Child Development

CHILD DEVELOPMENT

EMPIRICAL ARTICLE

The Dynamics of Caregiver Unpredictability Shape Moment-To-Moment Infant Looking During Dyadic Interaction

Tess Allegra Forest[1] | Layla Bradford[2] | Lorna Ginnell[3,4] | Maroussia Berger[1] | Donna Herr[2] | Emmie Mbale[3,5] | Kavindya Dalawella[1] | Chloë A. Jacobs[2] | Chikondi Mchazime[3] | Celia D'Amato[1] | Zamazimba Madi[2] | Pious Clifford Mkaka[3] | Claudia Espinoza-Heredia[1] | Tembeka Mhlakwaphalwa[2] | Vukiwe Ngoma[3] | Monique Gilmore[1] | Marlie Miles[2] | Jinge Ren[1] | Nwabisa Mlandu[2] | Reese Samuels[2] | Michal R. Zieff[2] | Melissa Gladstone[6] | Kirsten A. Donald[2] | Dima Amso[1]

[1]Department of Psychology, Columbia University, New York, New York, USA | [2]University of Cape Town, Cape Town, South Africa | [3]Kamuzu University of Health Sciences, Blantyre, Malawi | [4]University of Glasgow, Glasgow, UK | [5]ALMA DELTAS Fellowship Program, Blantyre, Malawi | [6]University of Liverpool, Liverpool, UK

Correspondence: Tess Allegra Forest (taforest@gmail.com)

Received: 12 January 2025 | **Revised:** 1 May 2025 | **Accepted:** 26 June 2025

Funding: This work was supported by Wellcome Trust, Wellcome LEAP K1D Program.

Keywords: attention | child development | early environment | infancy | learning | non-western populations

ABSTRACT

Cognitive development is associated with how predictable caregivers are, but the mechanisms driving this are unclear. One possibility is caregiver predictability initially shapes how infants gather information for learning. Here, caregiver-infant dyads ($N = 222$, 2-6-months-old, all female caregivers; data collected 2022–2023) in South Africa and Malawi engaged in naturalistic play before their interactions were hand-annotated to measure caregiver predictability and infant gaze. In both countries, temporal variation in caregiver predictability shaped infant looking dynamics—infants attended to specific sensory signals ($\eta^2 = 0.04$) and specific timepoints ($\eta^2 = 0.10$) that were useful for them based on their caregiver's typical behavior. These findings provide a framework by which the predictability of caregiver behavior may shape fundamental, early optimization of visual attention for learning in infancy and later cognitive development.

1 | Introduction

For infants, nearly every event is a chance to learn something new. Thus, characterizing how infants direct their attention amidst complexity has long been recognized as key to understanding their impressive learning feats (e.g., Gibson 1979). A rich literature shows that looking behavior is tightly coupled with information content in infancy, such that infants look at information that is useful for their learning across time (e.g., Poli et al. 2020). But what environmental cues might teach infants to efficiently allocate these resources? Caregivers, arguably the most prominent sources of regular environmental input in early life, are likely to be a crucial piece of the puzzle when it comes to understanding how infants learn to learn—but in what ways? In addition to providing infants with social and emotional support, caregivers provide either predictable or unpredictable sequences of behavior during daily dyadic interactions, with downstream consequences on later cognition (Davis et al. 2017). Here, we bridged the literatures on visual information seeking in infancy and caregiver predictability to test the idea that temporal

Kirsten A. Donald and Dima Amso contributed equally to this work.

variability in how predictable a caregiver is shapes the moment-to-moment dynamics of infant looking behavior.

Data have shown that caregiver predictability during everyday dyadic interactions influences cognitive development across species (Davis et al. 2017; Davis and Glynn 2024)—the more predictably mothers of 1-year-old infants transition between different behaviors (transitional probability of going from touching her baby to vocalizing, for example), the better children's general cognition, memory (Davis et al. 2017), and effortful control years later (Holmberg et al. 2022). Other data have shown that the predictability of caregiver behavior, subsequent to instances of vocalization, is associated with neural indices of statistical learning months later (Forest et al. 2024), and that the predictability of caregivers' emotional expressions (Vanoncini et al. 2022) and social gaze are related to performance on early language learning tasks (Vanoncini et al. 2024). But the link explaining how caregiver predictability becomes internalized in infants in a way that might reasonably impact later outcomes is still unclear. We hypothesized that predictable caregiver behavior may go on to shape cognitive development *by scaffolding how infants gather information*, with later consequences for more general cognition.

This suggestion is highly consistent with the fact that naturalistic caregiver-child interactions have a number of properties that support learning by directing infants' attention. For example, adults speak to young learners using Infant Directed Speech (IDS), which facilitates learning by capturing attention (Nencheva et al. 2021). Caregivers (Franchak et al. 2018) and infants (Yu and Smith 2016) move their bodies during naturalistic interaction in ways that facilitate joint and sustained attention (Suarez-Rivera et al. 2019; Yu and Smith 2016). Importantly, infants and caregivers *work together* to structure these interactions (Karmazyn-Raz and Smith 2023; Sameroff 2009), implying naturalistic behavior relies on prior experience between infants and their caregivers to support the infant's learning. These examples (and likely many others) demonstrate that the regular dyadic interactions infants have with their caregivers house a variety of properties that shape the temporal dynamics of looking behavior in early infancy. Still, whether a caregiver's *predictability* during these interactions is another such example has yet to be seen.

Why would changes in caregiver predictability shape infant looking? A parallel literature shows that early attention and learning rely on infants' ability to track predictable information in the environment (see Aslin 2017). *Entropy* (Shannon 1948), a widely-used measure of (un)predictability, has long characterized learning theory and human behavior (Miller 1956; Rieke et al. 1997). As higher entropy information sources are less predictable, they contain more to learn. Importantly, entropy predicts attention (Berlyne et al. 1971; Jensen et al. 2013), including in infancy: infants are most likely to look away from a stimulus when it is highly predictable or highly unpredictable (Kidd et al. 2012), suggesting medium entropy information is preferentially attended. The best trial-by-trial predictor of where infants attend in an experiment is how much they can learn from each location based on what they have already seen (Poli et al. 2020). Adults also attend to higher entropy sources of information as they gain experience in an environment, highlighting that the 'optimal' level of predictability is relative to one's past experience (Forest et al. 2022). Thus, learners prioritize looking towards medium predictability information in service of learning in lab-based contexts, and what counts as 'medium' is relative to someone's past experience. This suggests infants could use their prior experience with their own caregiver as a signal for when she is providing information that is neither too simple nor too complex *for them*, but this remains unknown.

There is one additional literature that converges with the idea that individual differences in infant looking behavior may be related to caregiver predictability. Individual differences in how infants allocate attention are also linked to later cognition (see Colombo et al. 2010, for review), in a parallel manner to early caregiver predictability. Briefly, individual differences in infants' looking behavior during habituation and paired-comparison tasks have been shown to predict a host of cognitive outcomes in childhood (Colombo 1993; McCall and Carriger 1993; Rose and Feldman 1995). For instance, infants who look at a stimulus for less time have better information processing capabilities than those who look for longer. In this way, both caregiver predictability and early visual attention allocation strategies predict later cognitive development. While this, of course, does not mean looking behavior and caregiver predictability must be related, probing whether there is a link between these two variables may allow us to connect these literatures more directly and better understand the unique predictive value they each hold for developmental outcomes.

Thus, we asked whether infants flexibly adjust their looking behavior as a function of their *own* caregiver's predictability, as a critical step in understanding how infants' early experiences may shape individual variability in allocating attention for learning and cognitive development more broadly on longer timescales. To do this, we simultaneously measured moment-to-moment caregiver predictability and infant looking during naturalistic, dyadic interactions. Importantly, we calculated caregiver predictability using basic caregiver actions that infants' sensory systems can process (Davis et al. 2017; touch, visual cues, auditory information) to ensure our approach was culturally agnostic.

If the predictability of infants' early caregiving experiences shapes their moment-to-moment visual behavior, an open question with implications for intervention is whether caregiver predictability varies as a function of the caregiver's environment. Most work in humans has studied environmental changes in caregiver predictability on longer timescales (e.g., how many job changes or instances of rehoming a family experiences, see Davis and Glynn 2024, for recent review). This means that our understanding of how more immediate contextual variability shapes moment-to-moment predictability is still quite limited. We know from rodent literature that limiting the availability of bedding resources increases the likelihood that rat mothers will act unpredictably (Demaestri et al. 2022). While this work suggests caregiver predictability may be shaped partly by local contexts, such manipulations are meant to understand how (long-term) stressful situations might shape caregiver predictability. We were interested in whether normative variability in a caregiver's environment would shape caregiver predictability, and thus infant looking behavior, in two ways. First, we characterized the moment-to-moment predictability of caregivers from two African countries with distinct cultures and environmental risks (Zieff et al. 2024). This allowed us to answer a basic open question of whether there are overarching differences in caregiver predictability which might impact infants' looking behavior across two cultures. Then, for a more granular analysis, we asked whether simple changes to environmental affordances (Gibson 1979) impact how predictable caregivers are. This manipulation might be more akin to understanding how subtle variations in context, such as where in a home dyads are interacting, or what daily routines they are completing, might modulate differences in caregiver predictability. While there are many ways to operationalize such changes in environmental affordances, we opted to record interactions between infants and their caregivers with and without toys present, as a relatively simple, experimentally tractable manipulation that would allow us to begin to understand potential changes in dyadic behaviors in humans. We hypothesized that if the dynamics of caregiver predictability shift in a context-dependent manner (with or without a toy), and infants are sensitive to this information, infant looking behavior would vary in relation to the caregiver behavior which provided the most information (i.e., the highest entropy) in each context. The alternative is that infant looking time is only related to a caregiver's average predictability, or is not related to a caregiver's predictability at all.

Finally, we modeled even finer-grained temporal coupling between caregiver predictability and infant looking by modeling how predictable each caregiver was over time during the naturalistic interactions and measuring how much her infant looked at her at each timepoint. We hypothesized that infants would attend most to their own caregiver when she was demonstrating a medium level of predictability *relative to her own average*. This would suggest that infant information gathering behavior develops to be exquisitely tuned to moments in time that are useful for them based on their own unique past experiences and start to link caregiver predictability to early information gathering strategies.

