only the events that received a response, we next determined how quickly parents responded to their child (Figure 4). Parent response latency was defined as the time from the onset of the child event to the first parent behavior (look, touch, or naming). The overall median response latency was 0.83s (SD=3.78s). An ANOVA on an LME of parent latency with fixed effects for child event type and random effects for participant revealed that response latency differed by event type, $F(2,1144)=93.61$, $p<0.01$). Look events had the shortest parent response latencies overall, 0.43s (0.69s), followed by Touch events, 1.00s (3.57s), followed by Look-Touch events 2.20s (5.38s). The latencies are most likely a function of the duration of child events, of which Look events are the shortest and Look-Touch events are the longest. Overall, parent attention and action are well-coordinated with the attention and action of the child.

Types of Parent Responses

We next compared which behaviors parents used to respond to their child: look, touch, or naming. A χ^2 test revealed a significant difference in the count of all child events that received a parent look, touch, or naming response, $\chi^2(2,N=3441)=1389.23$, $p<0.001$. We conducted post-hoc, pairwise $\chi2$ tests to determine if there was a relationship between parent response type and child event type. We also assessed chance levels of parent response types (Figure 5a).

Parent look responses were generated in response to child Look-Touch events more frequently than to child Look ($\chi^2(1,N=964)=18.66$, $p<0.001$) and child Touch ($\chi^2(1,N=581)=5.73$, $p=0.02$) events, which did not differ ($\chi^2(1,N=749)=1.28$, $p=0.26$). Child hand-eye coordination elicits more parent looking than other child behaviors.

Parent touch responses were generated equally to child Look, Touch, and Look-Touch events, $\chi^2(1,N=1147)=0.93$, $p=0.63$. Parent touch may be used in different ways (i.e.,

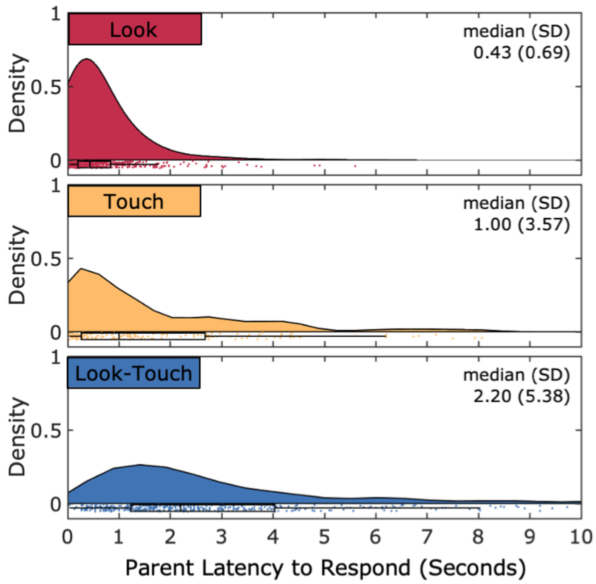

Figure 4: *Latency to Parent Response*. Raincloud plots represent the kernel density (y-axis) of latency to parent response (x-axis) and box plots below show the interquartile range. Latencies greater than 10s (n=43) are not depicted. Parent latency to response is quickest for Look events, followed by Touch events, followed by Look-Touch events.

moving a toy closer, helping the child use the toy) that differ by event type and are not captured in the current analysis.

Finally, parent naming was generated to child Look events less frequently than to Touch (χ^2(1,N=749)=5.04, p=0.02) and Look-Touch events ((χ^2(1,N=964)=27.01, p<0.001). Child Touch and Look-Touch events did not differ in the rate of parent naming response, χ^2(1,N=581)=2.83, p=0.09. The presence of touch in a child's behavior elicits more parent naming than the absence of touch.

We assessed chance levels of each modality of parent response to each child event type. We found that child Look events were less likely than expected by chance to receive a parent eye (d=7.59, p<0.001) or parent naming response (d=9.73, p<0.001). In contrast, Look-Touch events were more likely than expected by chance to receive a parent eye (d=2.17, p=0.02) or parent naming response (d=2.51, p=0.02). Child Touch events received responses at chance levels (d=0.14–0.87, all p>0.23). Child Look alone relates to less parent mutual gaze and naming opportunities compared to Look-Touch events and compared to chance. By contrast, child Look-Touch events relate to greater parent mutual gaze and naming than compared to other event types and compared to chance – further relating to more word learning opportunities for the child.

