DEVELOPMENTAL NEUROSCIENCE

Infants tailor their attention to maximize learning

F. Poli[1]*, G. Serino[1,2], R. B. Mars[1,3], S. Hunnius[1]

Infants' remarkable learning abilities allow them to rapidly acquire many complex skills. It has been suggested that infants achieve this learning by optimally allocating their attention to relevant stimuli in the environment, but the underlying mechanisms remain poorly understood. Here, we modeled infants' looking behavior during a learning task through an ideal learner that quantified the informational structure of environmental stimuli. We show that saccadic latencies, looking time, and time spent engaged with a stimulus sequence are explained by the properties of the learning environments, including the level of surprise of the stimulus, overall predictability of the environment, and progress in learning the environmental structure. These findings reveal the factors that shape infants' advanced learning, emphasizing their predisposition to seek out stimuli that maximize learning.

INTRODUCTION

By the moment they are born, infants have already learned astonishingly much about their environment: Newborns can discriminate their native language from an unfamiliar one (*1*) and recognize stimuli to which they were exposed while in the womb (*2*). Throughout the weeks and months that follow, infants continue to learn about the physical and social world around them at a breathtaking rate (*3*). Infants thus must have extremely flexible and sophisticated learning skills, but unexpectedly, little is known about the learning strategies that enable them to learn so much so quickly.

Previous research suggested that infants' learning might benefit from biases that guide their attention to relevant environmental stimuli. For instance, it has been proposed that infants direct their attention preferably to stimuli that are maximally novel (*4*) or that are neither too predictable nor too surprising (*5*, *6*). For example, Kidd *et al.* (*5*) presented infants with sequences of visual stimuli and quantified the complexity of each stimulus in terms of information-theoretic surprise. They found that infants were more likely to look away from the screen when stimuli were either too simple or too complex, while they were more likely to look at stimuli that had an intermediate level of complexity. However, others have proposed that organisms search for situations that are maximally unpredictable (*7*). In accordance with this, it has been found that infants avoid redundant information and rather prefer variable stimuli (*8*).

Simple strategies such as surprise maximization or redundancy avoidance might aid learning, as they allow agents to avoid spending time on challenges that have already been mastered. However, they have two major shortcomings. First, an organism who only seeks out stimuli with maximal novelty or unpredictability would risk getting stuck with too difficult tasks or in overly complex situations in which no regularities can be detected and nothing is learned (*9*). Second, it has been shown that the same level of surprise (or novelty) can have a different utility for learning depending on environmental volatility (*10*), task relevance (*11*), prior beliefs (*12*), and many other factors (*13*). A bias that focuses on surprise irrespective of any other environmental context is thus suboptimal (*14*).

Recently, researchers in the field of developmental robotics have proposed an alternative learning strategy that—rather than biasing the organism toward certain situations or stimuli—fosters learning itself (*9*, *15*). This strategy entails the maximization of the learning progress: It motivates the agent to explore situations that are initially unknown and to keep focusing on them only if they offer a learning opportunity. A learner aiming to maximize learning progress will not focus on stimuli that it can already predict (where prediction error is minimized) or stimuli that are completely unpredictable (where prediction error is maximized). Rather, the learner will focus on situations that offer an information gain, as long as they do so (*9*). Similarly, in the field of artificial intelligence and machine learning, newly developed curiosity-driven algorithms (*16*, *17*) challenged the most common heuristic exploration methods. Curiosity-driven algorithms drive artificial agents toward the maximization of the learning progress, leading to more efficient learning and superior learning outcomes (*18*). Here, we tested whether infants rely on similar learning strategies. Specifically, we hypothesized that infants as young as 8 months of age preferably attend to stimuli that offer the possibility to learn rather than simply to the most novel or unpredictable patterns.

RESULTS

General approach

We compared infants' behavioral performance, as indexed by eye-tracking data, to the performance of an ideal learner model. This process involved three main steps. The first step concerned the collection and processing of infant looking data on their learning from a sequence of visual stimuli. The second step entailed building an ideal learner model, feeding it with the same stimuli that infants saw and obtaining output estimates. These can be considered as parallel steps. The last step consisted of relating the infants' data to the estimates of the ideal learner model. If the model and the infants process the same environmental features, their performance should be correlated.

Quantifying infants' performance and the ideal learner model's performance

In an eye-tracking experiment, we presented 50 infants with sequences of cue-target trials (Fig. 1). In each sequence, the cue consisted of a simple shape appearing in the middle of the screen. The target was the same shape reappearing in one of four screen quadrants around

[1]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands. [2]Centre for Brain and Cognitive Development, Birkbeck, University of London, London, UK. [3]Wellcome Centre for Integrative Neuroimaging, Centre for Functional MRI of the Brain (FMRIB), Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford, UK.
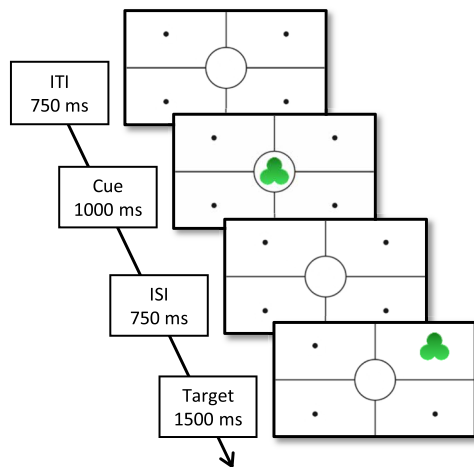*Corresponding author. Email: f.poli@donders.ru.nl

Fig. 1. A trial of the learning task. The cue looms in the center of the screen. After 750 ms, the target rotates in one of the four quadrants. ITI, intertrial interval; ISI, interstimulus interval.

the cue location. The shape (e.g., a trefoil or a star) was the same across all trials of the sequence but changed across sequences. Some sequences (i.e., 4) were deterministic, as the target always appeared in the same location. Most of the sequences (i.e., 12) were probabilistic: The target could appear in any location, but one location was more likely than the others (see movie S1 for an example). Infants could thus learn to predict the most likely target location of each sequence.

