RESEARCH ARTICLE

Individual differences in processing speed and curiosity explain infant habituation and dishabituation performance

Francesco Poli[1] | **Tommaso Ghilardi**[1,2] | **Roseriet Beijers**[3,4] | **Carolina de Weerth**[4] | **Max Hinne**[1] | **Rogier B. Mars**[1,5] | **Sabine Hunnius**[1]

[1]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands

[2]Centre for Brain and Cognitive Development, School of Psychological Sciences, Birkbeck, University of London, London, United Kingdom

[3]Behavioral Science Institute, Radboud University, Nijmegen, Netherlands

[4]Donders Institute for Brain, Cognition and Behaviour, Department of Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, Netherlands

[5]Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, UK

Correspondence
Francesco Poli, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands.
Email: francesco.poli@donders.ru.nl

Funding information
Wellcome Trust; Nederlandse Organisatie voor Wetenschappelijk Onderzoek; Marie Curie; Koninklijke Nederlandse Akademie van Wetenschappen; Biotechnology and Biological Sciences Research Council

Abstract

Habituation and dishabituation are the most prevalent measures of infant cognitive functioning, and they have reliably been shown to predict later cognitive outcomes. Yet, the exact mechanisms underlying infant habituation and dishabituation are still unclear. To investigate them, we tested 106 8-month-old infants on a classic habituation task and a novel visual learning task. We used a hierarchical Bayesian model to identify individual differences in sustained attention, learning performance, processing speed and curiosity from the visual learning task. These factors were then related to habituation and dishabituation. We found that habituation time was related to individual differences in processing speed, while dishabituation was related to curiosity, but only for infants who did not habituate. These results offer novel insights in the mechanisms underlying habituation and serve as proof of concept for hierarchical models as an effective tool to measure individual differences in infant cognitive functioning.

KEYWORDS
Bayesian modelling, curiosity, habituation, individual differences, infant cognitive development, processing speed

Research Highlights

- We used a hierarchical Bayesian model to measure individual differences in infants' processing speed, learning performance, sustained attention, and curiosity.
- Faster processing speed was related to shorter habituation time.
- High curiosity was related to stronger dishabituation responses, but only for infants who did not habituate.

1 | INTRODUCTION

When a stimulus is presented to an infant multiple times, the time they spend looking at it will gradually decrease. Once a novel stimulus is presented, the infant will look at it for a longer amount of time compared to the previous stimulus. These two phenomena are called habituation and dishabituation, respectively. Although apparently simple, they have given rise to diverse theories about their origin and function (Rankin et al., 2009), and multiple cognitive functions have been proposed to underlie them in a research effort that now spans

more than a century (Valentine, 1913). Habituation performance in infancy has considerable predictive power for school age, adolescence, and even adulthood intelligence (Bornstein & Sigman, 1986; Fagan et al., 2007). Infants who habituate faster and spend less time attending to the repeated stimulus presentation (also named 'short lookers') tend to score higher on IQ tests later in life and perform better in their academic career than infants who attend to the same stimuli for longer ('long lookers'; Bornstein & Colombo, 2012). The correlations between infants' looking behaviour and their later IQ have been shown to span up to 25 years of development (Fagan et al., 2007) and explain up to 40% of the variance in cognitive performance (Fagan & McGrath, 1981). The strong evidence for the predictive power of habituation performance for later cognitive functioning (Kavšek, 2004) emphasizes the importance of identifying the cognitive functions underlying it to understand and predict developmental change.

Various attentional, learning and memory processes have been associated with habituation and dishabituation. Influential research considers habituation to be mainly an attentional phenomenon: Several studies have related habituation to alerting (Gardner et al., 2003), orienting and disengagement (Blaga & Colombo, 2006; Colombo et al., 2001; McCall, 1994), and executive attention (Colombo & Cheatham, 2006). Taking a different stance on the underlying processes, Rose and Feldman (1997) argue that habituation and dishabituation are related to learning. Specifically, habituation is thought to reflect the formation and consolidation of memory traces for a stimulus, while dishabituation requires successfully retrieving the previous stimulus and comparing it to the novel one. However, in a following study, Feldman and Mayes (1999) argue that looking time to the stimuli can be taken as a measure of processing speed, as infants who are faster in processing visual stimuli might need less time to form a memory trace and thus habituate faster.

Recent theories reframe habituation as an active process: Infants might not only process information, but they might also actively decide whether and when to do so (Cao et al., 2022). Although previous research looked at curiosity as a state, with information gains triggering a momentary increase of curiosity, here we look at curiosity as a trait that differs across individuals. In this perspective, fast habituation is not only the result of efficient memory consolidation or high processing speed but also of infants' sensitivity to the informativity of a stimulus. Indeed, recent evidence shows that infants adapt their allocation of attention depending on the information gain offered by a stimulus (Ghilardi et al., 2023; Poli et al., 2020). This inherent bias towards information (and away from a lack thereof) is a central aspect of curiosity (Gottlieb & Oudeyer, 2018). Individual differences in habituation and dishabituation might thus be a result of individual differences in curiosity in addition to differences in attentional and learning processes. Understanding these mechanisms would also allow us to better identify the reason why habituation performance has robust predictive power of later IQ and academic achievements.

In the current paper, we devised a novel approach to investigating the cognitive functions underlying infant (dis)habituation. Specifically, we tested infants on a visual learning task and used a hierarchical Bayesian model to extract parameters indexing individual differences in attention, learning, processing speed, and curiosity. This allowed us to obtain estimates of individual differences that are more refined than the behavioural data themselves. Infants were also tested on a habituation task to relate the parameters of the visual learning task to their habituation and dishabituation performance. Hence, this model-based approach allowed us to assess multiple cognitive factors at the same time, and investigate their unique contribution to habituation and dishabituation. We explored whether attention (Blaga & Colombo, 2006), learning (Rose & Feldman, 1997), processing speed (Feldman & Mayes, 1999) or curiosity (Cao et al., 2022) would correlate with habituation and dishabituation, and also considered it possible that multiple cognitive factors concurrently predicted habituation and dishabituation.