Count of Parent Response Modalities

We then compared how many behaviors parents used when responding to their child's interest: one, two, or all three parent behaviors. A χ^2 test revealed a significant difference in the count of all child events that received a parent response using one, two, or three behaviors, χ^2(2,N=3441)=1008.63, p<0.001. We conducted post-hoc, pairwise χ2 tests to determine if there was a relationship between the number of modalities in a parent response type and child event type. We also assessed chance levels of parent responses (Figure 5b).

Unimodal parent responses were less likely to be generated in response to child Look-Touch events than to Look events (χ^2(1,N=964)=26.87, p<0.001) and Touch events (χ^2(1,N=581)=4.50, p=0.04), which did not differ (χ^2(1,N=749)=3.06, p<0.08). Unimodal child behaviors were more likely to generate unimodal parent behaviors.

By contrast, multi-modal parent responses were more likely to be generated in response to child Look-Touch events than child Look (Two-Behavior: χ^2(1,N=964)=17.26, p<0.001; Three-Behavior: χ^2(1,N=964)=5.83, p=0.02) events. The rate of multi-modal parent responses to child Touch events did not differ from response rate to Look (Two-Behavior: χ^2(1,N=749)=3.58, p=0.59; Three-Behavior: χ^2(1,N=749)=0, p=0.99) or Look-Touch events (Two Behavior: χ^2(1,N=581)=1.56, p=0.21; Three-Behavior: χ^2(1,N=581)=2.76, p=0.10). Child Look-Touch events generate richer, more multi-modal parent responses than child Look only events.

We assessed chance levels of the different counts of modalities of parent responses to each child event type. Consistent with cross-event type analyses, child Look events were more likely than expected by chance to receive a unimodal parent response (d=9.34, p<0.001) and less likely than expected by chance to receive a multimodal parent response (d=3.76-7.99, all p<0.002). By contrast, child Look-Touch events were less likely than expected by chance to receive a unimodal parent response (d=2.61, p=0.01) and more likely than expected by chance to receive a multimodal parent response (d=1.41-2.01, all p<0.04). Multimodal child behaviors relate to greater-than-chance levels of parent multimodal responses, suggesting that child behaviors can serve to organize parent responses.

Discussion

The current study aimed to determine if different child behaviors elicited different parent responses and, if so, how the parent responses differed. We utilized head-mounted eye trackers during toy play to capture parent and child moment-by-moment gaze shifts and manual activity, as well as parent speech. We found, overall, that children use multimodal behaviors to elicit parent responses at a rate greater than expected by chance and that parent responses differ in the amount and type of modalities used.

We found that events of child attention that included both look to and touch of an object (Look-Touch) were infrequent but highly successful in eliciting parent responses relative to other event types and to chance. The infrequency of these events may be related to the length of the events, such that there are few, extended moments of coordinated child action and attention. Parents took longer to respond to child Look-Touch events but tended to use multimodal responses that can

