

# Toddlers’ Hands Organize Parent-Toddler Attention across Different Social Contexts

Steven L. Elmlinger  
Psychology Department  
Cornell University  
Ithaca, New York - USA  
sle64@cornell.edu

Sumarga H. Suanda  
Department of Psychological  
Sciences  
University of Connecticut  
Storrs, Connecticut - USA  
s.suanda@uconn.edu

Linda B. Smith  
Department of Psychological and  
Brain Sciences  
Indiana University  
Bloomington, Indiana - USA  
smith4@indiana.edu

Chen Yu  
Department of Psychological and  
Brain Sciences  
Indiana University  
Bloomington, Indiana - USA  
chenyu@indiana.edu

**Abstract**— Toddlers and their parents achieve joint attention in many different social contexts. In some contexts, parents follow toddlers’ attention; in other contexts, toddlers follow parents. Using a dual head-mounted eye-tracking paradigm and micro-level analyses of behavior, we examined the sensorimotor properties of parent-toddler joint attention both in episodes where parents followed their toddlers’ focus of attention and episodes where parents directed their toddlers’ attention. Our results revealed that across both contexts the degree to which parents and toddlers engaged in sustained joint attention was predicted by toddlers’ manual engagement with the target object. These results deepen our understanding of the sensorimotor and micro-level processes that shape joint attention and underscore the interconnections between early motor and social developments.

**Keywords**—parent-child interaction, embodied cognition, social cognition, joint attention

## I. INTRODUCTION

Parents’ and toddlers’ ability to establish and sustain joint attention is foundational to several facets of development. Studies of parent-toddler joint attention have revealed, for example, that joint attention abilities are associated with more mature cognitive development [1], more precocious language abilities [2], and more advanced socio-cognitive reasoning [3]. Recent research has even revealed that joint attention positively shapes the moment-to-moment ebbs and flows of toddler visual attention [4]. A deep understanding of just how parents and toddlers achieve joint attention is thus an important and worthwhile pursuit. Such an understanding is the overarching goal of the current work.

Previous research on this topic suggests that joint attention is more likely to be achieved when parents follow into their toddlers’ focus of attention than when parents attempt to direct their toddlers’ focus of attention [5,6]. Recent sensorimotor and micro-level analyses of joint attention [7-9] might help explain why parent following is so potent in creating and sustaining joint attention. That is, toddlers’ manual actions on an object always attract parents’ attention to the object and therefore establish and sustain joint attention between the two social partners [7,8]. In one study for

example, Yu and Smith [8] found that individual differences in how toddlers attended to objects in their own hands predicted how much time toddlers and their parents spent in joint attention. If parent following of toddler attention is defined, at least in part, by the fact that toddlers are manually engaged with and focused on an object prior to the parent joining in on that focus [2,10], then toddler manual action may be the key for parents to follow toddlers’ attention to establish joint attention.

Although the power of parent following of toddlers’ attention for achieving joint attention is well documented, parents and toddlers are also able to establish joint attention in other contexts wherein parents *direct* their toddlers’ attention (see Figure 1). Even though some studies have explored the development of this ability in controlled laboratory experiments [11], little is known about how moments of parent directing create joint attention in free-flowing dyadic interactions. Analyzing the sensorimotor dynamics that create joint attention when parents direct is particularly interesting because unlike moments wherein parents follow, parents’ directing involves more *parent* than toddler manual engagement with the target object [7]. Thus, an analysis of how joint attention is shaped in moments of parent directing might reveal different sensorimotor properties of joint attention that do not depend on toddlers’ manual engagement. Alternatively, it is also possible that although parent directing and parent