2 | METHODS

Participants. We recruited 164 infants from a database of volunteer families and via social media advertisement. Twenty infants did not reach inclusion criteria for the visual learning task (i.e., a minimum of 20 trials across the whole task) and 25 infants did not finish the habituation task (e.g., because they became fussy). Ten additional infants were excluded because they were too old at the age of testing (i.e., older than 9 months of age). The final sample consisted of 106 infants (Mean age = 8.0 months, SD = 0.5 months, 59 females). The study was approved by the ethics committee of Radboud University, Nijmegen (NL). For 90 of the initial 164 infants, the data of the visual learning task has been analysed for different purposes in previous research (Poli et al., 2020, 2023). The remaining 74 infants were participants of a larger longitudinal cohort study (i.e., Smiley).

Procedure. Mothers and their infants were invited to the laboratory of the Baby and Child Research Centre (Nijmegen, The Netherlands). The visit consisted of a visual learning task and a habituation task. The order of the tasks was counter-balanced across infants: the first half of the infants performed the visual learning task first, while the second half performed the habituation task first. Infants were seated in a baby car seat and positioned on the parent's lap, at a 60 to 65 cm distance from the eye tracker (Tobii Pro TX300). During the two tasks, the looking behaviour of the infant was monitored using the eye tracker and an external video camera. Infants took a short break between the two tasks. Parents were instructed not to interact with their child, unless infants sought their attention and, and even in that case, not to try to bring infants' attention back to the screen.

2.1 | Materials

The visual learning task. The visual learning task is described in detail in Poli et al. (2020). Infants were shown 16 sequences of cue-target trials. In each sequence, the cue consisted of a simple shape appearing in the middle of the screen. The target was the same shape appearing in one of four screen quadrants around the cue location (see Figure 1a).
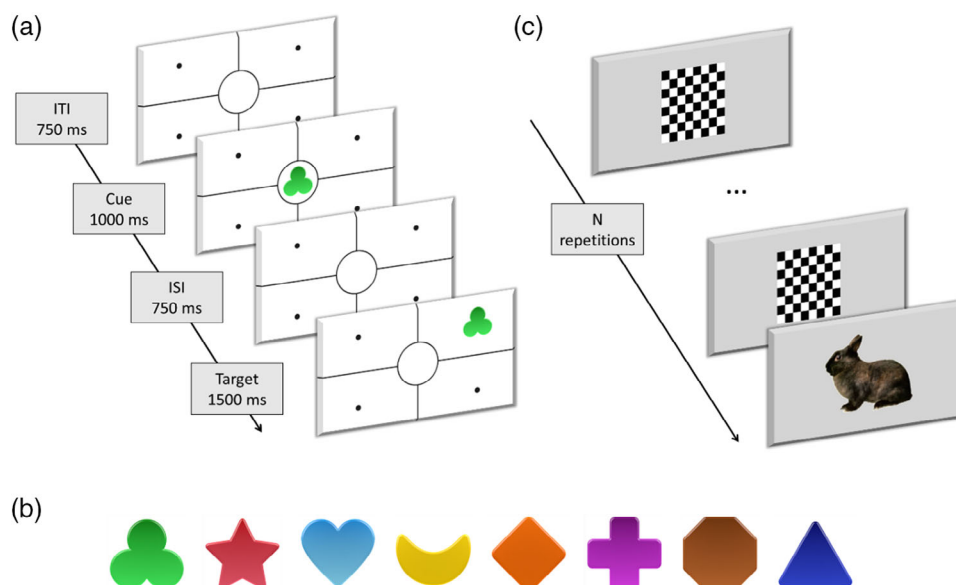
FIGURE 1 (a) One example trial of the visual learning task. (b) The shapes used in the visual learning task. Every sequence of the task contained a different shape. (c) The habituation task.

The shape was the same across all trials of the sequence but changed across sequences. In all sequences, the target was more likely to appear in one location than the others. Specifically, it could appear in the high-likelihood location 100% of the times (in four sequences), 80% of the times (in six sequences), and 60% of the times (in six sequences). Infants could, therefore, learn to predict the most likely target location of each sequence. The sequences were shown one after the other. When the infant looked away from any area of the screen for 1 s or more, the sequence was stopped. When the infant looked back at the screen, the following sequence was played. If infants did not look away from the screen, the sequence would automatically stop after 15 cue-target trials. The experiment lasted until the infant had watched all 16 sequences or became fussy. On average, infants watched 8.5 trials in each sequence (SD = 2.1) for a total of 7.3 sequences (SD = 3.0). Following previous work (Kidd et al., 2012; Poli et al., 2020), sequences in which infants looked at 3 trials or less were discarded, as it is not possible to infer the probabilistic structure of a sequence from such a limited number of observations.

Two independent variables were computed from the stimuli using information theory (Cover & Thomas, 1991). Specifically, we quantified trial-by-trial changes in the predictability of the sequences, as well as the information gain contained in each stimulus. If infants can track the levels of predictability and information gain that are present in the task, their looking behaviour should correlate with these information-theoretic measures (O'Reilly et al., 2013; Poli et al., 2023). This is addressed in detail in the Analysis section.

Three dependent variables were collected based on the infants' looking behaviour during this task:

- *Look-away.* After the presentation of each target stimulus, we registered a binary measure, noting whether infants kept looking at the screen or looked away within the next 4 s (i.e., one trial). This is taken

as a measure of the interest in the current stimulus (Kidd & Hayden, 2015) and since no external distractors were present, this is considered an active decision made by the infants to stop engaging with the task (cf. Kidd et al., 2012).
- *Saccadic latency.* We measured how quickly infants moved their eyes from the cue to the target location after the target appeared. This has been shown to be related to stimulus predictability: infants are faster in looking at predictable stimuli and slower in redirecting their attention to surprising events (Kayhan et al., 2019; Poli et al., 2020).
- *Looking time.* We measured how long infants looked at the target from the moment it appeared to 750 ms after its disappearance. Both in infants and adults, looking time has been found to be related to the informational value of the stimulus (O'Reilly et al., 2013; Poli et al., 2020).

The habituation task. The habituation task was taken from Addyman (2015; Figure 1c). A checkerboard was presented on the screen, accompanied by a sound. Every time the infants looked away from the screen for 1 s, the trial ended and the checkerboard disappeared from the screen. Then, a new trial started and the same stimulus was shown again, with the same procedure carried over for a maximum of 13 trials. If the overall looking time in the last three trials amounted to less than 50% of the looking time in the first three trials, the habituation criterion was met and the procedure was stopped. After this, a novel stimulus (a rabbit) was presented on the screen, accompanied by the same sound as before. When infants stopped looking at the stimulus for more than 1 s, the task ended.