For each trial, we measured three different variables to answer the following three questions. First, we asked which characteristics of a stimulus determined whether infants keep looking at it or look away. In infant habituation studies, looking away from the screen has been taken as a measure of lack of interest in the stimulus (*19*). More specifically, it can be considered an active decision made by the infant to stop allocating cognitive resources to the stimulus (*20*). We hypothesized that the main factor determining infants' looking away from the screen was the learning progress that a stimulus offered. Second, we examined whether infants' learning strategy was successful. When infants have an efficient model of the statistical environment, they become faster in directing their gaze toward predictable targets but slower in looking at events that are surprising or unpredictable (*21, 22*). Last, we tested whether the learning progress a trial offered affected infants' trial-by-trial information processing. When the offered learning progress is high, infants need more time to process the stimuli, and this should be reflected in an increase of looking time over the targets, as has already been shown in adults (*23*).

The sequences of cue-target associations seen by infants were also given as input to an ideal learner model (see Fig. 2). The model kept track of the probability with which targets appeared in each location and updated the probabilities at every trial. Using information theory (*24*), we then derived from the model for each trial an estimate of (i) how surprising the trial was (surprise), (ii) how predictable the sequence was at that moment in time (negative entropy), and (iii) how much learning progress the trial offered (information gain). Each of these factors might play a crucial role in determining infants' information-processing and attention during learning, and here, we were able to pit them against each other.

Relating infants' data to the estimates of the ideal learner model

We first used a generalized additive model to examine what factors predicted whether infants looked away from the sequence. Additive models are more flexible than linear models, as they allow for any type of nonlinear relation between dependent and independent variables. As reported in Fig. 3 (see also tables S1 to S3), we found that the strongest predictor of infants' looking away was the learning progress offered by the stimulus [Wald χ^2 = 48.87, edf (effective degrees of freedom) = 4, $P < 0.001$]. When the learning progress was high, infants were more likely to keep looking at the sequence. In turn, if a stimulus did not allow learning to proceed fast, then infants were more likely to look away. Consistent with previous research (*5, 22*), stimulus surprise (Wald χ^2 = 19.38, edf = 3, $P < 0.001$) and sequence predictability (Wald χ^2 = 21.87, edf = 1, $P < 0.001$) also played a role in determining infants' allocation of attention. Infants were more likely to look away if the sequence of stimuli was too predictable or too unpredictable (Fig. 3B). Moreover, they were more likely to look away if stimuli were more surprising (Fig. 3A). However, additional analyses showed that these effects were much weaker (five and three times, respectively) than the impact that learning progress had on the probability of looking away (table S3). These results reveal that infants are equipped with more advanced learning skills than known to date: From early in their lives, they allocate cognitive resources depending on the learning progress that a stimulus offers.

Second, we examined whether this learning strategy was effective. When infants learn statistical regularities, they get faster in anticipating predictable events and slower in directing their gaze to surprising events (*21*) or unpredictable patterns (*22*). As reported in Fig. 4A and table S4, this was the case in our experiment, and learning was indeed successful.

Last, we show that infants are engaged in a trial-by-trial updating process to constantly refine their model of the environment. A stimulus that offers the opportunity to learn also has a greater impact on how much the infant's model of the statistical environment must change. In other words, when infants' learning is proceeding at a faster pace, they need to update their model more. To examine whether infants were engaging in this trial-by-trial updating process, we performed a generalized linear regression with looking time on the target stimulus as dependent variable. As depicted in Fig. 4B and table S5, infants spent more time looking at a stimulus when it required greater updating, even after controlling for stimulus surprise, sequence predictability, and the passage of time. Conversely, surprise and predictability were unrelated to looking time. These results show not only that infants seek out situations that require maximal model updating but also that these situations imply and could lead to more learning. This would be in line with previous theoretical work (*14*), where better learning emerged from an artificial agent seeking out situations that required greatest adaptation.