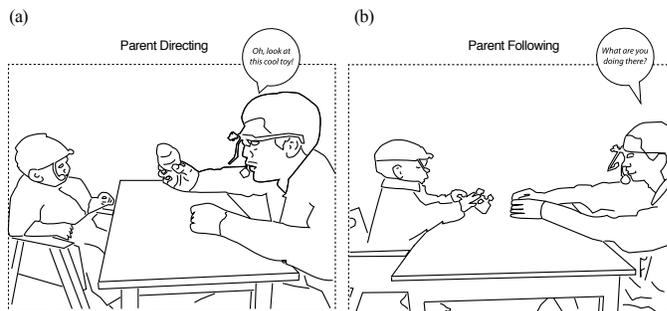


Figure 1. (a) Parents directing toddler’s attention to new objects. (b) Parents following their toddler’s attention.

following contexts differ in their initial states (in the former context parents start by trying to change their toddlers' focus of attention; in the latter context parents start by joining their toddlers' focus of attention), toddlers' manual engagement with the target object may still be key to creating sustained joint attention in *both* parent following and parent directing contexts.

## II. STUDY AND ANALYTIC OVERVIEW

To distinguish between these possibilities, we studied the sensorimotor dynamics of parents and toddlers as they engaged in free play with a set of toy objects. During the course of free-flowing object-play, we asked parents to occasionally attempt to direct their toddlers away from their focus of attention and toward a new object, creating opportunities to investigate parent directing contexts along with the more commonly occurring parent following contexts. Of particular interest was what led to sustained joint attention in directing contexts, and whether the correlates to joint attention in directing were different from the correlates to joint attention in parent following contexts. Given the previous research highlighted above [7-9], we focused on the role (and possibly lack thereof) of toddler object-directed manual engagement in sustaining joint attention.

We employed two analytic approaches in this study. First, we conducted micro-level analyses of individual parent following and parent directing episodes, querying the predictive role of toddler manual engagement on time spent in joint attention in those episodes. To the extent that toddlers' manual engagement is important to achieving joint attention regardless of social context, then we should observe that toddler manual engagement predicts sustained joint attention in both contexts. In contrast, if sustained joint attention in parent directing follows a different path, we should expect a much more reduced predictive power of toddler manual engagement on joint attention in directing contexts. In addition to this episode-level analysis, we also conducted an individual-differences analysis, testing the extent to which toddler manual engagement could explain why different follow-versus-direct dynamics lead to differences in joint attention [5]. That is, we ask whether the predictive power of a parent's tendency to follow toddlers' attention on parent-toddler joint attention can be explained by the degree to which toddlers are manually engaged with the object of joint attention.

We defined joint attention objectively as moments when parents and toddlers simultaneously looked at the object which was the target of a particular play episode. Because we sought to understand the objective sensorimotor underpinnings of joint attention and their potential differences across different social contexts, we defined joint attention without consideration of possible inferences regarding the knowledge states of the participants. It should be noted that this definition stands in contrast to previous definitions of joint attention [6]. However, our objective definition of joint attention, informed by high-density gaze data, was chosen in order to best disentangle the component parts which make up joint attention across differing social contexts.

## III. METHODS

### A. Participants

Sixteen toddlers ( $M_{age} = 21.8$  months,  $SD = 1.3$ ) and their parents were observed while they played with a set of objects in the laboratory. 8 toddlers were girls. 11 parents were mothers.

### B. Experimental Environment & Stimuli

Figure 2a depicts the experimental setup: toddlers sat in a high-chair across a table from their parents who sat on floor cushions. Parent-toddler dyads played with two sets of three novel objects. All objects were constructed in the lab and had a single main color. The objects' sizes were comparable (approximately 270 cm<sup>3</sup>) and allowed for toddlers' grasping, picking up, and playing. All objects were thoroughly pilot-tested to be interesting and engaging to toddlers.

### C. Apparatus

During the play session, both parents and toddlers wore head-mounted eye-trackers (Positive Science, LLC; [12]; see Fig 2). These eye-trackers include two cameras: (1) an infrared eye-camera pointed at the right eye for recording eye movements and (2) a scene camera pointed outward for recording first-person perspectives. Cameras had a 100 degree field of view. Both cameras recorded at a temporal resolution of 30 Hz and a spatial resolution of 640x480. This dual head-mounted eye-tracking system has been successfully used in previous parent-toddler interaction studies (for more details on the dual head-mounted set-up, see reference [7]; for more details on eye tracker specifications, see [12]). Three additional video cameras (one bird's eye view camera and two wall-mounted cameras positioned behind the right shoulders of toddlers and parents) captured the play session from viewpoints independent of participants' movements.