Two dependent variables were collected based on the infants' looking behaviour during this task:

- *Habituation time.* We measured the overall time (in seconds) spent looking at the checkerboard across all trials.
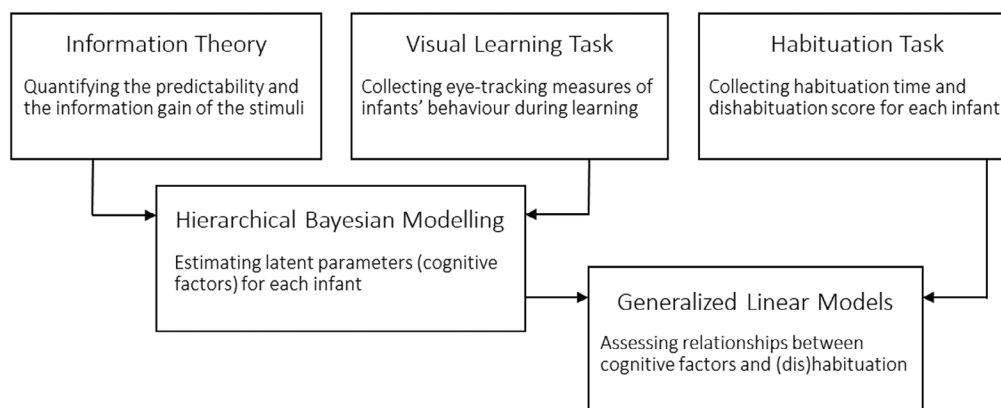
FIGURE 2 The research pipeline. We related the information-theoretic quantifications of predictability and information gain of the stimulus material to infants' behaviour during the visual learning task to infer their cognitive functioning in terms of processing speed, attention, learning, and curiosity. Then, we related these aspects of their cognitive functioning to habituation and dishabituation.

- *Dishabituation score.* We computed the proportion of looking time to the novel stimulus (i.e., the rabbit) over the sum of the looking time to the last habituation stimulus (i.e., the checkerboard on the last trial) and the novel stimulus.

2.2 | Analysis

Our aim was to identify individual differences in multiple cognitive factors (i.e., sustained attention, processing speed, learning performance and curiosity), and relate them to habituation and dishabituation. Figure 2 illustrates all the steps of our analysis. We develop a hierarchical Bayesian model that infers the values of latent parameters from the infants' looking behaviour during the visual learning task (i.e., look-away, saccadic latency and looking time). We argue that these latent parameters have two desirable features. First, they can be related to specific cognitive functions. As such, they should not be interpreted solely as statistical parameters, but also as cognitively meaningful factors. Second, these parameters may display different values across individuals. Together, these two features allow us to obtain reliable estimates of individual differences in cognitive factors.

The model was fitted using dependent and independent variables. The dependent variables were infants' looking behaviour during the visual learning task (look-away, saccadic latency and looking time), and the independent variables were trial-by-trial quantifications of the level of predictability of the stimuli and the amount of information gain that was contained in each stimulus. By relating the infants' behaviour to estimates of predictability and information gain in the stimulus material, we can assess individual differences in how infants process incoming stimuli.

First, we introduce the information theoretic (Cover & Thomas, 1991) variables that we used to quantify the predictability and information gain of the stimuli. Second, we describe how we related these variables to the infants' looking behaviour and the latent parameters (i.e., cognitive factors) that allowed us to do so. Third, we assess whether we find individual differences in these latent param-

eters. Finally, we related these latent parameters to habituation and dishabituation performance.

Information-theoretic measures. For every trial of the learning task, we quantified the trial-by-trial level of predictability of the sequence and the amount of information carried by each stimulus (Figure 3b, and cf. Poli et al., 2020). The level of predictability was computed from the probability that the target stimulus would appear in each target location. For example, if the stimulus is equally likely to appear in every location (i.e., 25%), the predictability is at its minimum. Using information theory (Cover & Thomas, 1991), we quantified predictability as the negative entropy (H) of a given trial:

$$-H = \sum_{i=1}^{k} P_t \log_2 (P_t)$$

Where k indicates the number of the target locations (i.e., 4), P indicates the probability of the target to appear in each of the four locations (e.g., [25%, 25%, 25%, 25%]), and t indicates the trial number.

The more a new event changes prior probabilities, the more information it carries. Hence, the information gain of each stimulus is computed by quantifying changes in predictability. For example, if the stimulus is presented in the upper-left corner, the probabilities will change from chance level (i.e., 25%), and the probability that the stimulus will appear in the upper-left corner will increase (e.g., 40%), while the probability that it will appear in any other location will decrease (e.g., 20%). As a result, thanks to the last observation, we know more about the target location. Kullback-Leibler Divergence offers a precise way of computing the amount of information that is contained in the new observation, or the information gain (IG):