DISCUSSION

We tested infants on a visual learning task, monitored their performance with eye tracking, and compared their behavior to the performance of an ideal learner model. We found evidence favoring the idea that infants build predictive models and constantly update their predictions based on what they observe in their environment. Moreover, they actively use the information about learning progress to allocate their cognitive resources strategically in a way that
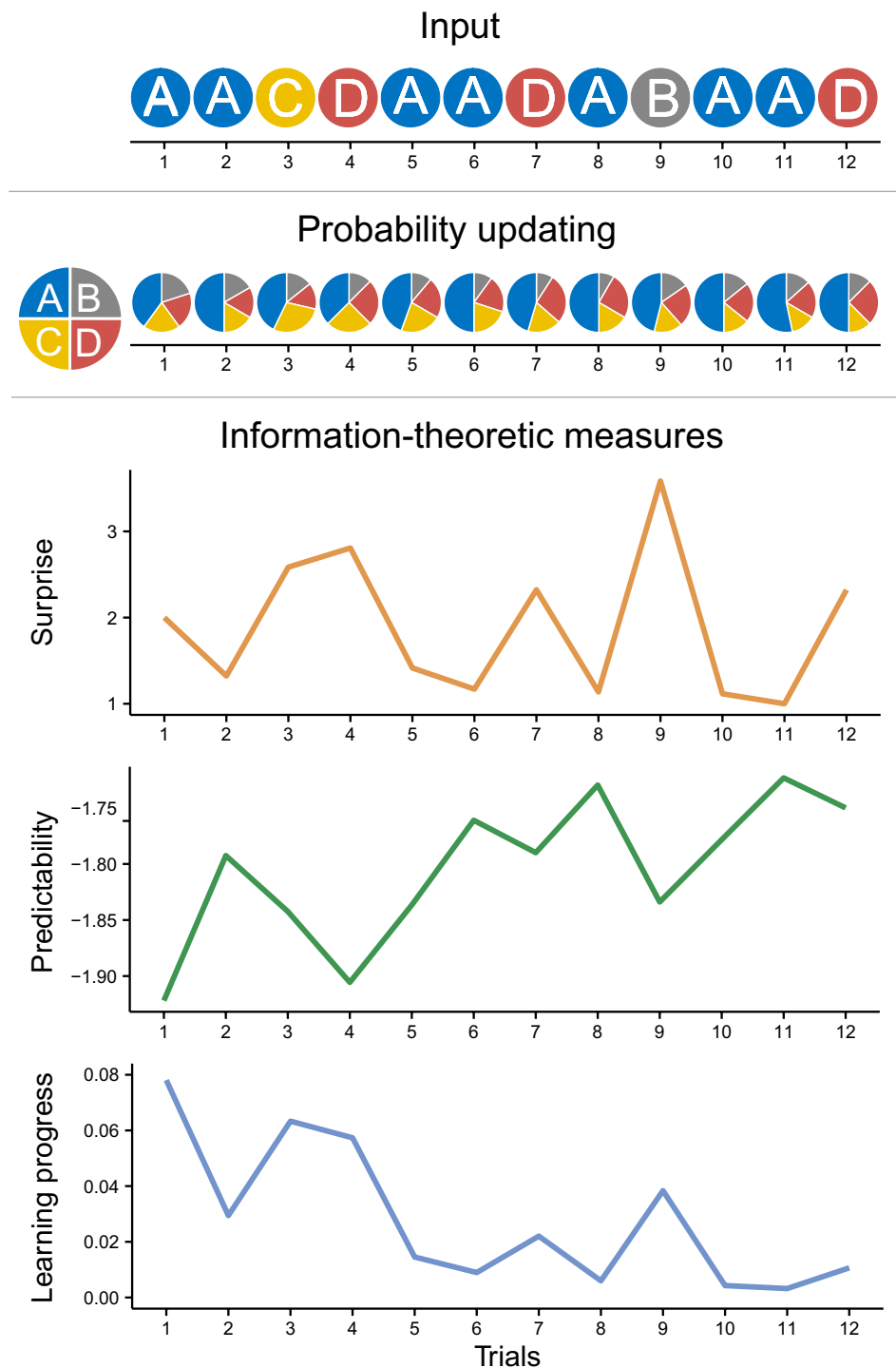
Fig. 2. The ideal learner model. At the beginning of a sequence, the ideal learner has no expectations on where the target will appear. The four locations (A, B, C, and D) are equally probable. At each following trial, probabilities are updated depending on the observed target location. The ideal learner then estimates surprise, predictability, and learning progress.

maximizes learning. These results challenge the idea that infants' exploration of the environment is solely biased by stimulus surprise or unpredictability. Rather, the learning progress framework offers a more extended explanation for why developing organisms might learn so much so quickly: They are driven to search for stimuli that maximize the learning progress.