### D. Procedure

Our goal was to observe a natural play session but one that consisted of both moments where parents followed their toddlers' attention and moments where parents directed their toddlers' attention. Thus, we first instructed parents to play with their toddlers and the objects as they might normally do at home. Because in the typical laboratory observation parents are more likely to follow than direct their toddlers' attention [5], before the study began we encouraged parents to occasionally bring their child's attention to new toys and to get their child to attend to each object at least once.

Once parents and their toddlers put on white smocks, the eye-trackers fitted and calibrated (for complete details of this process, see [7]), parents and their toddlers completed four brief trials of object play. Each trial lasted 1 to 1.5 minutes long ( $M = 1$  minute and 26.82 seconds;  $SD = 17.15$  seconds) and began with an experimenter putting one of two sets of three objects on the table. The object sets were swapped out after each trial.

### E. Coding

**Episode Coding.** From the video-recordings of these interactions, we first identified episodes of play with different objects, and classified each episode as either *parent following* or *parent directing*. We did this by parsing the continuous stream of play into individual episodes of play with a particular

target object. On average, these episodes of single object play accounted for 87% of play ( $M = .87$ ;  $SD = .56$ ). The remaining time of the experiment consisted of moments when parents and toddlers were engaged in non-object play (e.g., peek-a-boo), were off-task (e.g., looking around the experiment room), or played with two objects simultaneously. The onsets of episodes were coded as the moment in which toddlers or parents (whichever came first) initiated visual attention, manual contact, or vocalized about the target object. The offsets of episodes were coded as the moment in which toddlers and/or parents (whichever came last) terminated visual attention, manual contact, or vocalizations directed at the object.

Once play episodes were identified, we then determined whether the episode of play was initiated by toddlers’ attention or by parents’ bid for toddlers’ attention. Episodes were classified as “following episodes” if they were initiated by the toddler. Episodes were classified as “directing episodes” if they were initiated by a parent bid (usually a manual action, a gesture, or a vocalization) for toddlers’ attention. Table 1 reports the number of following and directing episodes we observed per dyad, the proportion of episodes classified as following vs. directing, and the mean duration of following and directing episodes.

(a)



(b)

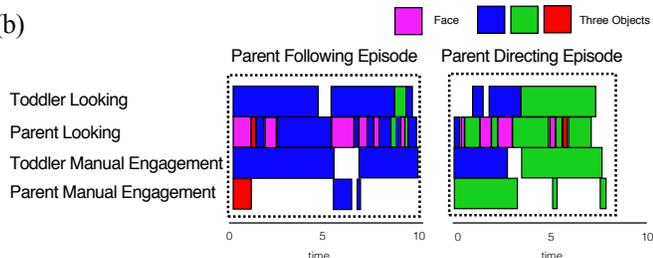


Figure 2. Toddlers and parents equipped with head-mounted eye trackers played with objects in the lab [7] (a). Example frames sampled from the Positive Science eye-trackers, which allowed for micro-level coding of toddlers and their parents gaze patterns (b). Categorical time-series plots illustrating gaze and manual engagement patterns across a representative parent-following and parent-directing episodes. Joint Attention was operationalized as moments when parents and toddlers were visually attending to the same target object [7].

**Sensorimotor Coding.** Because we were interested in the sensorimotor correlates of different types of episodes, we also utilized in-house software to code, frame-by-frame, the target of parent and toddlers’ gaze and the target of their manual actions. For determining the targets of gaze, trained coders utilized X- and Y-coordinates produced by the eye-tracker, as well as the raw footage of the eye and scene cameras (see Figure 2a, right two panels). For each frame, coders determined whether parents and toddlers gaze was on one of four mutually exclusive targets: the three objects in play or their partner’s face. For manual actions, trained coders watched frame-by-frame the play footage from multiple angles and determined for each frame whether and which objects were in contact by parents’ and toddlers’ hands. Figure 2b illustrates a time series of the coded sensorimotor data.