2 | Methods

2.1 | Participants

2.1.1 | Cohort Study Description

The reported data were collected as part of a larger longitudinal cohort study designed to characterize the development of executive function from 0 to 1000 days. The two study populations were recruited through clinics in Cape Town, South Africa ($N = 393$) and Blantyre, Malawi ($N = 415$). All protocols were approved by the local university Health Research Ethics Committees. All caregivers were female, and all but one was a mother. Caregivers signed informed consents on behalf of themselves and their infants. All consent forms and procedures were communicated in the preferred family language (Xhosa or English in South Africa, and Chewa in Malawi). As part of the longitudinal study, families participated in tasks including medical assessments, MRI, EEG, and behavioral questionnaires (Zieff et al. 2024). The particular assessments at each timepoint varied slightly. Here, we report caregiver-infant interaction data collected at families' first postnatal visit, which were collected between the spring of 2022 and the winter of 2023.

2.1.2 | Participant Subsample and Power

Of all dyads enrolled, 540 completed the caregiver-infant interaction session during their first postnatal visit, when infants were between 2 and 6 months old. Of these, 226 randomly selected dyads were hand-annotated for these relatively exploratory analyses. Four of these dyads were then excluded because infants were wearing MRI ear protection during the caregiver-infant interaction, or because the caregiver drew attention to toys during the No Toy condition. This resulted in a final sample size of 222 ($N = 116$ in South Africa, $N = 106$ in Malawi, see Table 1 for additional demographic information), which is larger than the minimum ($N = 188$) needed to achieve 95% power when detecting a small effect ($F = 0.1$) on an interaction term in a 5×2 ANOVA (see 'Analysis Plan'), as determined by a G*Power analysis (Faul et al. 2007).

2.2 | Procedure

2.2.1 | Caregiver-Infant Interaction

The main goal of this assessment was to observe infants and caregivers interacting naturally. All dyads were shown to a private space where caregivers sat facing infants, and were instructed to play as they would at home for ~10 min. These 10 min were divided into two, back-to-back, randomly-ordered segments: one with a selection of culturally appropriate toys for caregivers to choose from ("Toy" condition) and another without the toys available ("No Toy" condition). Each condition took place in the same space, the only difference was the presence or absence of toys; (see Supporting Information S1, "Session Script and Protocol", for exact instructions which were piloted and culturally adapted for each site and each condition order). Due to slight variability in the duration of the play sessions, recording time

TABLE 1 | Demographic information for each site.

	Malawi	South Africa
Total *N*	106 dyads	116 dyads
Infant age (in days)	$M = 99.82$, SD = 25.80	$M = 110.67$, SD = 28.22
Infant sex	31 Male, 36 Female, 39 missing	54 Male, 57 Female, 6 missing
Caregiver entropy	$M = 0.82$, SD = 0.34	$M = 1.0$, SD = 0.32

ranged from 1.42 min to 7.34 min ($M = 5.46$ min, $SD = 0.45$ min), and was included as a covariate in relevant analyses.

2.2.1.1 | Recording Apparatus. In both South Africa and Malawi, the space in which caregiver-infant interactions took place held three Logitech C920 Pro HD Webcam cameras, connected to a computer which temporally aligned audio and video input from all cameras (https://manycam.com/). One camera directly faced the infant, one faced the caregiver, and one captured a side view of the dyad.

2.3 | Hand-Annotation Protocol

Caregiver behavior was hand-annotated offline using DataVyu (Datavyu Team 2014), after videos were assessed for resolution and audio quality. All persons who hand-annotated the data ('coders') were rigorously trained and tested against other trained coders until their annotations on each behavior across all timepoints in eight videos matched an experienced coder's annotations with > 85% accuracy. Separately, we ensured high overall inter-rater reliability by implementing a quality control procedure on a random 10% of all videos. These videos were re-coded by a second coder to confirm interrater reliability was at least 80%. If any quality check resulted in less than 80% reliability systematically for any behavior for any coder, that behavior was re-coded in all videos that had been annotated by that coder. The final inter-rater reliability for these 10% of videos resulted in a Cohen's Kappa of 0.75, which reflects high inter-rater reliability (McHugh 2012).

To understand the predictability of caregiver behavior, we used a protocol developed by Davis and colleagues (Davis et al. 2017) to annotate five caregiver behaviors which could be processed by infants' sensory systems, and infant looking time towards the caregiver. This annotation was done separately for each condition, by marking the temporal onset and offset of the following five orthogonal caregiver behaviors (see Figure 1a and Table 2): vocalizing (auditory signal), holding the baby (i.e., supporting the entirety of the baby's weight, for example while breastfeeding or standing with the baby, tactile signal) or touching the baby (i.e., tapping or kissing the baby, separate tactile signal), holding an object or pointing to an object (visual signals, note that this object could be any object in the room, including the toy in the Toy condition), and one infant behavior (looking towards the caregiver). Note that while we included both "Holding the baby" and "Touching the baby" to align with the extant coding scheme (Davis et al. 2017) and support future comparisons between samples, the infants in our sample were almost always seated in a car seat. This meant that there were very few instances of caregivers holding their infants. Nonetheless, these two signals are still orthogonal—for instance a caregiver might have lifted her baby out of the seat and breastfed during the session ("Holding the Baby") and then given the baby a kiss while doing so ("Touching the Baby"). As in past work (Davis et al. 2017), instances of a behavior separated by less than 500 milliseconds (holding baby, touching baby, holding an object, or pointing) or 1000 milliseconds (vocalizing, infant eye gaze) were identified as a single event, and behaviors occurring further apart in time were annotated as separate occurrences. Therefore, every instance of each behavior was coded for both its occurrence and duration throughout the session, on a frame-by-frame (33 millisecond) basis. Each behavior was coded during a separate pass of the video (e.g., caregiver vocalizations were coded during one pass, and infant eye-gaze was coded in a completely separate pass), to avoid mistakes and to ensure all codes were independent. This approach ensures the relations between behaviors are based solely on true co-occurrence, not on coder interpretation of co-occurrence. Instances where caregiver and, or, infant behaviors were unclear were annotated as "unknown" and excluded from analysis (No Toy condition, caregiver behavior was not codable for $M = 1.18\%$ of frames, infant behavior $M = 1.67\%$ of frames; Toy condition, caregiver behavior $M = 0.92\%$ of frames, infant behavior $M < 0.01\%$ of frames).

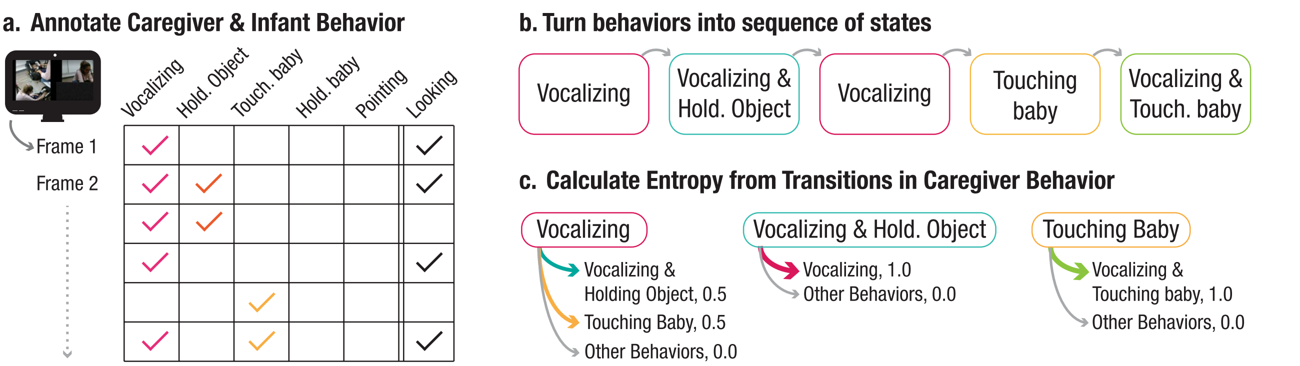

FIGURE 1 | Measuring caregiver predictability. (a) Caregiver-infant interactions were recorded from three cameras which were aligned temporally before hand-annotating five caregiver behaviors (Vocalizing, Holding object, Touching baby, Holding baby, Pointing) and one infant behavior (Looking towards caregiver) on each frame of the video. On each frame (y-axis of table), the presence or absence of each behavior (x-axis of table) was marked (colored check marks in each box), resulting in a frame-by-frame annotation of caregiver behavior. (b) These annotation matrices were then transformed into a sequence of states. Behavioral states could be combinations of behaviors or single behaviors. Sequences were generated independently for each caregiver. (b) shows an example transformation of the frame-by-frame annotations in (a) to a sequence of states. (c) The transitional probability from one state to another was then calculated to compute the entropy of caregiver behavior. In this example, the caregiver transitioned from Vocalizing (pink outline, first column), to Vocalizing and Holding Object once (50% of the time), and to Touching her baby once (50% of the time). Thickness of arrows reflects probability of the transition; color of arrows indicates the behavioral state the caregiver transitioned to. These transitional probabilities were then fed into the standard calculation for Shannon Entropy, using all possible behaviors (e.g., all transitions depicted in (c)), or just based on the transitions following particular behavioral states (e.g., just following 'Vocalizing' in (c)).

TABLE 2 | Prevalence of specific caregiver behaviors across conditions.