The current findings are in line with the predictive processing framework, which holds that the brain is essentially a prediction engine (25, 26). From what we expect to see in a visual scene to what we expect another person to do, our brain is constantly trying to predict how events unfold. In other words, the brain is engaged in a continuous attempt to reduce uncertainty. An organism that preferably
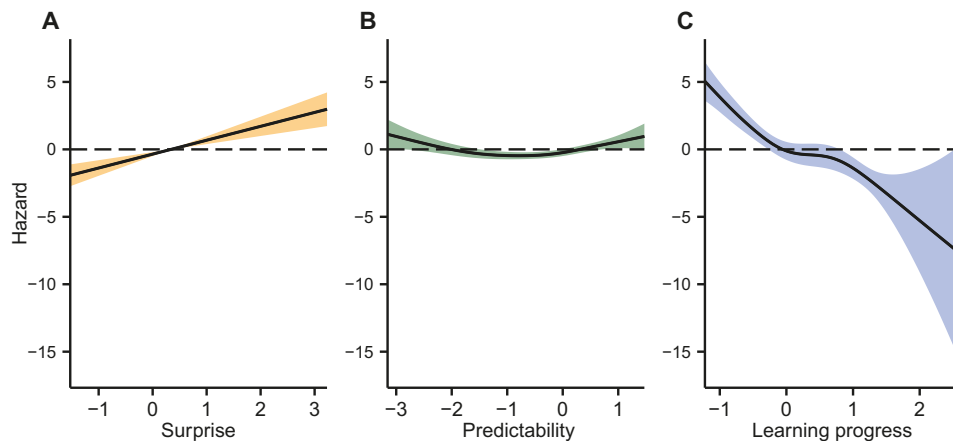
Fig. 3. Results of the generalized additive model. Surprise (**A**), predictability (**B**), and learning progress (**C**) independently affected infant's probability of looking away. Probability of looking away is expressed in terms of hazard. The effect of time was controlled while estimating the effects of the three information-theoretic variables. Model comparison showed that the full model with all regressors had a statistically better model fit than any other partial model (table S1). A follow-up generalized linear model (GLM) returned beta coefficients and effect sizes for each regressor (table S2).
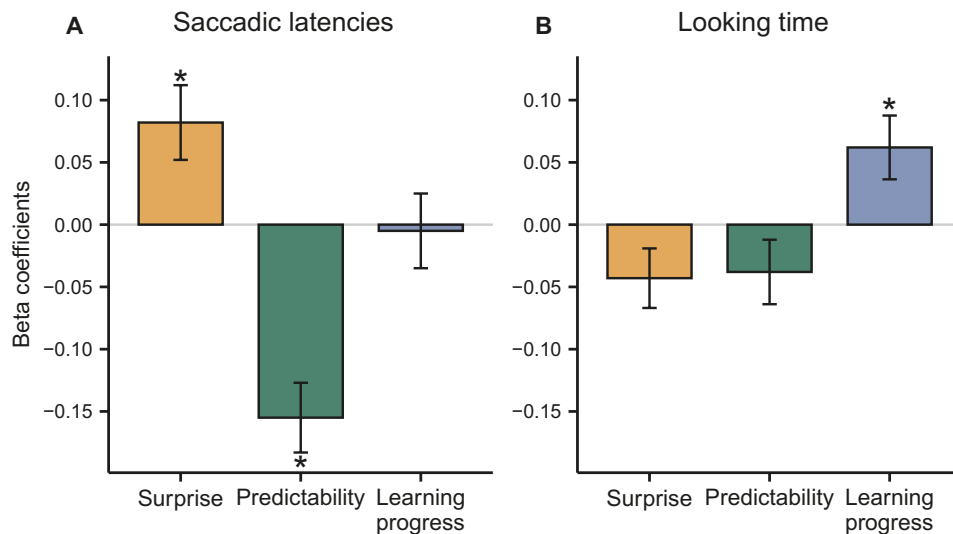

Fig. 4. Results of the GLMs for saccadic latency and looking time. Saccadic latency correlates with surprise and predictability (**A**), while looking time correlates with the learning progress (**B**). On the y axis, the size of beta coefficients for the three variables of interest (surprise, predictability, and learning progress) is given. Time was added as covariate to both models. Its effect was small but significant (see tables S4 and S5). Also, saccadic latency was added as covariate when estimating the effects of information-theoretic constructs on looking time (see Materials and Methods). *$P < 0.05$.

engages in activities that maximize learning will quickly minimize environmental unpredictability and will thus maximize its predictive power at a faster pace. Moreover, the avoidance of highly surprising stimuli as we found it can be explained within the predictive coding framework. When information is inconsistent with its current model of the environment, an agent has two viable policies (26). One is to change its model to fit the new information and the other is to disregard the new information. The latter alternative might explain why infants look away when information is too surprising.

The effect of learning progress on infants' attention might also help to reconcile the apparent contradiction between infants' familiarity and novelty preference (27). The idea that infant attentional preference is dynamic is well established in developmental research (28), and yet, it is still unclear why some studies report infants' preference for novel stimuli, while others find a preference for familiar

ones, even across very similar paradigms (29). Our results suggest that infants keep focusing on familiar stimuli as long as they offer learning progress (which explains the familiarity preference) but they switch to new stimuli when their learning progress drops (which explains the novelty preference). This is in line with previous work proposing that habituation emerges from the opponent, complementary processes of hippocampal selective inhibition and cortical long-term potentiation (30).

In the current paper, we showed that infants tailor their attention to maximize learning, but we remain agnostic with respect to the mechanisms that support this behavior. A likely possibility is that infants find gaining new information intrinsically rewarding (20). In this perspective, curiosity would be the internal motivation that drives infants' attention and learning. However, feedback and external rewards might also play a role (31), as learning might be

driven by the rewards obtained (or lack of rewards) following an action. Curiosity seems to activate similar brain areas as primary rewards, such as food (*32*), and dopaminergic neurons have been found to fire both for extrinsic and informational rewards (*33*). Hence, these two mechanisms might be intertwined (*34*). Some authors have suggested that analogous neural processes might account for reward-based decision making and other types of decision strategies (*35*). This hypothesis awaits further verification.

MATERIALS AND METHODS
Participants
Fifty 8-month-old infants (M = 8 m0d, SD = 12 d, 25 females) were recruited for the study from a database of volunteer families. Infants had to carry out at least 20 trials to be included in the analysis. Six infants failed to reach this threshold and were thus excluded from the analysis. One additional infant was excluded due to a MATLAB crash. The final sample consisted of 43 infants (M = 7 m29d, SD = 12 d, 24 females). Families received a book or 10 Euros for their participation. The local ethics review board approved the study (Ethical approval number: ECSW2017-3001-470), and the institutional review board guidelines were followed.