TABLE I. DESCRIPTIVE STATISTICS OF EPISODES TYPES

	Episode Types	
	<i>Following</i>	<i>Directing</i>
Avg. number of episodes per subject	9.12 (5.40)	18.81 (7.22)
Avg. prop. of episode type per subject	.33 (.17)	.66 (.17)
Avg. duration of episodes per subject	10.19s (3.41)	11.71s (2.81)

#### IV. RESULTS

Below we present two sets of results to reveal how joint attention is achieved in different social contexts. We first examined whether the amount of toddler manual engagement is associated with joint attention in both parent-directing as well as parent-following episodes. This analysis tests whether toddlers’ own manual actions are critical for sustained joint attention regardless of how joint attention was initiated (i.e., parent following or parent directing). We then asked whether toddler manual engagement may actually account for the differences in amount of joint attention commonly attributed to differences in how much parents follow vs. direct their toddlers’ attention. Through a mediation analysis, we examined whether the relation between parent following and amount of joint attention is mediated by the amount of toddler manual engagement.

*A. Toddler Manual Engagement is key to Joint Attention in Parent-Directing as well as Parent-Following.*

All parent-toddler dyads we observed produced some episodes where parents followed their toddlers' attention and some episodes where parents directed their toddlers' attention (see Table 1). Consistent with previous research and not surprisingly, following and directing episodes were characterized by different sensorimotor properties: parent following episodes consisted of both more joint visual attention (operationalized as parent and toddlers simultaneously looking at the target object) and more toddler object manual engagement (see Figure 3 and Table 2).

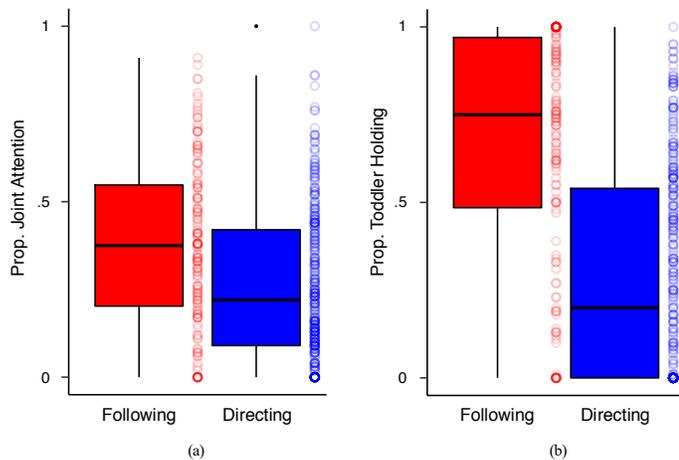


Figure 3. The proportion of frames both parent and toddler looked at (a), and toddler manually engaged with (b) the target object as a function of parent following the toddler's attention vs leading the toddler's attention.

TABLE II. COMPARISONS FROM RANDOM MIXED-EFFECTS MODELS

	Following	Directing	Statistical Test ^
Joint Attention	.37 (.23)	.26 (.21)	***
Toddler Manual Engagement	.66 (.33)	.28 (.30)	***

Note. ^ Statistical tests were mixed-models with sensorimotor property as the criterion variable (i.e., joint attention, toddler manual engagement), subjects as random effects, and social context (i.e., following vs directing) as fixed effects. \*\*\*parameter estimates for the effect of social context on sensorimotor property is significant at  $p < .001$ .