Behavior	Proportion of caregivers exhibiting behavior		Mean number of times behavior occurred in caregivers who exhibited behavior		Mean number of times behavior occurred in all caregivers	
	No toy	Toy	No toy	Toy	No toy	Toy
Single behaviors						
Vocalizing	1.0	1.0	117.82	75.55	117.82	75.55
Touching baby	0.99	0.92	93.82	42.94	92.55	39.28
Holding baby	0.30	0.11	32.6	19.56	7.37	2.20
Holding object	0.61	1.0	37.84	138.81	23.18	138.98
Pointing	0.09	0.05	2.67	2.82	1.06	0.14
Two- and three-way combinations of frequent behaviors						
Vocalizing and touching baby	0.97	0.37	47.20	3.30	52.11	1.22
Vocalizing and holding object	0.54	1.0	15.44	49.86	8.27	50.73
Touching baby and holding object	0.39	0.86	9.78	19.89	3.69	16.99
Vocalizing and touching baby and holding object	0.47	0.90	11.98	22.11	5.42	19.75

2.4 | Data Preprocessing and Analysis Plan

Our coding process resulted in a data frame which describes the presence or absence of the five caregiver behaviors and one infant behavior for each time point during a session (Figure 1a). To characterize the predictability of caregiver behavior, we transformed this data frame into a series of caregiver behavioral 'states,' which reflected the particular caregiver behavior(s) that were active at any time (Figure 1b). Operationally, a new behavioral state was generated each time any of the five behaviors started. As such, each new state's onset was the previous state's offset, and so every time point in a session was labeled with one caregiver state. For example, a caregiver might have been vocalizing while touching her baby and then switch to vocalizing while holding an object, which would be two separate states (i.e., the first three rows of Figure 1a and first two states of Figure 1b). Note that 'No behavior' is also a possible state, although this seldomly occurred. We then used this sequence of states to model caregiver behavior in each condition in two ways–to describe the frequency of different behaviors, and to describe their predictive structure. All data preprocessing was done in Python (version 3.9) via Jupyter Notebook (Kluyver et al. 2016).

2.4.1 | Frequency of Caregiver Behavioral States

First, we calculated the frequency of caregiver behaviors and behavior combinations by counting the number of occurrences of each state. Counts for single behaviors included occurrences that were part of combination states (e.g., counts for "Vocalizing" also included any instances of "Vocalizing and touching baby"). In contrast, counts for combinations of behaviors were restricted to instances of those exact combinations. This approach allowed us to understand how frequently behaviors occurred at all (single behavior states) and how frequently they occurred in tandem with each other. Two individual behaviors (Holding baby and Pointing) occurred in fewer than 12% of the caregivers in at least one of the conditions. Given this low prevalence, we did not analyze the frequency of combination behaviors involving these states, resulting in five single behavior states and four combination states included (see Table 2).

2.4.2 | Modeling the Predictability of Caregiver Behavior

The second way we characterized caregiver behavior was to model caregiver predictability, which we did both across all behaviors and for each behavior separately.

2.4.2.1 | Overall Caregiver Entropy. First, we examined the how predictable a caregiver was by modeling entropy across all her behaviors, in each condition separately. To do this, the sequence of states representing each caregiver's behavior was used to define a Markov model. These models allowed us to calculate how reliably a caregiver transitioned from one state to another, using the standard formula for Shannon Conditional Entropy (or, $H(Y|X) = \sum_{x \in X, y \in Y} p(x,y) log \frac{p(x,y)}{p(x)}$). This formula estimates the degree of unpredictability in transitioning from one state to another, given all states' frequency and between-state transitions (i.e., based on the transitional probabilities between states, as visualized in Figure 1c). The resulting output ('conditional entropy') reflects how unpredictable a caregiver's next action will be given her current state, across all of her behaviors (including Holding Baby and Pointing, despite their low prevalence). We refer to this metric as "Overall Entropy," to indicate that it was calculated across all behaviors (also see Table 1).

2.4.2.2 | (Un)predictability of Caregiver Transitions in Each Behavioral State. Conditional entropy can also be calculated *per behavior*, to reflect how predictable a caregiver is following any particular behavior. Mathematically, this involves using the same formula described above, but restricting the calculation to only include transitions that follow one particular behavior (e.g., following Vocalizing, Figure 1c). We used this approach to calculate how unpredictable a caregiver was after she displayed *each* behavior or behavior combination, to operationalize her behavior in a particular sensory domain in each condition (i.e., with and without toys present). We refer to these metrics as "Unpredictability" henceforth, to distinguish them from the overall entropy calculation above. As for frequency, we did not analyze caregiver unpredictability following Holding Baby nor Pointing (or behavior combinations including these states), but these states were included in the calculation of unpredictability following other behaviors.

2.4.3 | Analysis Plan

Our primary goals were to understand whether caregiver behavior differed as a function of environmental affordances, and how caregiver predictability impacted infant looking across time, in two African countries. All analyses were completed in R (R Core Team 2021). Linear Mixed Effects models were run using the lme4 package (Bates et al. 2015) and results are reported using the *Anova()* function from the Companion to Applied Regression package (Weisberg 2019). Effect sizes for the results of generalized linear models and linear mixed effects models were estimated using the *F_to_eta2()* function from the EffectSize package (Ben-Shachar et al. 2020).

2.4.3.1 | Analysis of Caregiver Behavior. First, we assessed whether frequencies of caregiver behaviors varied as a function of environmental affordances (i.e., the presence of toys). We compared the frequency of individual behaviors across conditions using a 5×2 ANCOVA with condition (Toy, No Toy) and behavior (Vocalizing, Holding object, Touching baby, Holding baby, and Pointing) as within-subjects factors and controlling for condition duration, before following up with paired two-sided *t*-tests where relevant. Then, we compared the frequency of combinations of behaviors using a 4×2 ANCOVA with condition (Toy, No Toy) and behavior combination (Vocalizing while holding object, Vocalizing while touching baby, Touching baby while holding object, and Vocalizing while holding object while touching baby) as within-subjects factors, while controlling for condition duration. As not every caregiver completed every behavior (some caregivers never touched their infant, for example), we also ensured that any effects we observed were still present in caregivers who performed those behaviors in both conditions. There were no major differences in the results, so these findings are presented in the Supporting Information S1.

Next, we examined whether overall entropy, as well as the unpredictability of particular behaviors, differed as a function of the country that caregivers lived in (South Africa or Malawi), or as a function of environmental affordances (Toy or No Toy condition). Entropy values for these metrics, as in other studies, range from 0 (i.e., perfectly predictable) to a theoretically infinite maximum based on how many behavioral states a caregiver demonstrated—thus, they are directly comparable across conditions in our study, and also to similar preexisting or future studies. Here, we first compared caregivers' overall entropy to a randomly permuted null distribution to understand if caregivers structure their behavior more than would be expected by chance in either condition. Then we directly compared the overall entropy of caregivers from Malawi to those from South Africa, and we directly compared overall entropy across the Toy and No Toy conditions (collapsed across countries). For specific behaviors, we could not calculate the unpredictability of each behavior for caregivers who did not complete that behavior. Thus, we compared the unpredictability of specific behaviors across conditions directly in caregivers who completed the behavior in both conditions using *t*-tests, and Bonferroni corrected for multiple comparisons across these seven two-sided *t*-tests; effects thus needed to meet or exceed a threshold of $p=0.007$ to be considered statistically significant.

2.4.3.2 | Analysis of Infant Behavior. In order to assess whether infant looking towards the caregiver varied as a function of caregiver overall entropy or specific behavior unpredictability across conditions, we examined the relation between caregivers' overall entropy and infant looking *across dyads* using a linear model that predicted infant looking as a function of caregiver entropy, condition, and their interaction, while controlling for the total time in each condition. We compared this model to one which also included a quadratic term for entropy and its interactions with condition, in order to understand any potential nonlinear relations between entropy and infant looking. However, the quadratic model was not a significantly better fit for the data (generalized linear model comparison: $F_{(2,437)}=2.52$, $p=0.08$), and so we report the results of the more parsimonious linear model. To ask if infant looking time was best predicted by different sensory behaviors across environmental affordance conditions, we ran two equivalent models which replaced caregiver entropy with the unpredictability of caregiver vocalization and the unpredictability of caregiver object holds, respectively. For these models, too, the parallel quadratic model did not fit better than the linear model (vocalizations, $F_{(2,324)}=1.28$, $p=0.28$; object holds, $F_{(2,312)}=0.66$, $p=0.52$), so we again report the results of the linear models.

We then tested the idea that *within* infant-caregiver dyads, infant looking time would be tightly coupled to variation in their caregiver's entropy. To do this, we binned the dyadic interaction data into 20 s intervals within each condition and computed the overall entropy of caregiver behavior within that interval. Thus, for each caregiver, we generated up to 21 separate entropy values, which reflect how unpredictable her behavior was during different points in time (for timepoints in which caregivers did not transition from one behavior to another, we could not compute this value, and thus some caregivers have fewer timepoints). We then mean-centered these entropy values within caregiver to result in a score for each time interval that reflected how unpredictable her behavior was *relative to her own average*. In other words, a low score reflects moments during which a caregiver was more predictable than her normal (lower entropy), while a high score reflects time intervals during which a caregiver was less predictable than her normal (high entropy). Importantly, this means that the within-caregiver entropy values used for this analysis were standardized relative to each caregiver and are thus not directly comparable to past work. We also calculated how much the infant looked at their caregiver during each time interval. Then, we modeled infant looking using a linear mixed effects model (Bates et al. 2015; Kuznetsova et al. 2017) which included within-caregiver entropy, condition, country, and total condition duration as fixed effects, and random by-subject effects for within-caregiver entropy and condition. Like above, we compared the fit of this model to one which also included a quadratic term for within-caregiver entropy and its interaction with condition and country. In this case, the data were better fit by a quadratic model (linear mixed effects model comparison: $\chi^2_7 = 33.45$, $p < 0.001$), and so we report those results below. We did not repeat this within-caregiver entropy calculation for the unpredictability of individual behaviors since there were many caregivers who did not exhibit the behavior of interest in enough time intervals to model those data in relation to infant looking.