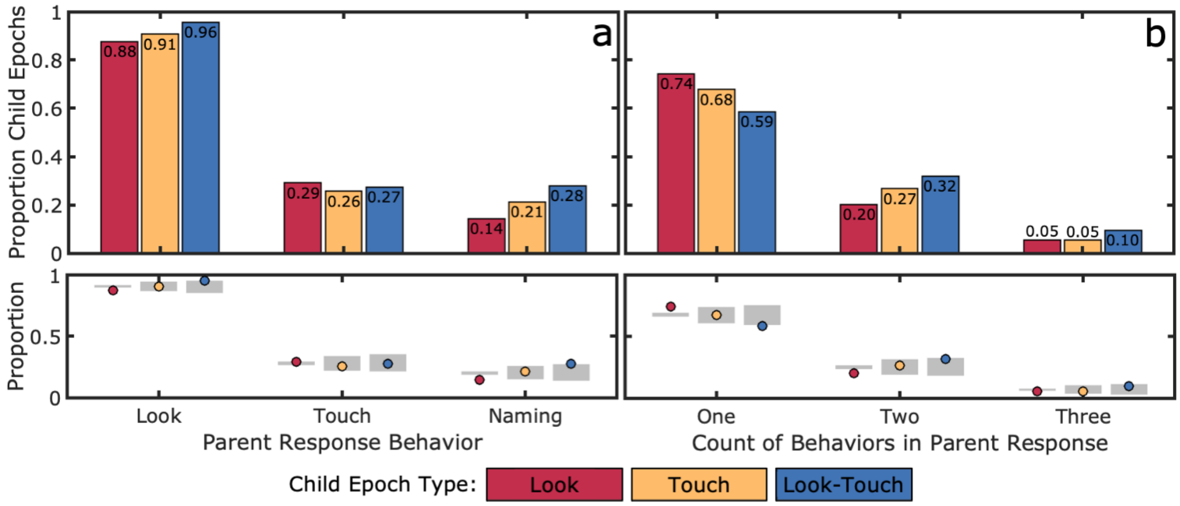

Figure 5: *Type and Amount of Parent Response*. (A) We examined the use of three parent response modalities: look, touch, and naming. The sum of the proportions across event types may not equal one because some events received multiple response types. Look events were responded to differently than Touch and Look-Touch events. The bottom panel shows the chance distributions in gray and the actual proportions in the colored dots. (B) We examined the multi-modality of parent responses by child behavior type. Look events were responded to unimodally more frequently and responded to multimodally less frequently than Look-Touch events. The bottom panel shows chance compared to actual response proportions.

provide a rich experience for the child, including increased rates of naming toys that can support child language learning. The long latency to respond and the use of multimodal cues most likely relate to the overall long duration of Look-Touch events. The child's coordination of their attention with their actions constrains and extends their attention (Bambach et al., 2016, 2017; Yurkovic et al., 2020). This extension provides the parent with an extended window within which to recognize the object of the child's attention and to respond (Elmlinger et al., 2019; Yu & Smith, 2013, 2017a), as well as providing opportunities for extended parent speech discourse about an object (Suanda et al., 2016). By coordinating their own actions, children provided more of an opportunity for parents to respond to their actions.

Further, we found that child events that included only Look before the parent response were the most common but were relatively ineffective at eliciting a parent response. Look events tended to be shorter and, when parents did respond, they did so quickly and less frequently with mutual gaze or naming of the object and more likely with unimodal responses compared to their responses to Look-Touch events and to chance levels of response types. Past research has shown that looks generated to toys while children are holding a different toy tend to be short and exploratory (Yurkovic et al., 2021), and many of the looks in the Look-only category may fall into that camp (i.e., quickly glancing at a nearby toy). A parent response to these looks may serve to redirect the child's attention from their primary focus, an action that is thought to have overall negative consequences for child development (Landry et al., 2001; Tomasello & Todd, 1983).

Finally, Touch-only events were relatively infrequent, were relatively ineffective at eliciting parent responses, and elicited response types that were more similar to Look-Touch events but were at chance levels of receiving different types and parent responses. Child touch, either with or without hand-eye coordination, may relate to some organization of parent attention above and beyond child look alone. Similar to child look, child touch alone does not relate to greater mutual gaze and parent naming instances that can support more positive developmental outcomes.

There are two key limitations to the current study. First, we examined a wide age range of children in a corpus analysis. Future analyses will aim to quantify how child age may impact the type and frequency of parent responsiveness, specifically as children transition through different physical (i.e., crawling to walking) and cognitive (i.e., play and sustained attention abilities) developmental stages. Additionally, there is some level of inaccuracy that is inherent when translating head-mounted eye tracking of a participant's gaze in a 3-dimensional space to a 2-dimensional screen for post-processing. Coders are trained to account for depth when processing the eye gaze data and both child and parent actions can be used to infer point of gaze. However, some looks may be coded improperly.

Future analyses will investigate individual differences in child behaviors and parent responses. Previous work has shown that when children do not frequently coordinate their gaze with their hands, parents do not frequently follow their child's hands into joint engagement, suggesting that parents are sensitive to the cues of their child (Yu & Smith, 2017b). We will use clustering analyses to identify shared patterns within children (i.e., types of cues), within parents (i.e., types of responses), or within dyads (i.e., unique behavior-response profiles). Additionally, future analyses will aim to capture the dynamics of child action, parent response, and child further action following the parent response. Research has demonstrated that children generate more sustained manual actions on toys when their parents are being more responsive

(McQuillan et al., 2019), highlighting the transactional nature of parent-child interactions.

References

Ainsworth, M. D., Bell, S. M., & Stayton, D. (1974). *Infant-mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals.*

Bambach, S., Crandall, D. J., Smith, L. B., & Yu, C. (2016). Active viewing in toddlers facilitates visual object learning: An egocentric vision approach. *Proceedings of the 38th Meeting of the Cognitive Science Society.*, 2-7.

Bambach, S., Smith, L. B., Crandall, D. J., & Yu, C. (2017). Objects in the center: How the infant's body constrains infant scenes. *2016 Joint IEEE ICDL-EpiRob*, 132-137.