Experimental design
The stimuli consisted of eight shapes (star, heart, trefoil, triangle, crescent, rhombus, octagon, and cross). The shapes were presented as cues and as targets within a frame (see Fig. 1). Vertical and horizontal lines divided the screen in four locations (the target locations), and a central circular area defined the cue location. Cues were looming in the center of the screen. Targets were rotating in one of the four locations around the center of the screen (see movie S1).

Infants were presented with 16 sequences of cue-target couplings. Each sequence was composed of 15 trials. Every trial consisted of a cue phase (1000 ms), an interstimulus interval (750 ms), a target phase (1500 ms), and an intertrial interval (750 ms). In the cue phase, a simple shape (e.g., a star) appeared in the middle of the screen. In the target phase, the same shape appeared in one of four quadrants.

Every sequence consisted of one only type of shape (e.g., only stars). The shape location during the target phase was systematically manipulated. In four of the 16 sequences, the target appeared in the same location in 100% of the trials; in six sequences, the target appeared in one quadrant 80% of the times and the remaining 20% of the times its appearance was distributed over the other three locations; in the remaining six sequences, the target appeared in one quadrant 60% of the times and the remaining 40% of the times its appearance was distributed over the other three locations. Hence, the cue was always predictive of the target location, but its degree of predictability varied.

We created the sequences in MATLAB. First, 16 sequences were sampled pseudo-randomly, with the probabilities specified above as only constraint. Then, the sequences were concatenated. To check that the target location could be predicted only by relying on cue-target conditional probabilities, we fed the result of the sampling into a machine learning random forest classifier. If the classifier was able to reliably predict the target location with no information about the cue-target conditional probabilities (e.g., it successfully predicted the target location at trial N only based on target location at trial N-1), then the entire process was repeated and new sequences were sampled. The sequences obtained through this procedure were pre-

sented to all participants. The only element of the sequences that was pseudo-randomized across participants was the exact location of the target. For example, participants 1 and 2 would see the same deterministic sequence, but for participant 1, the target always appeared in the upper left corner, while for participant 2, the target always appeared in the bottom right corner. In this way, every participant was exposed to the same statistical regularities, but we were able to control for other biases (e.g., toward the left side of the screen), that might have influenced participants' performance. Averaging across all trials of all sequences, each of the four target locations had the same probability of showing the target (25%).

Throughout the presentation of the stimuli, background music was played to increase overall attention toward the screen.

Procedure
The study was conducted in a quiet room without daylight. Infants were seated in a baby seat placed on their caregiver's lap, 60 to 65 cm from a 23″ monitor. Their looking behavior was recorded using a Tobii X300 eye-tracker (www.tobii.com). Infants' behavior was monitored through an external video camera. Stimulus presentation and data collection were carried out using MATLAB Psychtoolbox. For every infant, the eye-tracker was calibrated with a five-point calibration sequence. If more than two points were not accurately calibrated, the calibration was repeated for a maximum of three times.

The sequences were played one after the other. When the infant looked away from the screen for 1 s or more, the sequence was stopped. When the infant looked back to the screen, the following sequence was played. The experiment lasted until the infant had watched all 16 sequences or became fussy. Parents were instructed not to interact with their child, unless infants sought their attention and, even in that case, not to try to bring infants' attention back to the screen.

Data processing
Raw eye-tracking data were first processed through identification by 2-means clustering (I2MC) (*36*). Settings for the I2MC algorithm were the following: interpolation window of 100 ms; interpolation edge of 6.7 ms; clustering window size of 200 ms; downsampling was set to 150, 60, and 30 Hz; window step size of 20 ms; clustering-weight cutoff of 2 SDs above the mean; merge fixation distance of 0.7°; merge fixation time of 40 ms; and minimum fixation duration of 40 ms. It has been shown that, when sampling at 300 Hz, these settings make I2MC very robust to high-noise infant data (*36*). The output of I2MC is a list of fixation points, each consisting of x-y coordinates (expressed in pixels) and a timestamp (expressed in milliseconds).

Areas of interest of 400 × 400 pixels were then delineated around the four target locations and the central cue location. Saccadic latencies, looking times to the targets, and look-away trials were extracted using MATLAB. These variables were standardized for every individual participant by computing z-scores using each participant's mean and SD in lieu of the group-level mean and SD.

Statistical analysis
Look-aways
We examined what factors influenced infants' probability of looking away from a certain sequence. To do so, we used additive Cox models with time-varying covariates. This type of model allowed us to explore any kind of relationship between independent and dependent variables and not just linear relationships. It also allowed

us to analyze truncated data such as look-aways, which violated the assumptions of the more common generalized linear model (GLM). We fitted the models using the R-package "mgcv."

First, we performed a model comparison procedure to select the model with the highest goodness of fit. The aim of the model comparison is to identify which statistical model among the ones that are available better explains the pattern of the behavioral data. To score the goodness of fit of each model, we used Akaike's information criterion (AIC). However, AIC ignores uncertainty related to smoothing parameters, which makes larger models more likely to fit better. We solve this problem as suggested by Wood *et al.* (*37*). Specifically, we compute the conditional AIC in the conventional way (*38*) with an additive correction that accounts for the uncertainty of the smoothing parameters. This correction makes complex models less likely to win over simpler ones.

A common way of comparing two models is to check the difference between their AIC. Here, ΔAIC is computed as the difference between the AIC of a given model and the AIC of the best model. Hence, the higher the ΔAIC is, the worse the model is.