3 | Results

3.1 | Frequencies of Caregiver Behaviors Vary Across Toy Conditions

We first asked if the frequency of caregiver behaviors differed as a function of the affordances available to caregivers during naturalistic play across each condition (Toy vs. No Toy). A 5×2 ANCOVA comparing the frequency of each caregiver behavior across conditions, while accounting for the duration of the condition, indicated there was a main effect of caregiver behavior ($F_{(4,884)} = 666.29$, $p < 0.001$, $\eta^2 = 0.56$) such that some behaviors were performed a great deal more than others (Table 2 and Figure 2a). There was no main effect of condition ($F_{(4,221)} < 0.001$, $p = 1.00$, $\eta^2 < 0.001$), but there was an interaction between condition and caregiver behavior ($F_{(4,884)} = 504.91$, $p < 0.001$, $\eta^2 = 0.42$). Pairwise two-sided t-tests indicated that caregivers vocalized and touched their baby more in the No Toy condition (Vocalizing, $M_{Toy} = 75.55$, $M_{NoToy} = 117.82$, $t_{(221)} = -10.54$, $p < 0.001$, Cohen's $d = -0.91$; Touching baby, $M_{Toy} = 39.28$, $M_{NoToy} = 92.55$, $t_{(221)} = -14.09$, $p < 0.001$, Cohen's $d = -1.09$), but held an object more in the Toy condition ($M_{Toy} = 138.98$, $M_{NoToy} = 23.18$, $t_{(221)} = 26.83$, $p < 0.001$, Cohen's $d = 2.36$; Figure 3a). Although caregivers held their baby statistically more in the No Toy condition ($M_{Toy} = 2.20$, $M_{NoToy} = 7.37$, $t_{(221)} = -3.58$, $p < 0.001$, Cohen's $d = -0.29$), this effect size was small and should be interpreted with caution. There was no difference in caregivers' pointing between conditions ($M_{Toy} = 0.14$, $M_{NoToy} = 1.06$, $t_{(221)} = -1.29$, $p = 0.20$, Cohen's $d = -0.12$). As noted above, given the low prevalence of these two behaviors, we did not consider them further.

We next asked if the frequency of any *combinations* of common caregiver behaviors differed between conditions. A 4×2 ANCOVA comparing the frequency of caregiver behavior combinations across Toy and No Toy conditions (see Table 2 and Figure 2b), while controlling for condition duration, indicated there was a main effect of caregiver behavior ($F_{(3,663)} = 130.40$, $p < 0.001$, $\eta^2 = 0.19$, Figure 2b) such that some combinations of behaviors happened more frequently than others. There was no main effect of condition ($F_{(1,221)} < 0.001$, $p = 0.99$, $\eta^2 < 0.001$), but there was an interaction between condition and caregiver behavior ($F_{(3,663)} = 523.04$, $p < 0.001$, $\eta^2 = 0.49$), such that, as expected given the higher rate of vocalization alone in the No Toy condition, the frequency of vocalizing while touching the baby was higher in the No Toy condition ($M_{Toy} = 1.22$, $M_{NoToy} = 52.11$, $t_{(221)} = -22.30$, $p < 0.001$, Cohen's $d = -2.17$), while any behavior combination including holding

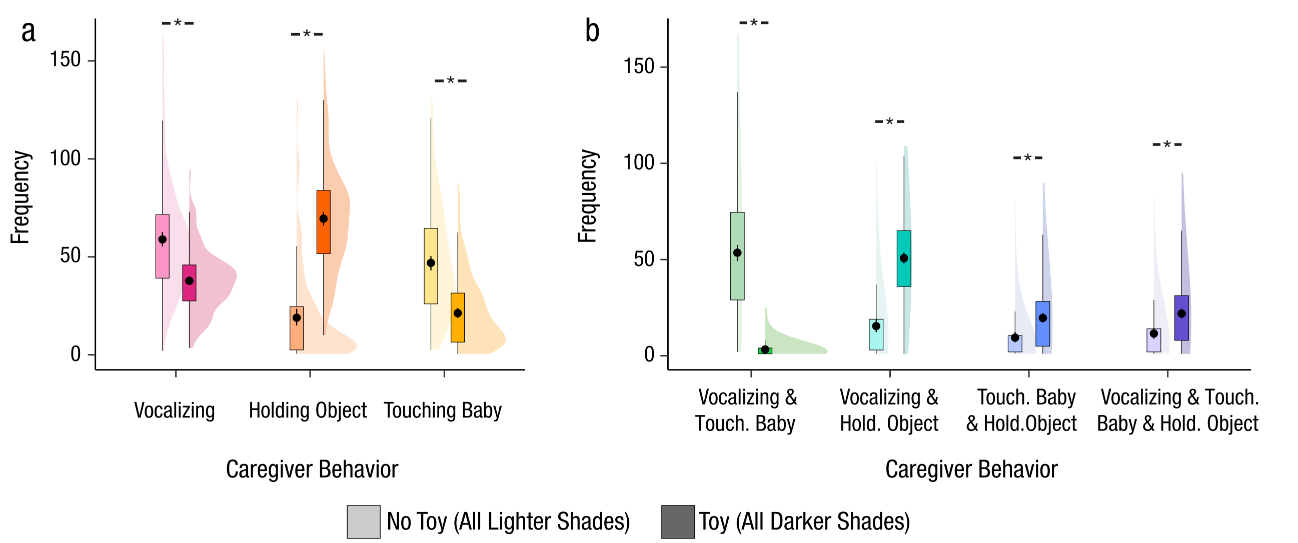

FIGURE 2 | Mean frequency (y-axis) of caregivers performing each behavior (x-axis) varies as a function of Toy condition (lighter versus darker shades) for (a) single behaviors (Vocalizing, pink; Holding Object, orange; Touching baby; yellow) and (b) combined behaviors (Vocalizing and Touching Baby, green; Vocalizing and Holding Object, teal; Touching Baby and Holding Object, blue; Vocalizing, Touching Baby, and Holding Object, purple) which were performed by the majority of caregivers (see Table 2). Box plots indicate interquartile ranges, black circles and error bars represent the mean and 95% confidence intervals around the mean, density plots indicate the distribution of values across participants in each condition. Significance stars (*) indicate p-values < 0.05 for pairwise t-tests comparing each behavior across conditions.

an object was more frequent in the Toy condition (Vocalizing while holding an object, $M_{Toy} = 50.73$, $M_{NoToy} = 8.27$, $t_{(221)} = 24.71$, $p < 0.001$, Cohen's $d = 2.22$; Touching the baby while holding an object, $M_{Toy} = 16.99$, $M_{NoToy} = 3.69$, $t_{(221)} = 10.84$, $p < 0.001$, Cohen's $d = 0.93$; Vocalizing while touching the baby while holding an object, $M_{Toy} = 19.75$, $M_{NoToy} = 5.42$, $t_{(221)} = 9.99$, $p < 0.001$, Cohen's $d = 0.90$). These results indicate that, perhaps unsurprisingly, the frequency of particular caregiver behaviors varies as a function of the context a caregiver is in.

3.2 | The Structure of Caregiver Behaviors Varies as a Function of Toy Condition

Having observed that the frequency of particular caregiver behaviors varies as a function of environmental affordances, we next asked whether the predictability of caregiver behavior differed by Toy and No Toy condition as well. We found that Overall Entropy was not different in the Toy and No Toy conditions ($M_{Toy} = 0.94$, $M_{NoToy} = 0.89$, $t_{(221)} = 1.77$, $p = 0.08$, Cohen's $d = 0.14$, Figure 3a). We also compared the entropy of caregivers' behavior to a randomly permuted null distribution of the entropy that would be expected by chance given the frequency of each behavior, and found that in both conditions, caregivers had lower entropy, or were more predictable, than would be expected by chance (see Supporting Information S2 for details). A caregiver's entropy was also correlated across conditions ($r = 0.30$, $t_{(220)} = 4.62$, $p < 0.001$), suggesting some stability in how predictable individual caregivers are. The broader cultural environment that caregivers lived in also related to their overall entropy, as Malawian caregivers had lower entropy behavior than South African caregivers across both conditions ($M_{Malawi} = 0.76$, $M_{SouthAfrica} = 1.01$, $t_{(218.23)} = -5.66$, $p < 0.001$, *Cohen's* $d = -0.76$).

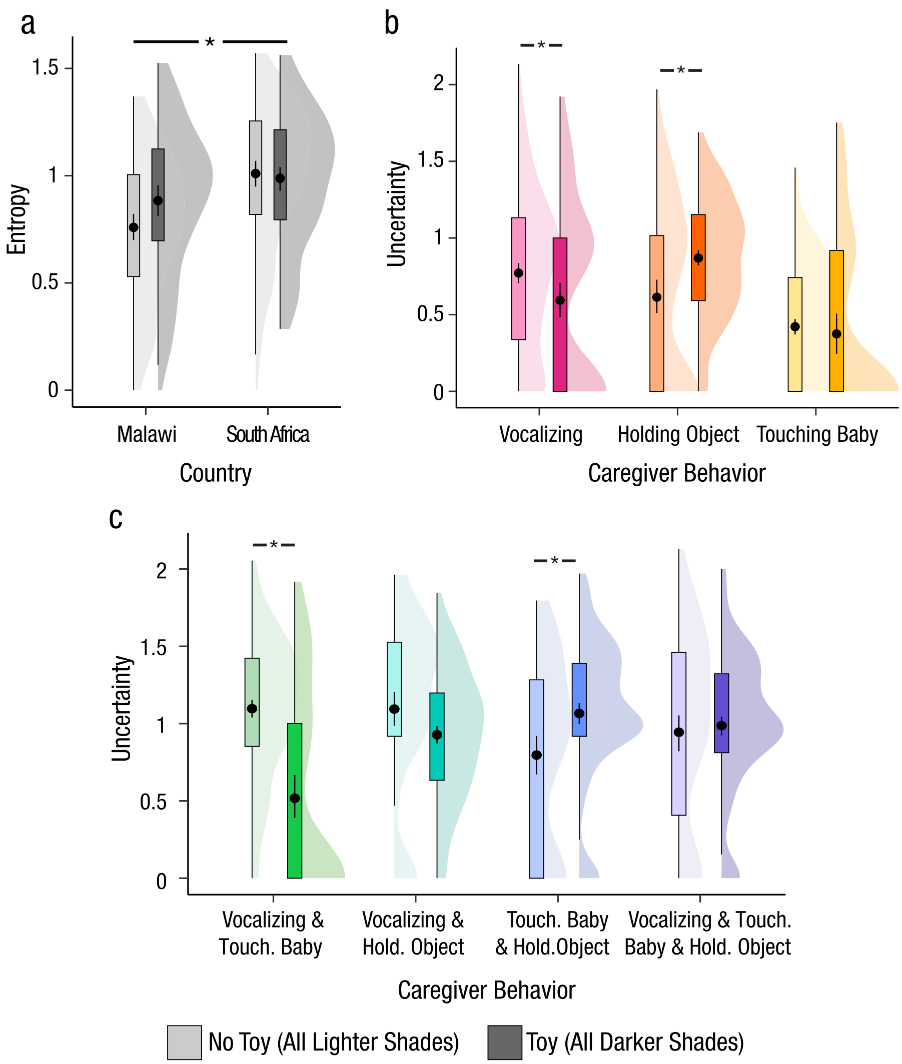

FIGURE 3 | (a) Caregivers' overall entropy (y-axis) differed by Country (x-axis), and marginally by Toy condition (lighter versus darker shades). (b, c) Caregivers' unpredictability (y-axis) following single behaviors (b, Vocalizing, pink; Holding Object, orange; Touching *B*aby; yellow) and combined behaviors (c, Vocalizing and *T*ouching *B*aby, green; Vocalizing and Holding Object, teal; Touching Baby and Holding Object, blue; Vocalizing, Touching Baby, and Holding Object, purple) differed between conditions (lighter versus darker shades) for some behaviors. In all panels, box plots indicate interquartile ranges, black circles and error bars represent the mean and 95% confidence intervals around the mean, density plots indicate the distribution of values across participants in each condition. Significance stars (*) indicate p-values < 0.05 for pairwise t-tests comparing each behavior across conditions.