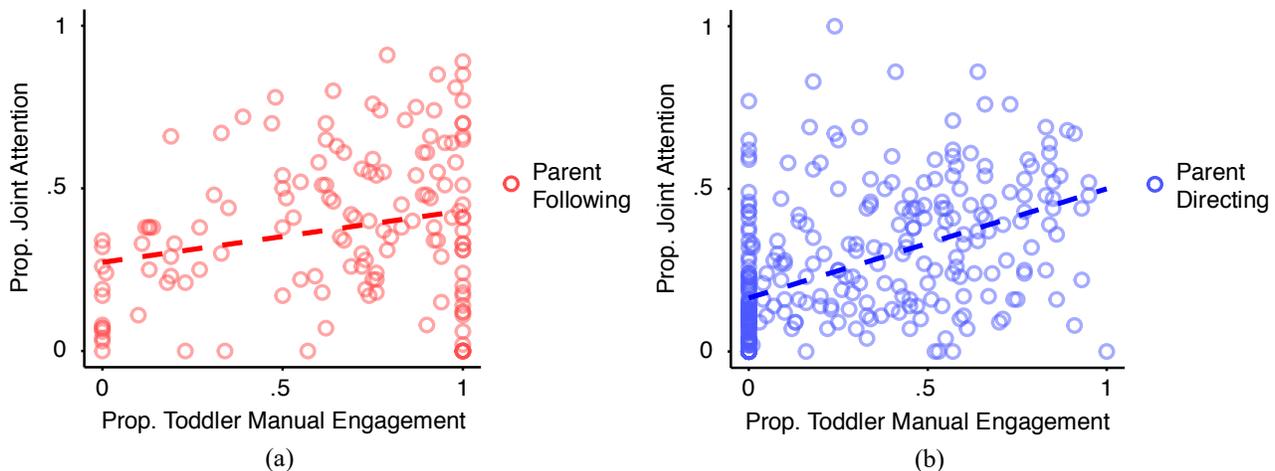


Figure 4. Relation between joint attention (y-axis) and toddler manual engagement (x-axis) in episodes of parents following (a) and directing (b) their toddler's attention.

Of particular interest was the association between toddler's manual engagement with objects and joint attention across both contexts. To explore this association, we analyzed the relation between toddler's manual engagement and joint attention during individual episodes of parent-following and individual parent directing episodes. Specifically, for both episode types, we computed the correlation between proportion of time toddlers manually engaged with objects and the proportion of time the toddlers and parents were in joint attention (see figure 4). Figure 4a illustrates the positive correlation between the amount of toddlers' manual engagement in a parent-following episode and the degree to which parent and toddler were in joint attention in that same episode. A mixed-model regression with subjects as random intercept effects, proportion of frames in joint attention during episodes of parent following and proportion of frames of toddler manual engagement with target objects during episodes of parent following as dependent variable showed that the relationship between these proportions was statistically significant ( $\chi^2(1) = 7.49, p < .01$ ). Illustrated in figure 4b is the finding that when parents direct toddler's attention to new objects, toddler's manual engagement with objects is positively associated with joint attention. A mixed-model regression with subjects as random intercept effects, proportion of frames in joint attention during episodes of parent directing and proportion of frames of toddler manual engagement with target objects during episodes of parent directing as dependent variable showed that the relationship between these proportions was statistically significant ( $\chi^2(1) = 72.99, p < .001$ ), increasing the proportion of joint attention by  $.31 \pm .03$  (standard errors) per standard deviation unit increase in toddler manual engagement with target objects. Thus, overall, toddler's manual engagement is positively associated with joint attention.