Bornstein, M. H., & Tamis-Lemonda, C. S. (1997). Maternal responsiveness and infant mental abilities: Specific predictive relations. *Infant Behav. Dev.*, *20*(3), 283-296.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1990). Activities and interactions of mothers and their firstborn infants in the first six months of life: Covariation, stability, continuity, Correspondence, and Prediction. *Child Dev.*, *61*(4), 1206.

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. *Dev. Psych*.

Brigham, N. B., Yoder, P., Jarzynka, M., & Tapp, J. (2010). The sequential relationship between parent attentional cues and sustained attention to objects in young children with autism. *J Autism Dev. Disord.*, *40*(2), 200-208.

Chang, L., de Barbaro, K., & Deák, G. (2016). Contingencies Between Infants' Gaze, Vocal, and Manual Actions and Mothers' Object-Naming: Longitudinal Changes From 4 to 9 Months. *Dev. Neuropsychol.*, *41*(5-8), 342-361.

Elmlinger, S. L., Suanda, S. H., Smith, L. B., & Yu, C. (2019). Toddlers' hands organize parent-toddler attention across different social contexts. *2019 Joint IEEE 9th ICDL-EpiRob*, 296-301.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *PNAS*, *100*(13), 8030-8035.

Hayhoe, M. M., McKinney, T., Chajka, K., & Pelz, J. B. (2012). Predictive eye movements in natural vision. *Exp. Brain Res.*, *217*(1), 125-136.

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? *Dev. Psych.*, *37*(3), 387-403.

McDuffie, A. S., & Yoder, P. J. (2010). Types of parent verbal responsiveness that predict language in young children with autism spectrum disorder. *J. Speech, Lang., Hear.*, *53*(4), 1026-1039.

McQuillan, M. E., Smith, L. B., Yu, C., & Bates, J. E. (2019). Parents influence the visual learning environment through children's manual actions. *Child Dev.*, *00*(0), 1-20.

Prime, H, Wade, M, & Gonzalez, A (2020). The link between maternal and child verbal abilities: An indirect effect through maternal responsiveness. *Dev. Sci.*, *23*(3), 1-14.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cogn. Sci.*, *29*(6), 1045-1060.

Richardson, D. C, Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psy. Sci.*, *18*(5), 407-413.

Siller, M., & Sigman, M. (2002). The behaviors of parents of children with autism predict the subsequent development of their children's communication. *J Autism Dev. Disord.*, *32*(2), 77-89.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *TiCS*, *7*(8), 343-348.

Suanda, S. H., Smith, L. B., & Yu, C. (2016). The multisensory nature of verbal discourse in parent-toddler interactions. *Dev. Neuropsychol.*, *41*(5-8), 324-341.

Suarez-Rivera, C., Smith, L. B., & Yu, C. (2019). Multimodal parent behaviors within joint attention support sustained attention in infants. *Dev. Psych.*, *55*(1), 96-109.

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Cur. Dir. in Psychol. Sci.*, *23*(2), 121-126.

Thelen, E., & Smith, L. B. (2007). Dynamic systems theories. *Handbook of Child Psychology*, 258-312.

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, *4*(12), 197-211.

West, K. L., & Iverson, J. M. (2017). Language learning is hands-on: Exploring links between infants' object manipulation and verbal input. *Cog. Dev.*, *43*, 190-200.

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS ONE*.

Yu, C., & Smith, L. B. (2016). The social origins of sustained attention in one-year-old human infants. *Curr. Bio.*, *26*(9), 1235-1240.

Yu, C., & Smith, L. B. (2017a). Multiple sensory-motor pathways lead to coordinated visual attention. *Cogn. Sci.*, *41*, 5-31.

Yu, C., & Smith, L. B. (2017b). Hand-Eye coordination predicts joint attention. *Child Dev.*, *88*(6), 2060-2078.

Yurkovic, J. R, Lisandrelli, G., Shaffer, R. C., Dominick, K. C., Pedapati, E., Erickson, C. A., Yu, C., & Kennedy, D. P. (In Press). Using head-mounted eye tracking to examine visual and manual exploration during naturalistic toy play in children with and without autism. *Sci. Rep*.

Yurkovic, J. R., Lisandrelli, G., Shaffer, R. C., Dominick, K. C., Pedapati, E. V., Erickson, C. A., Kennedy, D. P., & Yu, C. (2020). Examining sustained attention in child-parent interaction: A comparative study of typically developing children and chidlren with autism spectrum disorder. *Proceedings of the 42nd Meeting of the Cog Sci Society*.