Given that the frequency of some behaviors was higher in one condition than the other (Figure 2), we reasoned that the unpredictability of particular behaviors might differ across conditions even if the overall entropy did not. In other words, the structure of caregiver behavior might be driven by different behaviors across conditions. Indeed, planned pairwise comparisons of the unpredictability of each behavior across conditions indicated that caregiver vocalizations were more unpredictable in the No Toy condition ($M_{Toy} = 0.59$, $M_{NoToy} = 0.77$, $t_{(114)} = -2.70$, $p < 0.01$, Cohen's $d = -0.34$, Figure 3b), although this comparison became only marginally significant when considering our Bonferroni corrected threshold ($p = 0.008$, threshold was 0.007). Caregiver object holding was, instead, more unpredictable in the Toy condition ($M_{Toy} = 0.87$, $M_{NoToy} = 0.61$, $t_{(96)} = 4.44$, $p < 0.001$, Cohen's $d = 0.60$). There was no difference in the unpredictability of caregivers touching their baby between conditions ($M_{Toy} = 0.37$, $M_{NoToy} = 0.42$, $t_{(62)} = -1.32$, $p = 0.19$, Cohen's $d = -0.25$).

The unpredictability of combinations of behaviors also differed across conditions (Figure 3c). In this case, the unpredictability of caregivers vocalizing while touching their baby was higher in the No Toy condition ($M_{Toy} = 0.52$, $M_{NoToy} = 1.10$, $t_{(78)} = -7.25$, $p < 0.001$, Cohen's $d = -1.02$), while the unpredictability of caregivers holding an object while touching their baby was higher in the Toy condition ($M_{Toy} = 1.07$, $M_{NoToy} = 0.80$, $t_{(77)} = 4.50$ $p < 0.001$, Cohen's $d = 0.71$). There were no differences in the predictability of holding an object while vocalizing, nor vocalizing while holding an object while touching the baby across conditions (Holding an object while vocalizing, $M_{Toy} = 0.93$, $M_{NoToy} = 1.09$, $t_{(116)} = -1.59$, $p = 0.12$, Cohen's $d = -0.19$; Vocalizing while holding an object while touching the baby, $M_{Toy} = 0.99$, $M_{NoToy} = 0.94$, $t_{(96)} = 1.52$, $p = 0.13$, Cohen's $d = 0.21$). Still, this pattern of results indicates that the specific sensory signals caregivers provide their infants, and which infants might use to direct their attention, differ across time in relation to the affordances of the environment they are currently in.

3.3 | Across Dyads, Infant Looking Time Is Negatively Related to Caregivers' Overall Entropy

Having richly characterized the frequency and predictability of caregiver behavior, we asked how infant looking time to the caregiver during dyadic interaction relates to caregiver predictability. First, we investigated the relation between caregiver predictability and infant looking time by modeling the total amount of time infants looked at their caregiver as a function of her overall entropy and condition, while controlling for the duration for each condition. The results of the best fitting model indicated there were significant effects of caregiver entropy ($F_{(1,439)} = 48.93$, $p < 0.001$, $\eta^2 = 0.13$) and condition duration ($F_{(1,439)} = 36.42$, $p < 0.001$, $\eta^2 = 0.10$) on infant looking. There was also an effect of condition ($F_{(1,439)} = 54.52$, $p < 0.001$, $\eta^2 = 0.15$), but condition did not interact with caregiver entropy ($F_{(1,439)} = 1.02$, $p = 0.31$, $\eta^2 = 0.003$). Correspondingly, we found that overall caregiver entropy was significantly correlated with infant looking in both the Toy and No Toy conditions (Figure 4a; Toy, $r = -0.30$, $t_{(220)} = -4.83$, $p < 0.001$; No Toy, $r = -0.30$, $t_{(220)} = -4.71$, $p < 0.001$), such that the less predictable a caregiver was *on average*, the less her infant looked at her.

3.4 | Infant Looking Is Yoked to the Highest Entropy Caregiver Signal in Each Condition

Next, we turned to understanding the exact coupling between predictable caregiver behavior and infant looking across conditions. As the unpredictability of caregiver vocalizing and caregiver object holding differed across conditions, we asked whether infant looking time was related to caregivers' unpredictability following different sensory behaviors in each condition. Our model of infant looking as a function of caregiver vocalization showed significant effects of caregiver vocalization

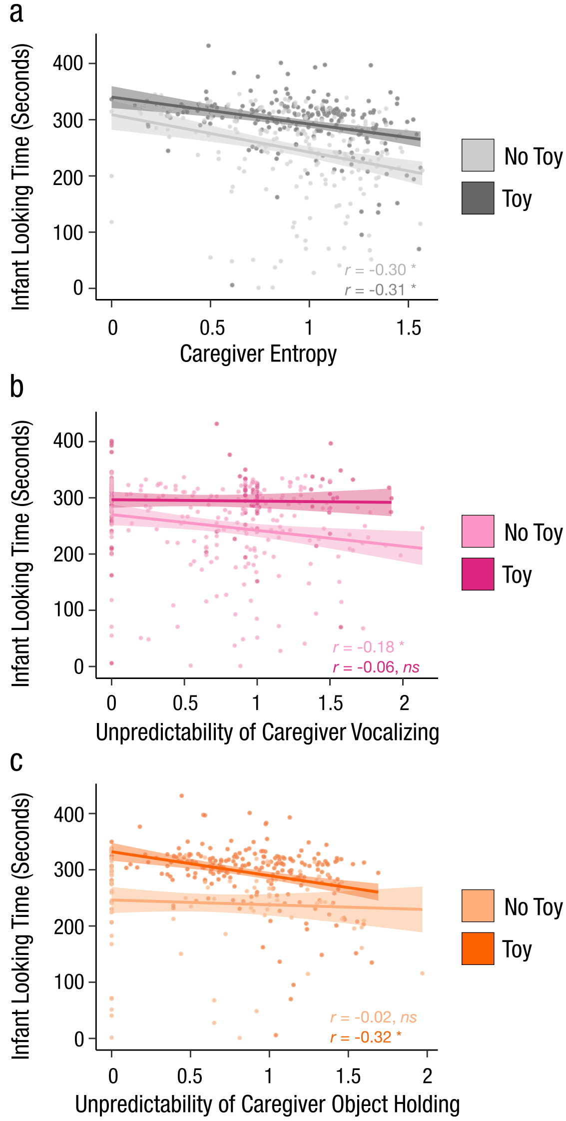

FIGURE 4 | (a) Total infant looking time (*y*-axis) correlates negatively with maternal entropy (*x*-axis) in both the No Toy (light gray) and Toy (dark gray) conditions. (b, c) Total infant looking time (*y*-axis) correlates with maternal unpredictability (*x*-axis) following maternal vocalizing (b, pink) in the No Toy condition (lighter shade), and with maternal object holding (c, orange) in the Toy condition (darker shades). In all panels, dots represent individual participants, shading around the linear fit represents 95% confidence intervals, *r* values indicate correlations in each condition, and significance stars (*) indicate correlation *p*-values < 0.05.

unpredictability ($F_{(1,326)} = 8.40$, $p = 0.004$, $\eta^2 = 0.03$), condition duration ($F_{(1,326)} = 25.70$, $p < 0.001$, $\eta^2 = 0.07$), and condition ($F_{(1,326)} = 20.59$, $p < 0.001$, $\eta^2 = 0.06$), like the overall entropy model. There was also a small but significant interaction between unpredictability following caregiver vocalization and condition ($F_{(1,326)} = 4.17$, $p = 0.04$, $\eta^2 = 0.01$), such that infant looking correlated significantly with unpredictability following caregivers' vocalizations in the No Toy condition ($r = -0.18$, $t_{(211)} = -2.71$, $p < 0.01$), but not in the Toy condition ($r = -0.06$, $t_{(96)} = -0.61$, $p = 0.54$, Figure 4b). Conversely, the results of a model examining how unpredictability following caregiver object holding related to infant looking demonstrated main effects of condition ($F_{(1,314)} = 56.72$, $p < 0.001$, $\eta^2 = 0.15$), duration ($F_{(1,314)} = 27.36$, $p < 0.001$, $\eta^2 = 0.08$), and unpredictability following object holding ($F_{(1,314)} = 14.60$, $p < 0.001$, $\eta^2 = 0.04$). There was also a significant interaction between condition and caregiver unpredictability following object holding ($F_{(1,314)} = 5.24$, $p = 0.02$, $\eta^2 = 0.02$), such that, in contrast to caregiver vocalization, infant looking behavior was significantly correlated with unpredictability following the caregivers' object holding in the Toy condition ($r = -0.32$, $t_{(219)} = -5.02$, $p < 0.001$), but not in the No Toy condition ($r = -0.02$, $t_{(116)} = -0.28$, $p = 0.78$; Figure 4c).