The result of model selection is reported in table S1. The winning model had surprise, predictability, learning progress, and time as covariates and subjects as random factor. Time was expressed in two ways. First, in terms of trial number within a certain sequence (sequence-wise time). Second, in terms of overall number of trials seen during the task (task-wise time). The parameters of the winning model are reported in table S2. All the independent variables had a significant effect on the probability of looking away.

Additive models provide the potential for better fits to data than purely parametric models but arguably with some loss of interpretability: The effect of additive parameters cannot be quantified as clearly as the effect of β parameters. Hence, we fitted another model where we specified the relation between independent and dependent variables, instead of leaving it unspecified. This allows us to obtain beta coefficients and effect sizes. Given the results of the additive model, we specify a linear effect of surprise and learning progress and a quadratic effect of predictability. The results confirm the effects found with the additive model and are reported in table S3. As in Kidd *et al.* (*5*), and to allow direct comparison across studies, we used $e^{|\beta|}$ as a measure of the effect size. Learning progress shows the strongest effect size ($e^{|\beta|} = 7.02$), followed by surprise ($e^{|\beta|} = 2.44$) and predictability ($e^{|\beta|} = 1.27$).

Saccadic latency and looking time

Since the distribution of both saccadic latency and looking time to the target was not normal [as is common for reaction time data, see (*39*)], we used GLMs rather than a linear model. GLMs allow the specification of the distribution of the data, leading to a better model fit and respecting the assumptions of linear regression. Specifically, we used a Cullen and Frey graph to check the distribution type that most closely resembled the ones of our data. We did so via bootstrapping 500 values from the distribution of each dependent variable. This method showed that saccadic latencies and looking time were distributed following a logistic distribution rather than a normal distribution. The logistic distribution is similar to the normal distribution but has heavier tails. The models were fitted in R using the GAMLSS package.

First, we estimated the effects of the information-theoretic measures on saccadic latencies. Time was added as a covariate, as saccadic latencies might decrease as a function of time just because of a practice effect or familiarity with the task. Participants were added

as a random factor to control for interindividual differences. As reported in table S4, the results show a significant effect of surprise and predictability. The selected model fitted better than a null model with no regressors (ΔAIC = 119). It also fitted better than a more common linear model, which assumes normally distributed data (ΔAIC = 147). Last, to make sure that the correlation between information-theoretic values would not hinder the estimation of beta coefficients, we computed the variance inflation factor (VIF) for every independent variable. When the VIF is above 5, there might be a problem with multicollinearity. However, the VIFs were all below 5 (2.28 for surprise, 2.45 for predictability, 2.78 for learning progress, and 1.04 for time).

A similar model was fitted for fixation times. The only difference was that, in addition to the information-theoretic measures and time, saccadic latencies were also added as covariate. Given that saccadic latencies and fixation times were related ($r = -0.59$), in this way, we estimated the relationship between information-theoretic measures and fixation time controlling for fluctuations in saccadic latencies. The results are reported in table S5. The selected model fitted better than a null model (ΔAIC = 59) and also better than a simple linear model (ΔAIC = 248). Last, also in this model, the VIFs were all below 5 (2.38 for surprise, 2.49 for predictability, 2.81 for learning progress, 1.05 for saccadic latencies, and 1.04 for time).

Ideal learner model

In the current study, we expected infants to keep track of the probabilities with which targets appeared in the four quadrants, update these probabilities at each trial, track the level of surprise of each event, the level of predictability of the sequence at each trial, and the amount of learning progress that the trial offered. Following previous literature (*23*, *40*), we developed an ideal learner model that performs the same computations.

The model is presented with a set of events *x*. An event is, for example, the target appearing in the upper left corner of the screen. The events followed each other until the sequence ended (or the infant looked away). The last event of a sequence, which also coincides with the length of the sequence, is named *j*, and the sequence can thus be denoted by $X^j = \{x^1, \ldots, x^j\}$. The first goal of the model is to estimate the probability with which a certain event *x* will occur. Given that the target can appear in one of four possible locations *k*, the distribution of probabilities can be parameterized by the random vector $p = [p_1, \ldots, p_k]$, where p_k is the probability of the target appearing in the *k*th location. In our specific case, the target locations are four, and thus, $p = [p_1, p_2, p_3, p_4]$. The ideal learner treats $p_{1:4}$ as parameters that must be estimated trial by trial given X^j. In other words, given the past events up until the current trial, the ideal learner will estimate the probabilities with which the target will appear in any of the four possible target locations.

At the very beginning of each sequence, the ideal learner expects the target to appear in one of the four target locations with equal probability. This is expressed here as a prior Dirichlet distribution

$$P(p \,|\, \alpha) = \mathrm{Dir}(p; \alpha_k) \qquad (1)$$

where all elements of α are equal to one, α = [1,1,1,1]. In this case, the parameter α determines prior expectations. If there is an imbalance between the values of α (e. g., α = [100,1,1,1]), this means that the model is biased into thinking that the target is more likely to appear in the one location (p_1 in the example) rather than the others.

Conversely, when the numbers are equal, the ideal learner has no biases toward any location. Moreover, high numbers indicate that the model has strong expectations, while low numbers indicate that the model will quickly change its expectations when presented with new evidence. Thus, specifying $\alpha = [1,1,1,1]$, we are defining a weak uniform prior distribution. In other words, the model has no bias toward any location but is ready to change these expectations if presented with contradicting evidence.