$$IG = \sum_{i=1}^{k} P_t \log_2 (P_t) - \sum_{i=1}^{k} P_t \log_2 (P_{t-1})$$

where P_t indicates the probabilities at the current trial (e.g., [40%, 20%, 20%, 20%]), while P_{t-1} indicates the probabilities at the previous trial (e.g., [25%, 25%, 25%, 25%]). Details on how probabilities are
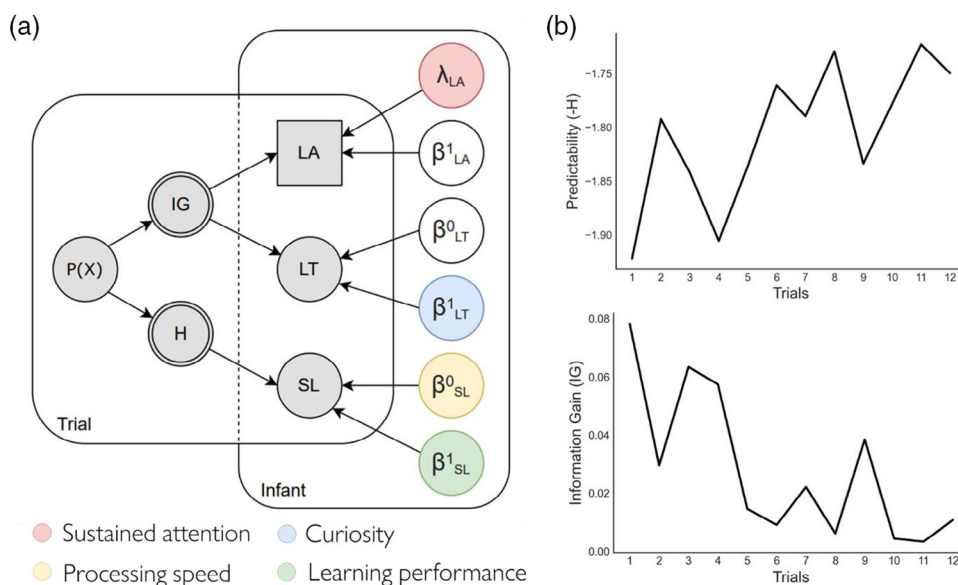
FIGURE 3 The hierarchical Bayesian model. (a) Simplified representation of the model parameters (for a complete representation, see supplementary materials). Information gain and predictability are computed from the probabilities of the target locations P(X), and then related to look-away (LA), looking time to the target (LT) and saccadic latency (SL) through latent parameters. The latent parameters of interest (in colours) are estimated for each infant. (b) The information-theoretic values of predictability and information gain for one example sequence of the visual learning task. Note that participants were presented with multiple sequences of stimuli.

updated on every trial can be found in Poli et al. (2020, 2023) and in the supplementary materials.

Latent parameters. The information theoretic measures of predictability and information gain as well as the measures collected from infants' eye movement behaviour are observable (i.e., their value is known). They can be related to each other to infer latent parameters that have unknown values (Figure 3a).

Previous research showed that infants are faster at looking at predictable events compared to unpredictable events (Kayhan et al., 2019). Accordingly, we assumed a linear relationship between saccadic latency and predictability. In the model, this linear relationship is controlled by two parameters, β_{SL}^0 and β_{SL}^1. The former is the intercept, which indicates the baseline speed of a saccade, while the latter is a correlation coefficient, which indicates how strong the relation is between saccadic latency and predictability. Thus, β_{SL}^0 captures what are infants' reaction times irrespective of the information gain of the stimulus, and offer a good proxy for processing speed. In fact, infants' saccadic latency has been related to processing speed later in life (Dougherty & Haith, 1997), as well as to more positive cognitive outcomes (Egger et al., 2020; Fernald & Marchman, 2012). β_{SL}^1 captures the strength of the correlation between saccadic latency and predictability, where higher values indicate that infants have better learned to predict the location of the target. Hence, β_{SL}^1 can be regarded as a proxy for learning performance.

Stimuli with higher information content have been associated with longer looking times both in adults (Bestmann et al., 2008; O'Reilly et al., 2013) and infants (Poli et al., 2020). Hence, we expect a linear relationship between stimulus informativity and looking time to the target. Also in this case, two parameters β_{LT}^0 and β_{LT}^1 regulate the

relation between the two variables. We did not relate the intercept parameter β_{LT}^0 to any cognitive factor, as we did not have an a priori hypothesis about what aspect of cognition may be related to it, nor has this measure been used in previous literature. Hence, it was not included in our further analyses but only used to ensure proper model fit. The correlation coefficient β_{LT}^1 captures the extent to which infants are sensitive to the information content of a given stimulus. A stronger bias towards informative stimuli results in higher β_{LT}^1 values. Hence, we used β_{LT}^1 as a proxy for curiosity, defined as a bias towards information-rich stimuli (Dubey & Griffiths, 2020; Goupil & Proust, 2023). It is important to stress that curiosity is a multidimensional construct that encompasses motivational and cognitive components (Goupil & Proust, 2023), and is driven by a multitude of factors, such as surprise (Iigaya et al., 2016), information gain (Bennett et al., 2016) and learning progress (Poli et al., 2022). The current paradigm taps into one key aspect of curiosity that received much attention in recent theoretical and experimental research (Poli et al., 2020; Twomey & Westermann, 2018), namely the degree to which infants are sensitive to information. However, it does not encompass every aspect of curious behaviour.

Finally, we used a survival function to estimate how long infants attended to the screen. This was captured by the parameter λ. Infants' looking duration to the screen has often been used as a measure of sustained attention (Colombo & Cheatham, 2006; Wass et al., 2018), and it has also been related to executive control later in life (Hendry et al., 2018). However, looking duration can be affected by other factors, unrelated to sustained attention, such as how interesting or surprising the stimulus is or by increasing fatigue due to the passage of time. Hence, we controlled for information-theoretic stimulus surprise (i.e.,

the negative logarithm of its probability of occurrence, see Poli et al., 2020), information gain, and time when estimating looking durations by adding them as covariates to the regression models. This allowed us to obtain a more refined proxy of sustained attention from looking duration.

We estimated these parameters at two levels, thus following a hierarchical structure. The lower level of analysis is the individual infant. This allowed us to measure whether the estimates of the latent parameters differed across infants. The higher level of analysis is the group level, because even in the presence of individual differences, there should be an average value at the population level. This approach offers the additional advantage of estimating each individual's values *in relation to* the other individuals' values, rather than in isolation. Although this may lead to a less accurate objective estimate of the values of the parameters, according to Stein's paradox (Efron & Morris, 1977), it leads to a more meaningful estimate of the differences across individuals. Hence, the hierarchical structure of the model allowed us to better capture individual differences.

Generalized linear models. As a last step, we used the latent parameters of the hierarchical Bayesian model to predict habituation and dishabituation performance. This step was carried out with generalized linear models. The estimates of the latent parameters during the visual learning paradigm were used as independent variables. It must be noted that the hierarchical Bayesian model estimates a probability distribution for each parameter of each infant, but generalized linear models only work on point-wise estimates. Hence, we inferred the true value of the parameters by minimizing the expected absolute loss function, which is equivalent to computing the median of the posterior distribution of each parameter.