Importantly, although the frequency of these caregiver behaviors also differed by condition, the relation between caregiver behavior and looking was not significant if frequency was used in place of unpredictability (see Supporting Information S2). Together, these results strongly suggest that infants are sensitive not only to caregiver entropy in a general sense, but their looking behavior is driven by the sensory signal which has the highest entropy at any time. And, this looking behavior reflects infants' sensitivity to structure, rather than the frequency of particular behaviors.

3.5 | Within Dyads, Infants' Looking Is Shaped by Variability in Their Caregiver's Predictability Over Time

Past research has indicated that the relation between infant attention and predictable structure is u-shaped, such that infants attend to information which is neither too simple nor too hard to predict (Kidd et al. 2012). Our final analysis examined whether infants are sensitive to variations in the predictability of their caregiver's behavior which would allow them to engage and disengage with her in a u-shaped manner across time. To examine this, we modeled infant looking during particular time intervals as a function of within-caregiver entropy across conditions (i.e., with or without toys) and across countries (South Africa and Malawi). As a reminder, this within-caregiver entropy reflects how unpredictable a caregiver is being in the moment, relative to her own average. The results of the best fitting model showed significant linear ($F_{(1,206.2)} = 297.81$, $p < 0.001$, $\eta^2 = 0.59$) and quadratic effects ($F_{(1,172.7)} = 13.67$, $p < 0.001$, $\eta^2 = 0.07$) of within-caregiver entropy and condition ($F_{(1,223.7)} = 83.02$, $p < 0.001$, $\eta^2 = 0.27$) such that infants looked more at their caregiver in the Toy condition. There was no main effect of country ($F_{(1,220.6)} = 2.62$, $p = 0.11$, $\eta^2 = 0.01$) or condition duration ($F_{(1,301.8)} = 0.19$, $p = 0.66$, $\eta^2 < 0.001$). There were no interactions between the linear within-caregiver entropy term and condition ($F_{(1,5145.3)} = 2.26$, $p = 0.13$, $\eta^2 < 0.001$) or the quadratic within-caregiver entropy term and condition ($F_{(1,4050.2)} = 0.55$, $p = 0.46$, $\eta^2 < 0.001$). While there was no interaction between the linear within-caregiver entropy term and country ($F_{(1,207.7)} = 0.75$, $p = 0.39$, $\eta^2 = 0.004$), there was an interaction between the quadratic within-caregiver entropy term and country ($F_{(1,177.3)} = 6.67$, $p = 0.01$, $\eta^2 = 0.04$), such that the quadratic effect was somewhat stronger in caregiver-infant dyads from Malawi than from South Africa. Together, the results of this model indicate that within each caregiver-infant dyad, infants look at their caregiver the most when she is exhibiting a medium level of entropy, relative to her own range (Figure 5). Conversely, when she is being highly predictable or highly unpredictable relative to herself, her infant looks at her less. This suggests that infants adaptably engage and disengage with caregivers based on dynamic changes in the information content of the input they are providing.

To more closely examine the interaction between the quadratic within-caregiver entropy term and country, we ran a set of parallel follow-up models in each country alone. It is important to note that these models used data from many fewer subjects and are thus less well powered. That said, in Malawi there was a significant linear effect of within-caregiver entropy ($F_{(1,96.21)} = 118.63$, $p < 0.001$, $\eta^2 = 0.55$), a significant quadratic effect of within-caregiver entropy ($F_{(1,90.40)} = 19.47$, $p < 0.001$, $\eta^2 = 0.18$), a significant effect of condition ($F_{(1,3006.18)} = 214.77$, $p < 0.001$, $\eta^2 = 0.07$), and a significant main effect of condition duration ($F_{(1,3041.89)} = 11.90$, $p = 0.001$, $\eta^2 = 0.004$, although note the very small effect size). There was no interaction between the linear entropy term and condition ($F_{(1,2628.76)} = 0.004$, $p = 0.95$. $\eta^2 < 0.001$), although there was a small but significant interaction between the quadratic entropy term and condition ($F_{(1,2243.53)} = 5.81$, $p = 0.02$, $\eta^2 = 0.003$). In South Africa alone there were also significant linear effects of within-caregiver entropy ($F_{(1,107.44)} = 188.20$, $p < 0.001$, $\eta^2 = 0.64$) and condition ($F_{(1,114.68)} = 50.91$, $p < 0.001$, $\eta^2 = 0.31$). There was not, however, a significant quadratic effect of within-caregiver entropy ($F_{(1,89.06)} = 0.41$, $p = 0.52$, $\eta^2 = 0.005$). As in Malawi, there was no significant effect of condition duration ($F_{(1,177.92)} = 0.35$, $p = 0.55$, $\eta^2 = 0.002$), nor interactions between within-caregiver entropy and condition (linear: $F_{(1,2664.86)} = 2.56$, $p = 0.11$, $\eta^2 < 0.001$; quadratic: $F_{(1,2059.47)} = 0.85$, $p = 0.36$, $\eta^2 < 0.001$). One potential explanation for these differences is that the overall entropy of caregivers in South Africa was higher than that in Malawi (see Section 3.2). This means that the relatively low entropy timepoints for caregivers in South Africa are still higher entropy than those in Malawi and may still be providing infants with useful, learnable information. This idea is supported by the fact that the linear effect of entropy is stronger in South Africa than in Malawi. We return to these points in the discussion.

4 | Discussion

Here, we show that the temporal dynamics of caregiver predictability shape infant looking in caregiver-infant dyads from two African countries. Specifically, we showed the predictability of caregiver behavior during naturalistic interaction varies across caregivers, sensory behaviors, environmental contexts, and across time within the same caregiver. Strikingly, *all* of these fluctuations in caregiver predictability shaped

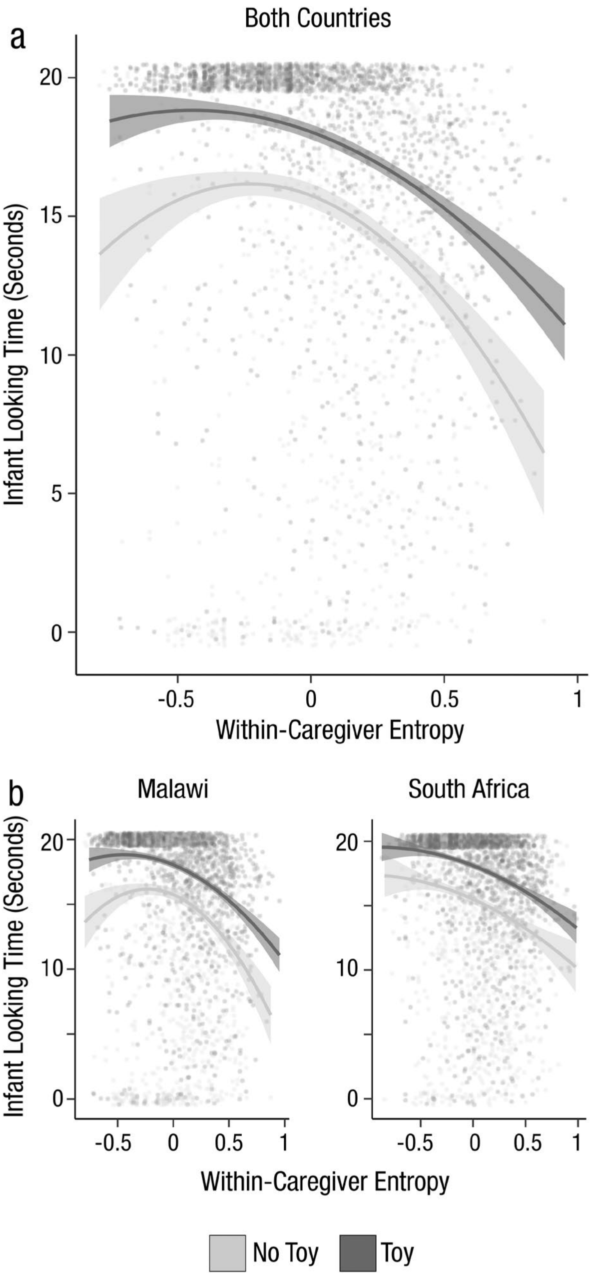

FIGURE 5 | Total infant looking time (y-axis, in seconds) is related to within-caregiver entropy values (x-axis) across 20 s time intervals in (a) the full sample, and (b) in Malawian (left graph) and South African (right graph) dyads, across the No Toy (light gray) and Toy (dark gray) conditions. Note that within-caregiver Entropy was standardized within each subject and thus 0 on the x-axis reflects each caregiver's mean. Each dot represents one 20-second interval for one caregiver-infant dyad during which we calculated caregiver entropy and infant looking. Shading around the quadratic fit represents 95% confidence intervals, and time interval dots are jittered by up to 500 milliseconds to illustrate many time windows during which infants looked at their caregiver for the duration of the 20 seconds.

moment-to-moment infant information gathering: infant looking time, an age-appropriate metric of information gathering, was yoked to the most informative sensory signal in each context, and infants attended to their caregiver most when she was slightly more predictable than *her norm*, suggesting infants use their prior experience to direct their attention during naturalistic interactions. We discuss each finding in turn.

First, the structure of caregiver behavior varied substantially. Caregiver predictability was correlated across environmental contexts (with or without toys present), such that some caregivers consistently provided a great deal of predictable information to their infants, while others were unpredictable in both contexts. Caregivers in Malawi were also more predictable than caregivers in South Africa, on average. That said, the predictability of particular behaviors varied similarly as a function of the Toy versus No Toy manipulation across the two countries. Without toys present, caregiver vocalizations were the least predictable signal, but caregiver object holding was the least predictable with toys present. Thus, while some caregivers are more or less predictable than others, the sensory signals they provide are not static. Instead, the information content of different caregiver behaviors, and therefore the behaviors from which infants could learn the most, shifts from situation to situation.