At every trial, the prior distribution is updated given the observation of the new event x from the set X^j. The posterior distribution of such update is given by

$$P(p|X^j, \alpha) = \text{Dir}(p^j; n^j_k + \alpha_k) \qquad (2)$$

where n^j_k refers to the number of outcomes of type k observed up until the trial j. As a practical example, imagine that, at trial 1, the model observes a target in the location 1 (i.e., [1, 0, 0, 0]). The values of α will be updated with the evidence accumulated, thus moving from [1, 1, 1, 1] to [2, 1, 1, 1]. This implies that now it is slightly more likely to see the target in location 1 than in any of the other locations. Specifically, the probability of the target appearing in any location can be computed from the posterior distribution $P(p|X^j, \alpha)$ in the following fashion

$$p(x^j = k|X^{j-1}, \alpha) = \frac{n^{j-1}_k + 1}{j - 1 + K} \qquad (3)$$

In words, how likely the target is to appear in a certain corner is given by the total number of times it appeared in that corner, plus one (the value of α), divided by the total number of observations, plus 4 (the sum of the values of α). This updating rule implies that as evidence accumulates, new evidence will weigh less. Given that our sequences are stationary (i.e., the most likely location does not change within the same sequence), this assumption is justified for the current task.

At every trial j, the posterior Dirichlet distribution of trial $j - 1$ becomes the new prior distribution. The new prior is updated using (2) and the probabilities estimates are computed using (3). When infants look away and a new sequence is played, the prior is set back to (1). This means that we assume that when infants start looking to a new sequence, they consider it as independent of the previous ones. Previous research in adults demonstrated the suitability of this assumption (24).

The ideal learner model uses information theory (24) to compute the surprise of each event, the predictability of the sequence at each trial, and the learning progress at each trial. Surprise is quantified in terms of Shannon Information, I

$$I(x^j = k) = -log_2 p(x^j = k|X^{j-1}, \alpha) \qquad (4)$$

where $p(x^j = k)$ is the probability that an event x (i.e., the appearance of the target) will occur in a given location k (e.g., the upper left corner). This probability depends on the prior α and on the evidence accumulated on the previous trials, X^{j-1}. By taking the negative logarithm of a probability, events that are highly probable will have low levels of surprise, while low-probability events will have a high level of surprise.

Predictability is quantified it terms of negative entropy, $-H$

$$-H(p^j) = \Sigma^K_{k=1} p(x^j = k | X^j, \alpha) \, log_2 p(x^j = k | X^j, \alpha) \qquad (5)$$

Note that, different from surprise, here, predictability is estimated considering also the event j, and not just up to $j - 1$. This formula was applied when relating predictability to infants' looking away and looking time, as they have the information relative to trial j when they decide whether to look away and when they look at the target of trial j. However, saccadic latencies do not depend on X^j but rather on X^{j-1}, as when planning a saccade toward the target of trial j, the target has not appeared yet. Hence, a formula slightly different from (5) was used when relating predictability to saccadic latencies, in which X^j was replaced by X^{j-1}.

Last, the learning progress is quantified in terms of Kullback-Leibler Divergence (or information gain), D_{KL}

$$D_{KL}(p^j \| p^{j-1}) = \Sigma^K_{k=1} p(x^j = k|X^j, \alpha) \, log_2 \frac{(x^j = k|X^j, \alpha)}{(x^j = k|X^{j-1}, \alpha)} \qquad (6)$$

where p^j is the estimate of the parameters $p_{1:k}$ at trial j, while p^{j-1} is the estimate of the parameters $p_{1:k}$ that was performed on the previous trial $j - 1$. Learning progress has been defined as the reduction in the error of an agent's prediction (15). D_{KL} is the divergence between a weighted average of prediction error at trial j and a weighted average of prediction error at trial $j - 1$, and hence, it is a suitable way to model learning progress in this task.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/39/eabb5053/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES
1. C. M. Moon, H. Lagercrantz, P. K. Kuhl, Language experienced in utero affects vowel perception after birth: A two-country study. *Acta. Paediatr.* **102**, 156–160 (2013).
2. D. K. James, Fetal learning: A critical review. *Infant Child Dev.* **19**, 45–54 (2010).
3. G. Westermann, D. Mareschal, M. H. Johnson, S. Sirois, M. W. Spratling, M. S. C. Thomas, Neuroconstructivism. *Dev. Sci.* **10**, 75–83 (2007).
4. R. L. Fantz, Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science* **146**, 668–670 (1964).
5. C. Kidd, S. T. Piantadosi, R. N. Aslin, The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLOS ONE* **7**, e36399 (2012).
6. D. K. Kinney, J. Kagan, Infant attention to auditory discrepancy. *Child Dev.* **47**, 155–164 (1976).
7. M. M. Botvinick, Y. Niv, A. G. Barto, Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition* **113**, 262–280 (2009).
8. C. Addyman, D. Mareschal, Local redundancy governs infants' spontaneous orienting to visual–temporal sequences. *Child Dev.* **84**, 1137–1144 (2013).
9. J. Gottlieb, P.-Y. Oudeyer, M. Lopes, A. Baranes, Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends Cogn. Sci.* **17**, 585–593 (2013).
10. M. Browning, T. E. Behrens, G. Jocham, J. X. O'Reilly, S. J. Bishop, Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* **18**, 590–596 (2015).
11. M. R. Nassar, R. Bruckner, M. J. Frank, Statistical context dictates the relationship between feedback-related EEG signals and learning. *eLife* **8**, e46975 (2019).
12. K. Krishnamurthy, M. R. Nassar, S. Sarode, J. I. Gold, Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. *Nat. Hum. Behav.* **1**, 0107 (2017).
13. A. Zénon, Eye pupil signals information gain. *Proc. Biol. Sci.* **286**, 20191593 (2019).
14. K. E. Twomey, G. Westermann, Curiosity-based learning in infants: A neurocomputational approach. *Dev. Sci.* **21**, e12629 (2018).
15. P.-Y. Oudeyer, F. Kaplan, V. V. Hafner, Intrinsic motivation systems for autonomous mental development. *IEEE Trans. Evol. Comput.* **11**, 265–286 (2007).
16. M. Bellemare, S. Srinivasan, G. Ostrovski, T. Schaul, D. Saxton, R. Munos, Unifying count-based exploration and intrinsic motivation, in *Advances in Neural Information Processing Systems (NIPS)* (2016), pp. 1471–1479.