Habituation time and dishabituation score were the dependent variables. Habituation time data was fitted using a lognormal distribution because it was positively skewed. Dishabituation scores were fitted with a quasibinomial distribution, which is an appropriate distribution for proportion data. Age in months was always added as a covariate in the regression models. Additionally, when predicting dishabituation scores, we included whether infants had habituated or not as an interaction term with the cognitive factors. The reason is that infants' looking behaviour to the novel stimulus may differ depending on wether infants had habituated in the first place.

3 | RESULTS

Infants track information gain and predictability. Our preliminary analysis focuses on the group-level relationship between infants' behaviour and information-theoretic values. We replicate the results from Poli et al. (2020) with an extended sample. Specifically, we found a negative correlation between information gain and look-away probability (mean = −1.79, SD = .15, 89%HDI = [−2.04, −1.55]) indicating that infants were more likely to disengage when the stimuli carried less information gain. Moreover, we find a negative correlation between saccadic latency and predictability (mean = −0.18, SD = .09, 89%HDI = [−0.26, 0.01]) indicating faster eye movements

when the environment was more predictable, and a positive correlation between looking time to the target and information gain (mean = 0.06, SD = 0.08, 89%HDI = [−0.08, 0.19]) indicating that infants tended to look longer to stimuli that carried more information. The effects for saccadic latency and looking time are weak at the group level. In the next paragraph, we show that shifting the focus on individual differences reveals that some infants can track changes in predictability and information gain, while others do not.

Bayesian latent parameters capture individual differences in cognitive functioning. We compared the performance of the hierarchical model (with individual and group levels) to the performance of a simpler model (group level only). We found that the goodness of fit of the hierarchical model was higher compared to the simpler model (for details on the model sampling, convergence and comparison, see the supplementary materials). This indicates that the individual differences across infants were present and that they were successfully captured by the hierarchical model. An impression of the individual differences in the parameter estimates is given in Figure 4.

When considering the association between saccadic latency and predictability (i.e., learning performance), we identified 57 infants (40%) with a significant coefficient, as indexed by 89% credible intervals different from zero (where zero indicates lack of an effect). Of these infants, 98% showed a positive coefficient, and 2% showed a negative coefficient. When focusing on the association between looking time and information gain (i.e., curiosity), we identified 31 infants (22%) with a significant coefficient, as indexed by 89% credible intervals that do not include the zero. Of these infants, 84% showed a positive coefficient, and 16% showed a negative coefficient. This indicates that infants were performing the task as expected, but with considerable individual differences. As such, additional evidence is provided that individual differences in task performance can be detected with this model-based approach.

When focusing on the association between look-away and information gain, we identified 140 infants (99%) with a significant positive coefficient, as indexed by 89% credible intervals that do not include zero . However, all infants showed comparable effects, which indicated a lack of individual differences in this specific parameter (see Figure S2). For this reason, we do not analyse this parameter further.

Individual differences in cognitive factors predict habituation and dishabituation. We analysed how z-scored values of processing speed, curiosity, sustained attention, and learning performance related to habituation and dishabituation using two generalized linear models. When focusing on habituation, we found that habituation times were significantly related to individual differences in processing speed (Figure 5a) ($\beta = -0.21$, SE = 0.05, $p < 0.001$), but not to curiosity ($\beta = -0.05$, SE = 0.05, $p = 0.34$), sustained attention ($\beta = 0.09$, SE = 0.05, $p = 0.06$), or learning performance ($\beta = -0.10$, SE = 0.05, $p = 0.06$). Moreover, as expected, infants who habituated showed reduced habituation time compared to the ones who did not ($\beta = -0.27$, SE = 0.10, $p = 0.006$). Age did not have a significant effect in predicting habituation times ($\beta = -0.005$, SE = 0.003, $p = 0.10$).

When analysing dishabituation scores, we found a significant interaction between curiosity and whether infants had habituated (see
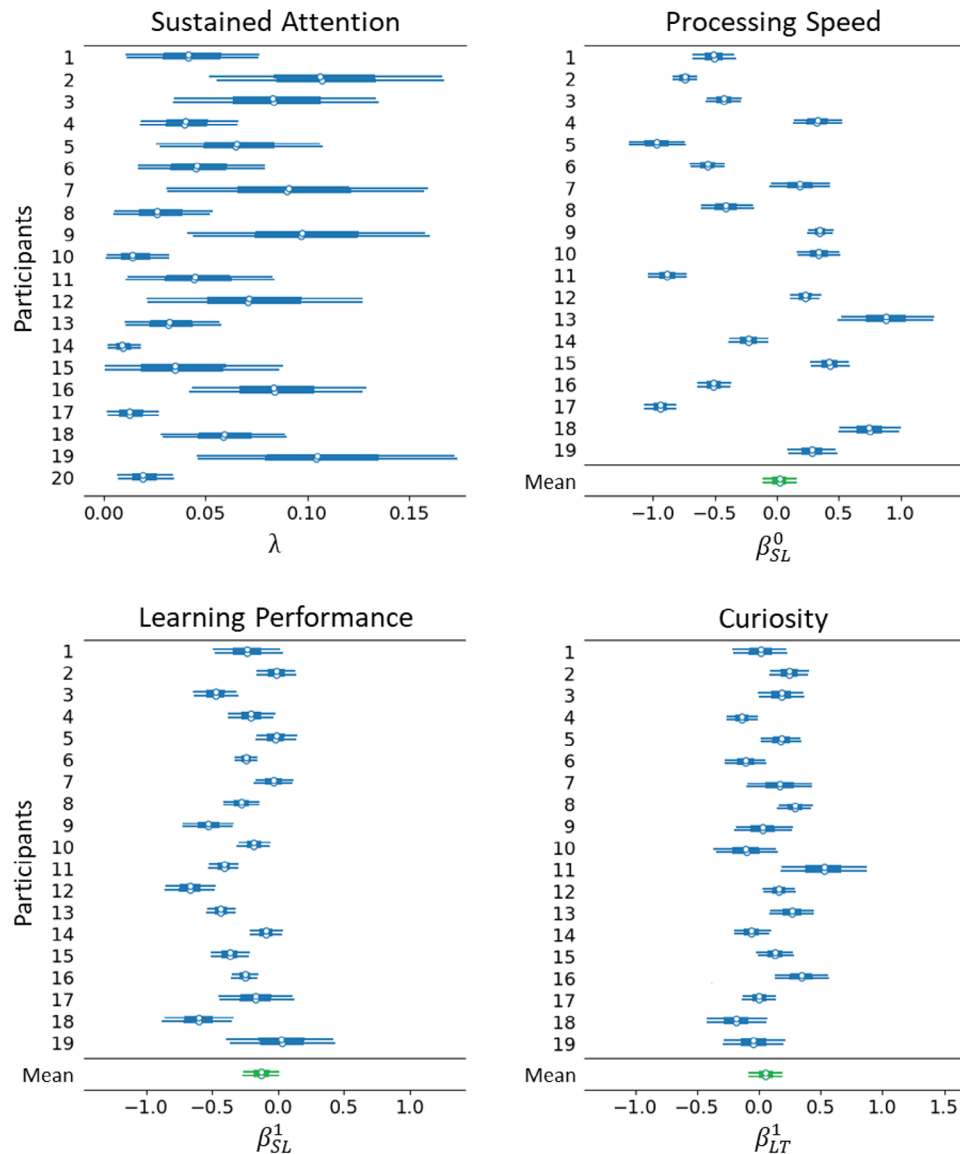
FIGURE 4 The latent parameters of the hierarchical Bayesian model, estimated for each infant. The figure shows a subset of infants sampled randomly from the full sample (19 out of 106). For each infant, the 89% Highest Density Interval (HDI) is represented with thin lines. The standard deviation of the estimates is represented with thick lines. The white points represent the mean values of the estimates. Mean values in green indicate the estimates for the hyperparameters of the model.