This variability in caregiver predictability then related to infant looking time. Across dyads, infants looked at their caregiver less if she was less predictable. And, this pattern in infant looking was driven by the least predictable behavior in each toy condition, meaning infants dynamically adjust their looking towards their caregiver in a manner yoked to the predictability of specific behaviors. As our measure of predictability (entropy) measures the amount of uncertainty in (and therefore how much someone can learn from) a behavior, this second finding demonstrates infants are sensitive to, and change their behavior as a function of, which sensory signal is currently providing the most information for them to learn. Notably, the relation between infant looking and caregiver entropy was still negative in each case. This suggests that although infants are using the highest entropy signal to shape their looking behavior, this does not mean they look at their caregiver all the time when she is engaging in that behavior, but instead use that behavior as a metric for when to engage or disengage with their caregiver.

Finally, infants shifted their looking behavior to and from their caregiver as her entropy fluctuated across time. Specifically, in our full sample, variability in caregivers' entropy across 20-second-windows related quadratically to variability in infant looking during those same windows—in fact, infants looked at their caregiver most when she was slightly more predictable than normal. In other words, each infant learned when it is most useful for them to look at their caregiver, *relative to their own regular input*. This finding adds real-world applicability to the few lab-based studies that have measured attention as a function of input predictability (Cubit et al. 2021; Kidd et al. 2012; Nussenbaum and Amso 2016). Information theoretic approaches (Gottlieb et al. 2013; Itti and Baldi 2009; Siegelman et al. 2019; Twomey and Westermann 2018) argue that this 'Goldilocks Effect' emerges because medium predictability information is neither too simple nor too complicated for someone's learning (Wade and Kidd 2019), and suggest that what counts as 'medium predictability' is relative to an individual's past experience (Forest et al. 2022). Our results mirror this, suggesting infants flexibly allocate their attention to caregivers based on when their caregiver is providing information that is useful for the infant in the moment.

While infants in both South Africa and Malawi showed temporal fluctuations in their looking behavior as a function of their caregiver's entropy, the exact function describing this link varied across sites. Specifically, South African infants looked less at their caregiver when she was being less predictable than normal, while infants in Malawi looked less at their caregiver when she was being more or less predictable than normal. Importantly, these analyses used within-caregiver entropy, which is our subject-specific measure of entropy reflecting how predictable a caregiver is being relative to her own typical behavior. However, one possible explanation for this between-country difference is that the overall entropy of South African caregivers (i.e., before normalizing within-caregiver) is higher than in Malawi. This means that moments during which South African caregivers are relatively predictable (i.e., moments with low within-caregiver entropy) are still relatively higher entropy than moments of low within-caregiver entropy for the Malawian dyads. This might help explain why South African infants do not look away from their caregivers in moments with low within-caregiver entropy. These results help provide context to the relation between infant looking and caregiver entropy: infants are sensitive to temporal fluctuations in their own caregiver's behavior, *and* the overall predictability of an infant's environment constrains the extent to which they attend to their caregiver across time. This finding helps to link these within-dyad results to our across-dyad finding that overall caregiver entropy is negatively related to infant looking, and past work suggesting a negative relation between caregiver entropy and children's long-term cognitive outcomes (Davis et al. 2017; Davis and Glynn 2024).

An open question is how caregiver predictability might shift over developmental time, as children and caregivers work together to scaffold continued learning. Other aspects of the environment become less predictable with age. For example, early visual environments shift from mostly consisting of simple contrasts to more complex scenes (Clerkin et al. 2017; Jayaraman and Smith 2019) which impacts what infants learn at different ages (Smith et al. 2018). Caregiver entropy may likewise increase with age if dyads collectively structure input to allow for more complexity. This suggestion is in line with work showing that the entropy of infant-directed speech decreases as children age, such that the way caregivers speak to their children becomes less predictable with time (Tal et al. 2024). On the other hand, dyadic interactions in toddlerhood are driven primarily by toddlers, not adults (Karmazyn-Raz and Smith 2023), suggesting caregiver predictability could become less important for shaping the development of core learning mechanisms as children play larger roles in structuring their interactions. This is a fruitful area for future research.

More broadly, our data open the door to answering many important questions about the documented relation between early caregiver predictability and later cognitive development (Birnie and Baram 2022; Davis and Glynn 2024). Here, we have shown there are temporally specific links between early infant looking and their caregiver's predictability. But is this really a viable mechanism for understanding how caregiver predictability relates to longer-term cognitive development? One possible answer is "No"—there could simply be a correlation between caregiver predictability and infant looking, and understanding the link between caregiver predictability and cognitive development might not be improved by accounting for early looking behavior. Alternatively, early caregiver behaviors could actually *train or tune* the development of basic attention and learning systems to expect particular environments at a time when the brain is highly malleable. At its most extreme, this would mean that the structure of early caregiver behavior becomes a model from which infants generalize how to learn from other sources of information down the line. This is consistent with recent work showing that caregiver predictability is related to infants neural responses during statistical learning months later (Forest et al. 2024), but a great deal more research is needed to flesh out this claim. Other important questions to answer in service of this goal will be to understand how tightly coupled individual dyads are, whether caregiver predictability or infant looking typically drives this relation, over what time-scales these relations emerge, what the exact cognitive processes affected are, or what the clinical implications of early disruptions to this process might be.

In support of this more mechanistic possibility, our data hint that moment-to-moment caregiver predictability is uniquely poised to tune infant development around the world by teaching infants how to allocate information-gathering resources efficiently. While other aspects of infants' environment are predictable (Saffran et al. 1996; Smith et al. 2018), it is unlikely that cross-culturally they provide the level of reward or motivation that could drive changes in what infants attend to. Visual attention emerges over the first year of life (Forest and Amso 2023; Hendry et al. 2019; Oakes and Amso 2018), but understanding how infants learn to allocate that emerging attention is less clear. That said, infants learn particularly well from social stimuli—spatial relations (Tummeltshammer et al. 2019), rule-guided behavior (Werchan and Amso 2021), and audio-visual pairings (Wu and Kirkham 2010) are all better learned when a social stimulus serves as the learning cue. Caregiver predictability may sit at the intersection of statistical and reinforcement learning to flexibly support early development.

Future work aside, the research presented here represents a shift in our understanding of the learning that dominates early life around the world. These findings help to forego a one-size-fits-all characterization of 'good' or 'bad' development that has historically marginalized those in non-western contexts (Draper et al. 2023; Nketia et al. 2021) in favor of the notion that infants' experience helps them adapt to common and motivating sources of information in their own environment. These findings may also center caregiver-child interactions as a globally-relevant metric of behavioral health, by highlighting the key role of predictability during early dyadic interaction for shaping core cognitive processes.

Acknowledgments

The materials, analytic code, and data for this work are not currently publicly accessible.

Data Availability Statement

We wish to thank the many research assistants who helped with data collection for this project, including Ringie Gulwa and Pamela Madikane in South Africa, and Grace Baloyi, Emmie Gausi, Colleta

Mphasa, Meffa Saukira, Friday Nantongwe, and Innocent Mpakiza in Malawi. We are also very grateful to all the families who participated, and Wellcome Leap for providing funding. The materials, analytic code, and data for this work are not currently publicly accessible. This work was not preregistered.

References

Aslin, R. N. 2017. "Statistical Learning: A Powerful Mechanism That Operates by Mere Exposure." *Wiley Interdisciplinary Reviews: Cognitive Science* 8: 1–7. https://doi.org/10.1002/wcs.1373.

Bates, D., M. Machler, B. M. Bolker, and S. C. Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67: 1–48. https://doi.org/10.18637/jss.v067.i01.

Ben-Shachar, M. S., D. Lüdecke, and D. Makowski. 2020. "Effectsize: Estimation of Effect Size Indices and Standardized Parameters." *Journal of Open Source Software* 5, no. 56: 2815. https://doi.org/10.21105/joss.02815.

Berlyne, D. E., J. C. Ogilvie, and L. C. C. Parham. 1971. "The Dimensionality of Visual Complexity, Interestingness, and Pleasingness." *Canadian Journal of Psychology* 25: 195–206. https://doi.org/10.1037/h00823811971.

Birnie, M. T., and T. Z. Baram. 2022. "Principles of Emotional Brain Circuit Maturation." *Science* 376: 1055–1056. https://doi.org/10.1126/science.abn4016.

Clerkin, E. M., E. Hart, J. M. Rehg, C. Yu, and L. B. Smith. 2017. "Real-World Visual Statistics and Infants' First-Learned Object Names." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 372: 20160055. https://doi.org/10.1098/rstb.2016.0055.

Colombo, J. 1993. *Infant Cognition: Predicting Later Intellectual Functioning*. SAGE Publications.

Colombo, J., L. Kapa, and L. Curtindale. 2010. "Varieties of Attention in Infancy." In *Infant Perception and Cognition: Recent Advances, Emerging Theories, and Future Directions*, edited by L. Oakes, C. Cashon, M. Casasola, and D. Rakison. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195366709.003.0001.

Cubit, L. S., R. Canale, R. Handsman, C. Kidd, and L. Bennetto. 2021. "Visual Attention Preference for Intermediate Predictability in Young Children." *Child Development* 92: 1–703. https://doi.org/10.1111/cdev.13536.

Datavyu Team. 2014. *Datavyu: A Video Coding Tool [Computer Software]*. Databrary Project. http://datavyu.org.

Davis, E. P., and L. M. Glynn. 2024. "Annual Research Review: The Power of Predictability – Patterns of Signals in Early Life Shape Neurodevelopment and Mental Health Trajectories." *Journal of Child Psychology and Psychiatry* 65: 508–534. https://doi.org/10.1111/jcpp.13958.

Davis, E. P., S. A. Stout, J. Molet, et al. 2017. "Exposure to Unpredictable Maternal Sensory Signals Influences Cognitive Development Across Species." *Proceedings of the National Academy of Sciences* 114: 10390–10395. https://doi.org/10.1073/pnas.1703444114.

Demaestri, C., M. Gallo, E. Mazenod, et al. 2022. "Resource Scarcity but Not Maternal Separation Provokes Unpredictable Maternal Care Sequences in Mice and Both Upregulate Crh-Associated Gene Expression in the Amygdala." *Neurobiology of Stress* 20: 100484. https://doi.org/10.1016/j.ynstr.2022.100484.