17. D. Pathak, P. Agrawal, A. A. Efros, T. Darrell, Curiosity-driven exploration by self-supervised prediction, in *Proceedings of the 34th International Conference Machine Learning* (2017), pp. 2778–2787.

18. R. Houthooft, X. Chen, Y. Duan, J. Schulman, F. De Turck, P. Abbeel, Vime: Variational information maximizing exploration, in *Advances in Neural Information Processing Systems (NIPS)* (2016), pp. 1109–1117.

19. C. H. Rankin, T. Abrams, R. J. Barry, S. Bhatnagar, D. F. Clayton, J. Colombo, G. Coppola, M. A. Geyer, D. L. Glanzman, S. Marsland, F. K. McSweeney, D. A. Wilson, C.-F. Wu, R. F. Thompson, Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiol. Learn. Mem.* **92**, 135–138 (2009).

20. C. Kidd, B. Y. Hayden, The psychology and neuroscience of curiosity. *Neuron* **88**, 449–460 (2015).

21. E. Kayhan, S. Hunnius, S. J. X. O'Reilly, H. Bekkering, Infants differentially update their internal models of a dynamic environment. *Cognition* **186**, 139–146 (2019).

22. K. S. Tummeltshammer, N. Z. Kirkham, Learning to look: Probabilistic variation and noise guide infants' eye movements. *Dev. Sci.* **16**, 760–771 (2013).

23. J. X. O'Reilly, U. Schüffelgen, S. F. Cuell, T. E. J. Behrens, R. B. Mars, M. F. S. Rushworth, Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E3660–E3669 (2013).

24. F. Attneave, *Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods, and Results* (Holt, Rinehart and Winston, 1959).

25. A. Clark, Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).

26. K. J. Friston, J. Daunizeau, J. Kilner, S. J. Kiebel, Action and behavior: A free-energy formulation. *Biol. Cybern.* **102**, 227–260 (2010).

27. E. Mather, Novelty, attention, and challenges for developmental psychology. *Front. Psychol.* **4**, 491 (2013).

28. C. Houston-Price, S. Nakai, Distinguishing novelty and familiarity effects in infant preference procedures. *Infant Child Dev.* **13**, 341–348 (2004).

29. L. Gerken, C. Dawson, R. Chatila, J. Tenenbaum, Surprise! Infants consider possible bases of generalization for a single input example. *Dev. Sci.* **18**, 80–89 (2015).

30. S. Sirois, D. Mareschal, An interacting systems model of infant habituation. *J. Cogn. Neurosci.* **16**, 1352–1362 (2004).

31. J. Triesch, C. Teuscher, G. O. Deák, E. Carlson, Gaze following: Why (not) learn it? *Dev. Sci.* **9**, 125–147 (2006).

32. M. J. Gruber, B. D. Gelman, C. Ranganath, States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* **84**, 486–496 (2014).

33. E. S. Bromberg-Martin, O. Hikosaka, Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* **63**, 119–126 (2009).

34. W. Schultz, P. Dayan, P. R. Montague, A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).

35. M. F. Rushworth, R. B. Mars, C. Summerfield, General mechanisms for making decisions? *Curr. Opin. Neurobiol.* **19**, 75–83 (2009).

36. R. S. Hessels, D. C. Niehorster, C. Kemner, I. T. C. Hooge, Noise-robust fixation detection in eye movement data: Identification by two-means clustering (I2MC). *Behav. Res. Methods* **49**, 1802–1823 (2017).

37. S. N. Wood, N. Pya, B. Säfken, Smoothing parameter and model selection for general smooth models. *J. Am. Stat. Assoc.* **111**, 1548–1563 (2016).

38. T. Hastie, R. Tibshirani, Exploring the nature of covariate effects in the proportional hazards model. *Biometrics* **46**, 1005–1016 (1990).

39. S. Lo, S. Andrews, To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Front. Psychol.* **6**, 1171 (2015).

40. R. B. Mars, S. Debener, T. E. Gladwin, L. M. Harrison, P. Haggard, J. C. Rothwell, S. Bestmann, Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *J. Neurosci.* **28**, 12539–12545 (2008).

Citation: F. Poli, G. Serino, R. B. Mars, S. Hunnius, Infants tailor their attention to maximize learning. *Sci. Adv.* **6**, eabb5053 (2020).