Figure 5b) ($\beta = -0.33$, SE $= 0.14$, $p = 0.02$). For low levels of curiosity (standardized scores $= -2$), there was a significant difference in dishabituation scores between infants who habituated and infants who did not ($\Delta\beta = 1.53$, SE $= 0.32$, $p < 0.001$). Specifically, confidence intervals for dishabituation scores were at chance (i.e., they included 0.5) for infants who did not habituate, while scores were above chance for infants who did habituate (Figure 5b). This difference disappeared for high levels of curiosity (standardized scores $= 2$) (Figure 5c) ($\Delta\beta = 0.25$, SE $= 0.29$, $p = 0.40$), indicating above-chance performance for both groups. There were no significant effects of sustained attention ($\beta = -0.10$, SE $= 0.07$, $p = 0.12$), processing speed ($\beta = -0.13$, SE $= 0.07$, $p = 0.08$), learning performance ($\beta = -0.02$, SE $= 0.07$, $p = 0.63$), and age ($\beta = 0.006$, SE $= 0.004$, $p = 0.17$).

4 | DISCUSSION

Although habituation and dishabituation are widely used measures in the field of infant cognition, their underlying mechanisms have been unknown to date. Previous research suggested that habituation and dishabituation might be related to processes as diverse as attention (Blaga & Colombo, 2006), learning (Rose & Feldman, 1997), processing speed (Feldman and Mayes, 1999), or curiosity (Cao et al., 2022). Here, we were able to test these competing hypotheses. First, we tested infants on a habituation task and a visual learning task. From the visual learning task, we subsequently successfully measured individual differences in sustained attention, processing speed, learning performance, and curiosity using a hierarchical Bayesian model. Finally,
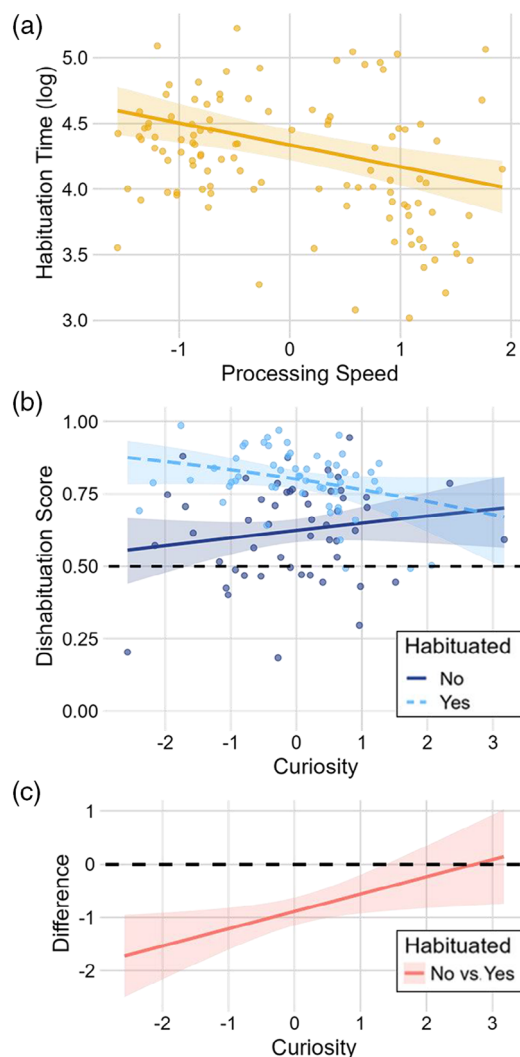
FIGURE 5 Significant predictors of individual differences in habituation and dishabituation performance. The values of the predictors are z-scored. (a) Greater processing speed predicts faster habituation. (b) Greater curiosity predicts higher dishabituation scores, but only for infants who did not habituate. (c) Difference between the marginal means at the two different groups (not habituated vs. habituated) (i.e., contrast analysis). Infants who did not habituate and scored low on curiosity also scored lower on dishabituation performance.

we associated these individual differences in cognitive functioning with habituation and dishabituation, as measured by the habituation task.

We found that habituation times were lower for infants with a higher processing speed. This is in line with the ideas by Feldman and Mayes (1999), who theorized that better (i.e., faster) habituation performance should be explained by individual differences in processing speed. Specifically, they argued that fast information-processers would learn faster and thus habituate faster. Moreover, Fry and Hale (2000) sketched a developmental cascade where processing speed determines working memory capacity (Kail, 1992), with both processing speed and working memory having a direct impact on intelligence

(Kail & Salthouse, 1994). Given the link between faster habituation in infancy and higher intelligence later in life (Kavšek, 2004), the early differences in processing speed that we observe might explain the association between habituation and intelligence. To unravel such underlying mechanisms, a longitudinal follow-up study is needed that investigates intelligence in our sample at a later age.