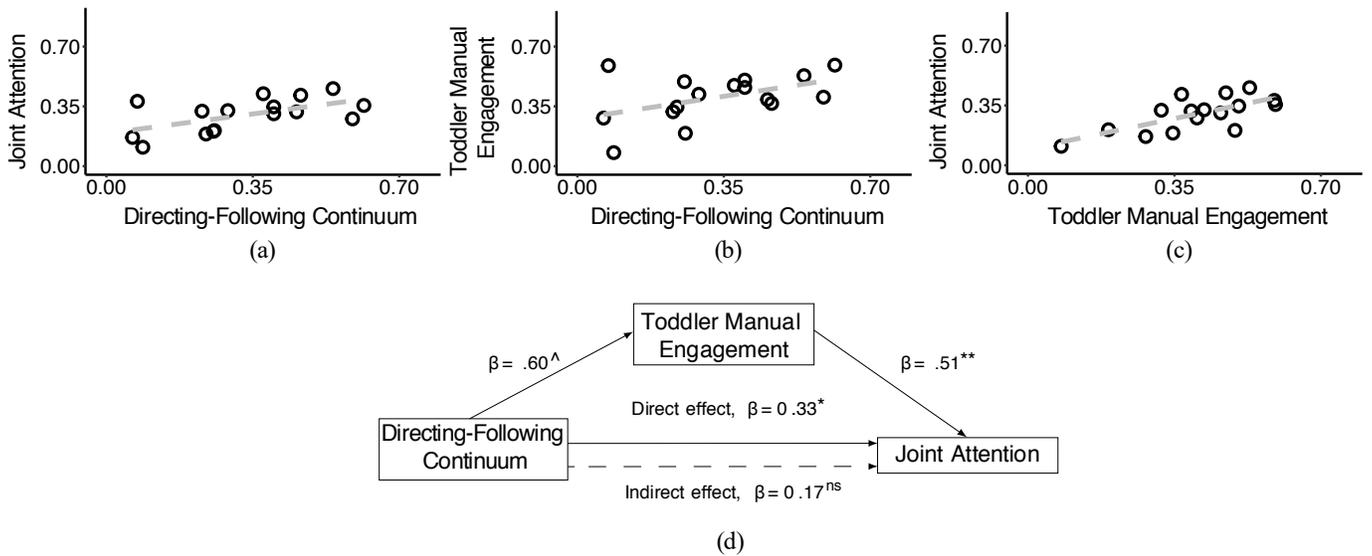


Figure 5. Relation between joint attention (y-axis) and proportion of episodes where parents followed toddler’s attention (a). Relation between toddler manual engagement (y-axis) and proportion of episodes where parents followed toddler’s attention (b). Relation between joint attention (y-axis) and toddler manual engagement (y-axis) (c). Mediation model predicting the mediation of the effect of parent following on joint attention through toddler manual engagement (d). Parent manual engagement was not positively related to joint attention in episodes of parent following or directing and thus was not included in the model.

*B. Toddler manual engagement accounts for the link between follow-direct dynamics and parent-toddler joint attention.*

Toddlers’ manual engagement thus appears to be a driver of joint attention in both social contexts. This result suggests an interesting possibility: the reason why individual differences in whether parents follow has previously been associated with individual differences in joint attention [13] may be accounted for entirely by differences in toddler manual engagement in parent-following vs. parent-directing contexts. To explore this possibility, we employed a mediation analysis [14] testing whether differences in toddler’s manual engagement during both parent following and parent directing contexts accounted for the differences in joint attention in those contexts.

The three scatter plots presented in Figure 5 show that the first three necessary conditions for mediation were met in these data: (1) individual differences in parents’ propensity to follow vs. direct episodes of play (i.e., the proportion of play episodes classified as “following” out of all play episodes) correlated with individual differences in joint attention (i.e., the mean proportion of joint attention between parents and toddlers during play episodes; Figure 5a); (2) individual differences in parents’ propensity to follow also correlated with individual differences in toddler manual engagement (i.e., the mean proportion of toddler manual engagement with the target object during play episodes; Figure 5b); and (3) individual differences in toddler manual engagement correlated with individual differences in time spent in joint attention (Figure 5c).

The final critical condition for mediation is that the path coefficient between the predictor variable (parent propensity to follow) and the outcome variable (joint attention) be significantly reduced when the mediator variable (toddler manual engagement) is introduced in a multivariate regression model. As shown in Figure 4d, the coefficient for the effect of parent’s propensity to follow on joint attention is reduced from .33 to .17 in a model where toddler manual engagement is

included as an additional predictor variable. This result suggests that the mediation was significant and that parent following of toddler’s attention promotes joint attention at least in part via its influence on toddler’s manual engagement with objects.