Draper, C. E., L. M. Barnett, C. J. Cook, et al. 2023. "Publishing Child Development Research From Around the World: An Unfair Playing Field Resulting in Most of the World's Child Population Under-Represented in Research." *Infant and Child Development* 32: e2375. https://doi.org/10.1002/icd.2375.

Faul, F., E. Erdfelder, A.-G. Lang, and A. Buchner. 2007. "G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39: 175–191. https://doi.org/10.3758/BF03193146.

Forest, T. A., and D. Amso. 2023. "Neurodevelopment of Attention, Learning, and Memory Systems in Infancy." *Annual Review of Developmental Psychology* 5: 45–65. https://doi.org/10.1146/annurev-devpsych-120321-011300.

Forest, T. A., S. A. McCormick, L. Davel, et al. 2024. "Early Caregiver Predictability Shapes Neural Indices of Statistical Learning Later in Infancy." *Developmental Science* 28, no. 1: 1–15. https://doi.org/10.1111/desc.13570.

Forest, T. A., N. Siegelman, and A. S. Finn. 2022. "Attention Shifts to More Complex Structures With Experience." *Psychological Science* 33: 09567976221114055. https://doi.org/10.1177/09567976221114055.

Franchak, J. M., K. S. Kretch, and K. E. Adolph. 2018. "See and Be Seen: Infant–Caregiver Social Looking During Locomotor Free Play." *Developmental Science* 21: e12626. https://doi.org/10.1111/desc.12626.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Mifflin and Company.

Gottlieb, J., P. Y. Oudeyer, M. Lopes, and A. Baranes. 2013. "Information-Seeking, Curiosity, and Attention: Computational and Neural Mechanisms." *Trends in Cognitive Sciences* 17: 585–593. https://doi.org/10.1016/j.tics.2013.09.001.

Hendry, A., M. H. Johnson, and K. Holmboe. 2019. "Early Development of Visual Attention: Change, Stability, and Longitudinal Associations." *Annual Review of Developmental Psychology* 1: 251–275. https://doi.org/10.1146/annurev-devpsych-121318-085114.

Holmberg, E., E.-L. Kataja, E. P. Davis, et al. 2022. "Unpredictable Maternal Sensory Signals in Caregiving Behavior Are Associated With Child Effortful Control." *PLoS One* 17: e0279384. https://doi.org/10.1371/journal.pone.0279384.

Itti, L., and P. Baldi. 2009. "Bayesian Surprise Attracts Human Attention." *Vision Research* 49: 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007.

Jayaraman, S., and L. B. Smith. 2019. "Faces in Early Visual Environments Are Persistent Not Just Frequent." *Vision Research* 157: 213–221. https://doi.org/10.1016/j.visres.2018.05.005.

Jensen, G., R. D. Ward, and P. D. Balsam. 2013. "Information: Theory, Brain, and Behavior." *Journal of the Experimental Analysis of Behavior* 100: 408–431. https://doi.org/10.1002/jeab.49.

Karmazyn-Raz, H., and L. B. Smith. 2023. "Sampling Statistics Are Like Story Creation: A Network Analysis of Parent-Toddler Exploratory Play." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 378: 20210358. https://doi.org/10.1098/rstb.2021.0358.

Kidd, C., S. T. Piantadosi, and R. N. Aslin. 2012. "The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple nor Too Complex." *PLoS One* 7: 1–8. https://doi.org/10.1371/journal.pone.0036399.

Kluyver, T., B. Ragan-Kelley, F. Pérez, et al. 2016. *Jupyter Notebooks – A Publishing Format for Reproducible Computational Workflows*, edited by F. Loizides and B. Scmidt, 87–90. IOS Press. https://doi.org/10.3233/978-1-61499-649-1-87.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. "lmerTest Package: Tests in Linear Mixed Effects." *Journal of Statistical Software* 82, no. 13: 1–26. https://doi.org/10.18637/jss.v082.i13.

McCall, R. B., and M. S. Carriger. 1993. "A Meta-Analysis of Infant Habituation and Recognition Memory Performance as Predictors of Later IQ." *Child Development* 64: 57–79. https://doi.org/10.1111/j.1467-8624.1993.tb02895.x.

McHugh, M. L. 2012. "Interrater Reliability: The Kappa Statistic." *Biochemia Medica* 22: 276–282.

Miller, G. A. 1956. “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information.” *Psychological Review* 63: 81–97. https://doi.org/10.1037/h0043158.

Nencheva, M. L., E. A. Piazza, and C. Lew-Williams. 2021. “The Moment-To-Moment Pitch Dynamics of Child-Directed Speech Shape Toddlers' Attention and Learning.” *Developmental Science* 24: e12997. https://doi.org/10.1111/desc.12997.

Nketia, J., D. Amso, and N. H. Brito. 2021. “Towards a More Inclusive and Equitable Developmental Cognitive Neuroscience.” *Developmental Cognitive Neuroscience* 52: 101014. https://doi.org/10.1016/j.dcn.2021.101014.

Nussenbaum, K., and D. Amso. 2016. “An Attentional Goldilocks Effect: An Optimal Amount of Social Interactivity Promotes Word Learning From Video.” *Journal of Cognition and Development: Official Journal of the Cognitive Development Society* 17: 30–40. https://doi.org/10.1080/15248372.2015.1034316.

Oakes, L., and D. Amso. 2018. “Development of Visual Attention.” In *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, 1–33. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119170174.epcn401.

Poli, F., G. Serino, R. B. Mars, and S. Hunnius. 2020. “Infants Tailor Their Attention to Maximize Learning.” *Science Advances* 6, no. 39: eabb5053. https://doi.org/10.1126/sciadv.abb5053.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing [Computer Software]. R Foundation for Statistical*. R Core Team. https://www.R-project.org/.

Rieke, F., D. Warland, R. de van Ruyter Steveninck, and W. Bialek. 1997. *Spikes: Exploring the Neural Code*. MIT Press.

Rose, S. A., and J. F. Feldman. 1995. “Prediction of IQ and Specific Cognitive Abilities at 11 Years From Infancy Measures.” *Developmental Psychology* 31: 685–696. https://doi.org/10.1037/0012-1649.31.4.685.

Saffran, J. R., R. N. Aslin, and E. L. Newport. 1996. “Statistical Learning by 8-Month-Old Infants.” *Science* 274: 1926–1928. https://doi.org/10.1126/science.274.5294.1926.

Sameroff, A. 2009. “The Transactional Model.” In *The Transactional Model of Development: How Children and Contexts Shape Each Other*, edited by A. Sameroff, 3–21. American Psychological Association. https://doi.org/10.1037/11877-001.

Shannon, C. E. 1948. “A Mathematical Theory of Communication.” *Bell System Technical Journal* 27: 623–656. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.

Siegelman, N., L. Bogaerts, and R. Frost. 2019. “What Determines Visual Statistical Learning Performance? Insights From Information Theory.” *Cognitive Science* 43: e12803. https://doi.org/10.1111/cogs.12803.

Smith, L. B., S. Jayaraman, E. Clerkin, and C. Yu. 2018. “The Developing Infant Creates a Curriculum for Statistical Learning.” *Trends in Cognitive Sciences* 22: 325–336. https://doi.org/10.1016/j.tics.2018.02.004.

Suarez-Rivera, C., L. B. Smith, and C. Yu. 2019. “Multimodal Parent Behaviors Within Joint Attention Support Sustained Attention in Infants.” *Developmental Psychology* 55: 96–109. https://doi.org/10.1037/dev0000628.

Tal, S., E. Grossman, and I. Arnon. 2024. “Infant-Directed Speech Becomes Less Redundant as Infants Grow: Implications for Language Learning.” *Cognition* 249: 105817. https://doi.org/10.1016/j.cognition.2024.105817.

Tummeltshammer, K., E. C. H. Feldman, and D. Amso. 2019. “Using Pupil Dilation, Eye-Blink Rate, and the Value of Mother to Investigate Reward Learning Mechanisms in Infancy.” *Developmental Cognitive Neuroscience* 36: 100608. https://doi.org/10.1016/j.dcn.2018.12.006.

Twomey, K. E., and G. Westermann. 2018. “Curiosity-Based Learning in Infants: A Neurocomputational Approach.” *Developmental Science* 21: e12629. https://doi.org/10.1111/desc.12629.

Vanoncini, M., N. Boll-Avetisyan, B. Elsner, S. Hoehl, and E. Kayhan. 2022. “The Role of Mother-Infant Emotional Synchrony in Speech Processing in 9-Month-Old Infants.” *Infant Behavior and Development* 69: 101772. https://doi.org/10.1016/j.infbeh.2022.101772.

Vanoncini, M., S. Hoehl, B. Elsner, S. Wallot, N. Boll-Avetisyan, and E. Kayhan. 2024. “Mother-Infant Social Gaze Dynamics Relate to Infant Brain Activity and Word Segmentation.” *Developmental Cognitive Neuroscience* 65: 101331. https://doi.org/10.1016/j.dcn.2023.101331.

Wade, S., and C. Kidd. 2019. “The Role of Prior Knowledge and Curiosity in Learning.” *Psychonomic Bulletin and Review* 26: 1377–1387. https://doi.org/10.3758/s13423-019-01598-6.

Weisberg, F. J. 2019. *An R Companion to Applied Regression*. 3rd ed. Sage. https://www.john-fox.ca/Companion/index.html.

Werchan, D. M., and D. Amso. 2021. “All Contexts Are Not Created Equal: Social Stimuli Win the Competition for Organizing Reinforcement Learning in 9-Month-Old Infants.” *Developmental Science* 24: e13088. https://doi.org/10.1111/desc.13088.

Wu, R., and N. Z. Kirkham. 2010. “No Two Cues Are Alike: Depth of Learning During Infancy Is Dependent on What Orients Attention.” *Journal of Experimental Child Psychology* 107: 118–136. https://doi.org/10.1016/j.jecp.2010.04.014.

Yu, C., and L. B. Smith. 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.

Zieff, M. R., M. Miles, E. Mbale, et al. 2024. “Characterizing developing executive functions in the first 1000 days in South Africa and Malawi: The Khula Study.” *Wellcome Open Research*. https://doi.org/10.12688/wellcomeopenres.19638.1.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.