We did not find evidence for an association between habituation and curiosity (Cao et al., 2022). However, we found an association between curiosity and dishabituation scores, but only for infants who did not habituate. This finding suggests that curiosity might have a preventive effect: Infants who are slower in processing information and fail to habituate still show a preference for a novel stimulus if they score high on curiosity. These results are also informative in relation to violation-of-expectations paradigms. In these paradigms, infants are first familiarized with a certain event (e.g., an object falling) and are then exposed to a surprising event (e.g., an object floating in mid-air). Longer looking times have usually been associated with a reaction of surprise (Stahl & Feigenson, 2015). More recently, they have also been used to measure individual differences in curiosity (Perez & Feigenson, 2021). Although dishabituation scores are related to the aspect of curiosity that we measured (i.e., sensitivity to information), we find that this is true only under specific conditions. Given that results change across tasks and measures, we call for an integrated approach to curiosity, in which multiple indexes derived from different tasks contribute to a more comprehensive assessment of the overarching construct of curiosity.

Focusing on the methodological aspects of the current work, we showed how a model-based approach to the measurement of individual differences allows us to obtain more direct measures of cognitive functioning. In this paper, we did not exploit the full potential of Bayesian estimates, because we reduced probability distributions to point-wise estimates when analysing the relation between individual differences in latent parameters and (dis)habituation. Future work might address this issue using fully-Bayesian approaches. One additional aspect that remains to be tested is whether the individual differences that we find in the current modelling work reflect stable traits or temporary states (Steyer et al., 1999). We tested infants on both tasks on the same day, and individual differences might thus reflect specific states (e.g., fatigue) rather than reliable traits. Although this does not invalidate the results of the current study, it remains to be determined whether these parameters are stable across sessions (test-retest reliability) and across development.

This novel model-based approach offers new avenues to assess differences in diverse cognitive functions in young infants. This approach can also be used with other tasks, as any behavioural, physiological, or neuroimaging data can be modelled in a similar way to infer the value of latent parameters, thus obtaining estimates of individual differences that are more refined than the data themselves. Although noisy and incomplete data might be difficult to use directly, they are still valuable input in hierarchical Bayesian models, as they bring additional information to infer the values of latent parameters. For this reason, longitudinal studies might especially benefit from hierarchical Bayesian models, and approaches like ours can open up new opportunities in the

difficult endeavour of predicting developmental outcomes from early in life.

AUTHOR CONTRIBUTIONS

Francesco Poli: Conceptualization; Investigation; Data Curation; Methodology; Software; Formal Analysis; Writing—Original Draft; Visualization. **Tommaso Ghilardi**: Formal Analysis; Writing—Review & Editing. **Roseriet Beijers**: Supervision; Writing—Review & Editing; Funding acquisition. **Carolina de Weerth**: Supervision; Writing—Review & Editing; Funding acquisition. **Max Hinne**: Supervision; Formal Analysis; Writing—Review & Editing. **Rogier B. Mars**: Conceptualization; Supervision; Writing—Review & Editing; Funding acquisition. **Sabine Hunnius**: Conceptualization; Supervision; Writing—Review & Editing; Funding acquisition.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data, computational models, and statistical analyses scripts are available on OSF: https://osf.io/zux9v/

ORCID

Francesco Poli https://orcid.org/0000-0001-9841-0056
Tommaso Ghilardi https://orcid.org/0000-0003-0855-2814
Roseriet Beijers https://orcid.org/0000-0002-4033-6620
Carolina de Weerth https://orcid.org/0000-0002-0921-1811
Max Hinne https://orcid.org/0000-0002-9279-6725
Rogier B. Mars https://orcid.org/0000-0001-6302-8631
Sabine Hunnius https://orcid.org/0000-0001-5510-4032

REFERENCES

Addyman, C. (2015). *Infant-Habituation-in-Matlab: Version 1 (Version v1)*. Zenodo. https://doi.org/10.5281/zenodo.16627

Bennett, D., Bode, S., Brydevall, M., Warren, H., & Murawski, C. (2016). Intrinsic valuation of information in decision making under uncertainty. *PLOS Computational Biology*, *12*(7), e1005020. https://doi.org/10.1371/journal.pcbi.1005020

Bestmann, S., Harrison, L. M., Blankenburg, F., Mars, R. B., Haggard, P., Friston, K. J., & Rothwell, J. C. (2008). Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Current Biology*, *18*(10), 775–780. https://doi.org/10.1016/j.cub.2008.04.051

Blaga, O. M., & Colombo, J. (2006). Visual processing and infant ocular latencies in the overlap paradigm. *Developmental Psychology*, *42*, 1069–1076. https://doi.org/10.1037/0012-1649.42.6.1069

Bornstein, M. H., & Colombo, J. (2012). Infant cognitive functioning and mental development. In *Early childhood development and later outcome* (pp. 118–147). Cambridge University Press.

Bornstein, M. H., & Sigman, M. D. (1986). Continuity in mental development from infancy. *Child Development*, *57*(2), 251–274.

Cao, A., Raz, G., Saxe, R., & Frank, M. C. (2022). Habituation reflects optimal exploration over noisy perceptual samples. *Topics in Cognitive Science*, *15*(2), 290–302. https://doi.org/10.1111/tops.12631

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol., 34, pp. 283–322). JAI. https://doi.org/10.1016/S0065-2407(06)80010-8

Colombo, J., Richman, W. A., Shaddy, D. J., Follmer Greenhoot, A., & Maikranz, J. M. (2001). Heart rate-defined phases of attention, look duration, and infant performance in the paired-comparison paradigm. *Child Development*, *72*(6), 1605–1616. https://doi.org/10.1111/1467-8624.00368

Cover, T. M., & Thomas, J. A. (1991). Entropy, relative entropy and mutual information. In Schilling D. L. (Ed.) *Elements of information theory*. p. 12–49. John Wiley & Sons Ltd.

Dougherty, T. M., & Haith, M. M. (1997). Infant expectations and reaction time as predictors of childhood speed of processing and IQ. *Developmental Psychology*, *33*, 146–155. https://doi.org/10.1037/0012-1649.33.1.146

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. *Psychological Review*, *127*, 455–476. https://doi.org/10.1037/rev0000175

Efron, B., & Morris, C. (1977). Stein's paradox in statistics. *Scientific American*, *236*(5), 119–127.