V. DISCUSSION

In the present study, we set out to examine the pathways to joint attention across different social contexts. Using dual head-mounted eye-tracking during dyadic object play [7], we measured moment-by-moment parent-toddler joint attention and how toddlers’ object-direct manual engagement correlated with joint attention in different social contexts. We found that across contexts, toddler manual engagement was key to sustained joint attention; the more toddlers were manually engaged with the target object, the more time parents and toddlers were in joint attention. Moreover, differences in the degree to which individual toddlers were manually engaged with objects explained the effect of parents’ tendency to follow their toddlers’ attentional lead on joint attention. These results speak to the sensorimotor pathways to joint attention, ground the research on parent following and directing in its sensorimotor elements, and highlight the importance of understanding how joint attention was achieved through toddlers’ bodily actions.

Researchers have previously demonstrated the importance of toddler manual engagement on parent-toddler joint attention [8,15]. The current study goes beyond that data by illustrating that this effect holds across multiple social contexts. That is, we observed that even in a context where toddlers are not initially manually engaged with the target object, toddlers’ manual engagement is the key factor that determines whether sustained joint attention happens. Why is toddler manual engagement so important? We hypothesize that there are likely multiple reasons for this. First, manual engagement may impact joint attention indirectly by first stabilizing and sustaining toddlers’ own visual attention [8,16]. Stable and sustained toddler attention in turn provides parents an easy target with which they

can coordinate their own attention (see [8,17,18]). Second, manual engagement may also impact joint attention more directly by simply providing a clear cue to toddler's focus of attention. That is, compared to toddlers' visual attention, manual engagement is a clearer and more stable signal to parents regarding possible opportunities for joint attention. Future research that investigates the dynamics of how toddler manual engagement and joint attention unfold in real time may be in a better position to reveal the precise mechanisms by which toddler manual engagement influences joint attention.

Although the positive benefit of parent "following-in" and parent responsivity on child development is unassailable [19], the negative impact of parent directing is not as settled a matter. For example, when researchers have carefully distinguished between "successful" and "unsuccessful" parent directing [20,21,2], researchers find that the frequency of successful directing is actually linked to positive outcomes (e.g., more rapid language development). The current study highlights that episodes where parents follow and episodes where parents successfully direct may have more in common than their labels pre-suppose. The sensorimotor correlates of these two kinds of episodes (a lot of toddler manual engagement and a lot of joint visual attention) are highly similar. Thus, although some have viewed the findings of parent responsivity and parent successful directing as being at odds with one another [20], we suggest that they may in fact be tightly related given their micro-level, sensorimotor underpinnings. Consistent with several other recent lines of research [22], the current study highlights that there is a lot to be gained theoretically from understanding well established macro-level constructs in more grounded micro-level terms.

In the developmental sciences, joint attention has historically been viewed as a root causal variable, a key capacity that impacts several facets of development in important ways [1,2,3]. Not surprisingly then, several research programs have sought to understand how differences in joint attention early in development predict differences in outcomes later in development [1-3,6,10], to identify the precise pathways through which joint attention shapes those outcomes [9,13,23], and to test whether interventions that focus on molding joint attention work for developmental cases gone awry [19]. A different view of joint attention however is that it is not only a developmental cause but also a development product of several inter-related pieces. Under this view, unpacking joint attention and identifying the key pieces that make up joint attention [9] may turn out to be a better strategy for shedding light on its developmental significance.

## ACKNOWLEDGMENTS

We thank Charlotte Wozniak, Melissa Elston, Seth Foster, and the IU Computational Cognition and Learning Laboratory for assistance with data collection, coding and management. We also thank all the families that participated in this study.