Egger, J., Rowland, C. F., & Bergmann, C. (2020). Improving the robustness of infant lexical processing speed measures. *Behavior Research Methods*, *52*(5), 2188–2201. https://doi.org/10.3758/s13428-020-01385-5

Fagan, J. F., Holland, C. R., & Wheeler, K. (2007). The prediction, from infancy, of adult IQ and achievement. *Intelligence*, *35*(3), 225–231. https://doi.org/10.1016/j.intell.2006.07.007

Fagan, J. F., & McGrath, S. K. (1981). Infant recognition memory and later intelligence. *Intelligence*, *5*(2), 121–130. https://doi.org/10.1016/0160-2896(81)90002-7

Feldman, R., & Mayes, L. C. (1999). The cyclic organization of attention during habituation is related to infants' information processing. *Infant Behavior and Development*, *22*(1), 37–49. https://doi.org/10.1016/S0163-6383(99)80004-6

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*(1), 203–222. https://doi.org/10.1111/j.1467-8624.2011.01692.x

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*(1), 1–34. https://doi.org/10.1016/S0301-0511(00)00051-X

Gardner, J. M., Karmel, B. Z., & Flory, M. J. (2003). Arousal modulation of neonatal visual attention: Implications for development. In *Visual information processing* (pp. 125–153). Praeger Publishers/Greenwood Publishing Group.

Ghilardi, T., Poli, F., Meyer, M., Colizoli, O., & Hunnius, S. (2023). *Early roots of information-seeking: Infants predict and generalize the value of information.*

Gottlieb, J., & Oudeyer, P. Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, 19(12), 758–770. https://doi.org/10.1038/s41583-018-0078-0

Goupil, L., & Proust, J. (2023). Curiosity as a metacognitive feeling. *Cognition*, 231, 105325. https://doi.org/10.1016/j.cognition.2022.105325

Hendry, A., Jones, E. J. H., Bedford, R., Gliga, T., Charman, T., Johnson, M. H., Baron-Cohen, S., Blasi, A., Bolton, P., Cheung, H. M. C., Davies, K., Elsabbagh, M., Fernandes, J., Gammer, I., Green, J., Guiraud, J., Lloyd-Fox, S., Liew, M., Maris, H., … the BASIS Team. (2018). Developmental change in look durations predicts later effortful control in toddlers at familial risk for ASD. *Journal of Neurodevelopmental Disorders*, 10(1), 3. https://doi.org/10.1186/s11689-017-9219-4

Iigaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J., & Dayan, P. (2016). The modulation of savouring by prediction error and its effects on choice. *eLife*, 5, e13747. https://doi.org/10.7554/eLife.13747

Kail, R. (1992). Processing speed, speech rate, and memory. *Developmental Psychology*, 28, 899–904. https://doi.org/10.1037/0012-1649.28.5.899

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, 86(2–3), 199–225. https://doi.org/10.1016/0001-6918(94)90003-5

Kavšek, M. (2004). Predicting later IQ from infant visual habituation and dishabituation: A meta-analysis. *Journal of Applied Developmental Psychology*, 25(3), 369–393. https://doi.org/10.1016/j.appdev.2004.04.006

Kayhan, E., Hunnius, S., O'reilly, J. X., & Bekkering, H. (2019). Infants differentially update their internal models of a dynamic environment. *Cognition*, 186, 139–146. https://doi.org/10.1016/j.cognition.2019.02.004

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, 88(3), 449–460. https://doi.org/10.1016/j.neuron.2015.09.010

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, 7(5), e36399. https://doi.org/10.1371/journal.pone.0036399

McCall, R. B. (1994). What process mediates predictions of childhood IQ from infant habituation and recognition memory? Speculations on the roles of inhibition and rate of information processing. *Intelligence*, 18(2), 107–125. https://doi.org/10.1016/0160-2896(94)90022-1

O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), E3660–E3669. https://doi.org/10.1073/pnas.1305373110

Perez, J., & Feigenson, L. (2021). Stable individual differences in infants' responses to violations of intuitive physics. *Proceedings of the National Academy of Sciences of the United States of America*, 118(27), e2103805118. https://doi.org/10.1073/pnas.2103805118

Poli, F., Ghilardi, T., Mars, R. B., Hinne, M., & Hunnius, S. (2023). Eight-month-old infants meta-learn by downweighting irrelevant evidence. *Open Mind*, 7, 141–155. https://doi.org/10.1162/opmi_a_00079

Poli, F., Meyer, M., Mars, R. B., & Hunnius, S. (2022). Contributions of expected learning progress and perceptual novelty to curiosity-driven exploration. *Cognition*, 225, 105119. https://doi.org/10.1016/j.cognition.2022.105119

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science Advances*, 6(39), 5053–5076. https://doi.org/10.1126/sciadv.abb5053

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., Coppola, G., Geyer, M. A., Glanzman, D. L., Marsland, S., McSweeney, F. K., Wilson, D. A., Wu, C. F., & Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, 92(2), 135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Rose, S. A., & Feldman, J. F. (1997). Memory and speed: Their role in the relation of infant information processing to later IQ. *Child Development*, 68(4), 630–641. https://doi.org/10.2307/1132115

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, 348(6230), 91–94. https://doi.org/10.1126/science.aaa3799

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state–trait theory and research in personality and individual differences. *European Journal of Personality*, 13(5), 389–408. https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5⟨389::AID-PER361⟩3.0.CO;2-A

Twomey, K. E., & Westermann, G. (2018). Curiosity-based learning in infants: A neurocomputational approach. *Developmental Science*, 21(4), e12629. https://doi.org/10.1111/desc.12629

Valentine, C. W. (1913). Colour perception and preferences of an infant at three months. *Section on Ophthalmology*, 18, 689–690.

Wass, S. V., Clackson, K., Georgieva, S. D., Brightman, L., Nutbrown, R., & Leong, V. (2018). Infants' visual sustained attention is higher during joint play than solo play: Is this due to increased endogenous attention control or exogenous stimulus capture? *Developmental Science*, 21(6), e12667. https://doi.org/10.1111/desc.12667

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.