## REFERENCES

- [1] A. Bigelow, K. MacLean, and J. Proctor, "The role of joint attention in the development of infants' play with objects," *Dev. Sci.*, vol. 7.5, pp. 518–526, 2004.
- [2] M. Tomasello, and M. Farrar, "Joint attention and early language," *Child Dev.*, pp.1454–1463, 1986.
- [3] T. Charman, S. Baron-Cohen, J. Swettenham, G. Baird, A. Cox, and A. Drew, "Testing joint attention, imitation and play as infancy precursors to language and theory of mind," *Cog. Dev.*, vol. 15(4), pp. 481-498, 2000.
- [4] C. Yu, and L. Smith, "The social origins of sustained attention in one-year-old human infants," *Current Bio.*, vol. 26.9, pp 1235-1240, 2016.
- [5] S. Mendive, M. Bornstein, and C. Sebastián, "The role of maternal attention-directing strategies in 9-month-old infants attaining joint engagement," *Inf. Behav. And Dev.*, vol 36.1, pp. 115-123, 2013.
- [6] M. Tomasello, "Joint attention as social cognition," *Joint attention: Its origins and role in development.*, 103130, 1995.
- [7] C. Yu, and L. Smith, "Joint attention without gaze following: human infants and their parents coordinate visual attention to objects through eye-hand coordination," *PLoS one*, vol. 8.11, 2013.
- [8] C. Yu, and L. Smith, "Hand-eye coordination predicts joint attention," *Child Dev.*, vol 88, pp. 2060-2078. 2017.
- [9] C. Yu, and L. Smith, "Multiple sensory-motor pathways lead to coordinated visual attention," *Cog. Sci.*, vol 41, pp. 5-31, 2017.
- [10] M. Tomasello, and J. Todd, *Joint attention and lexical acquisition style.* *First Lang.*, vol 4(12), pp. 197-211, 1983.
- [11] P. Mundy, and L. Newell. "Attention, joint attention, and social cognition." *Cur. Dirs. in Psych. Sci.*, vol 16.5, pp. 269-274, 2007.
- [12] J. Franchak, K. Kretch, K. Soska, and K. Adolph, "Head-mounted eye tracking: A new method to describe infant looking," *Child Dev.*, vol 82(8), pp. 1738-1750, 2011.
- [13] G. Mason, F. Kirkpatrick, J. Schwade, and M. Goldstein. "The Role of Dyadic Coordination in Organizing Visual Attention in 5-Month-Old Infants." *Infancy*, 2018.
- [14] R. Baron, and D. Kenny, "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations." *J. of Pers. and Soc. Psych.*, 51(6), p. 1173, 1986.
- [15] N. Rader, and P. Zukow-Goldring. "How the hands control attention during early word learning." *Gesture*, vol 10, pp. 202-221, 2010.
- [16] L. Yuan, L. Xu, C. Yu, and L. Smith, "Sustained visual attention is more than seeing," *J. of Exp. Child Psych.*, 179, pp. 324-336, 2019.
- [17] S. Wass, and V. Leong. "Developmental Psychology: how social context influences infants' attention," *Cur. Bio.*, vol 26, pp. R357-R359, 2016.
- [18] C. Suarez-Rivera, L. Smith, and C. Yu. "Multimodal parent behaviors within joint attention support sustained attention in infants." *Dev. Psych.*, vol 55, pp. 96, 2019.
- [19] C., Tamis-LeMonda, Y. Kuchirko, and L. Song, "Why is infant language learning facilitated by parental responsiveness?" *Cur. Dirs. in Psych. Sci.*, vol 23(2), pp.121-126, 2014.
- [20] M. Shimpf, and J. Huttenlocher, "Redirective labels and early vocabulary development," *J. of Child Lang.*, 34(4), pp. 845-859, 2007.
- [21] K. Scott, E. Sakkalou, K. Ellis-Davies, E. Hilbrink, U. Hahn, and M. Gattis, "Infant contributions to joint attention predict vocabulary development," 35<sup>th</sup> *Cog. Sci. Conf.*, pp. 3384-3389, 2013.
- [22] I. Nomikou, M. Schilling, V. Heller, and K. Rohlfing, "Language-at all times," *Interaction Studies*, vol 17(1), pp.120-145, 2016.
- [23] G. Deak, A. Krasno, J. Triesch, J. Lewis, and L. Sepeta, "Watch the hands: infants can learn to follow gaze by seeing adults manipulate objects." *Developmental Science*, vol. 16 (2), pp. 270 – 281, 2014.