

# Early roots of information-seeking: Infants predict and generalize the value of information

## Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

## About eLife's process

## Reviewed preprint version 1

January 10, 2024 (this version)

## Sent for peer review

September 6, 2023

## Posted to preprint server

August 31, 2023

Tommaso Ghilardi , Francesco Poli, Marlene Meyer, Olympia Colizoli, Sabine Hunnius

Centre for Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London, London, United Kingdom • Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands

 [https://en.wikipedia.org/wiki/Open\\_access](https://en.wikipedia.org/wiki/Open_access)

 Copyright information

## Abstract

Humans face the challenge of making sense of a complex world. Learning where to find information is crucial to filter through the abundance of stimuli, distinguish relevant from irrelevant sources, and optimize our learning. Here, we examined the developmental roots of information-seeking by testing whether 8-month-old infants can predict where to find information. We presented infants with visual cues indicating whether they will later receive information about the location of a rewarding stimulus. We analyzed the dynamics of pupil dilation when the cues were presented, but before the actual information was delivered. By combining additive Bayesian models with reinforcement learning, we show that infants learn to successfully predict what cues have a greater informational value and that they generalize these predictions to novel cues that share the same perceptual features. These results reveal the fundamental learning processes that support information-seeking from early in life.

### eLife assessment

This **valuable** study addresses the question of whether pupil size can serve as a sensitive physiological marker of information anticipation in human infants. The authors present **solid** experimental findings indicating that pupil size differs depending on the expected information content of a visual signal and that this effect might rapidly generalize to new visual information. The results could be further strengthened by additional eye movement processing and statistical analyses to rule out confounding effects of saccades and other artifacts as well as a stronger and more consistent rationale for excluding data.

## Introduction

“Information does not stream into us from the environment. Rather, it is we who explore the environment and suck information from it actively, like food.”

The ability to find valuable information is essential for successfully navigating our complex environment. In today's information-rich world, this can be a challenge, and failure to do so can result in being misinformed, misguided, or even deceived (Case, 2007 [↗](#); Hwang and Jeong, 2023 [↗](#)). The need to distil useful information from the environment is particularly evident when considering the experience infants have in their first year of life. Although infants' experience of the world is vastly different from our own, they also encounter a vast amount of novel input they have to make sense of (Johnson and Hannon, 2015 [↗](#)). To do so, they have to learn to detect relevant stimuli amidst a constant stream of multisensory input (Hunnius, 2022 [↗](#)). As they learn to interact with their environment, infants develop expectations about where and how to find valuable information. This enables them to focus their attentional resources on stimuli that are likely to provide useful information while ignoring irrelevant ones, resulting in more efficient learning. This process of information-seeking is the foundation for a lifetime of learning and discovery, but its origins are not yet well understood. Previous research has demonstrated that even infants as young as 8 months of age guide their attention based on the informativity of stimuli (Poli et al., 2020 [↗](#)) and can identify which social partner is more reliable in delivering information (Tummeltshammer et al., 2014 [↗](#)). Nevertheless, the cognitive mechanisms that enable infants to detect or infer where to find information have yet to be described. Here, we propose that statistical learning (Kirkham et al., 2002 [↗](#); Krogh et al., 2012 [↗](#)) and generalization processes (Aslin, 2017 [↗](#); Kemp et al., 2007 [↗](#); Lake et al., 2015 [↗](#); Yuan et al., 2020 [↗](#)) jointly support infants' ability to learn where information can be found.

Research has shown that from early in life infants possess the ability to extract statistical regularities from streams of sounds or visual patterns (Krogh et al., 2012 [↗](#); Saffran et al., 1996 [↗](#)). Saffran and colleagues demonstrated that 8-month-old infants can detect statistical regularities in a continuous auditory stream consisting of four three-syllable nonsense words repeated in random order (Saffran et al., 1996 [↗](#)). The stream provided no acoustic cues to word boundaries except for the transitional probabilities between syllable pairs. At test, infants were able to discriminate between “words” and “nonwords” with the same syllables in a different order. This indicates that infants can extract and utilize statistical regularities from speech after only a short exposure. While previous research has demonstrated that infants can infer structures from statistical regularities in sensory stimuli, we propose that infants may also exploit this ability to learn regularities in informativity. In other words, infants may learn which input is consistently informative and use this knowledge to generate expectations about where to find information. Infants' remarkable ability to generalize and apply learned knowledge to new situations has been demonstrated in several studies (Baram et al., 2021 [↗](#); Werchan et al., 2016 [↗](#), 2015 [↗](#)). For example, Werchan and colleagues (2015) [↗](#) showed that 8-month-old infants could learn which feature (colour and shape) was informative about the location of a subsequent cartoon, as they became faster at predicting that location with their gaze over trials (Werchan et al., 2015 [↗](#)). Importantly, the infants were able to apply this knowledge to new stimuli that followed the same rule, indicating their generalization ability. In this study, we investigated whether infants can use their generalization skills not only to predict stimuli but also the informativity of stimuli. This would allow infants to optimize their learning by focusing on informative stimuli and thereby drastically saving time and effort in the long run. Specifically, infants might use perceptual features of stimuli (like colour or shape) and generalize their learned expectations about informativity to novel stimuli that share the same perceptual features.

To test our hypotheses, we designed a novel experimental paradigm where infants could form expectations about the informativity of incoming stimuli. Infants were presented with static cues that varied in their border type (see **Fig. 1B** [↗](#)), indicating whether they would receive information about the location of a rewarding stimulus (see **Fig. 1A** [↗](#)). Then, the shapes moved in a way that either did or did not signal the location of a subsequent reward. We measured infants' pupil dilation in response to the presentation of informative and uninformative cues before the information itself was delivered. During stimulus presentation, infants' pupil dilation was continuously measured using an eye-tracker. Previous research has shown that pupil dilation

highly correlates with stimulus uncertainty (Joshi and Gold, 2020 [↗](#); Lavín et al., 2013 [↗](#); Preuschoff et al., 2011 [↗](#)), which refers to the unpredictability or ambiguity surrounding a stimulus outcome, and with the expected information gain of a stimulus in a task (Zénon, 2019 [↗](#)). In particular, it has been observed that pupil dilation increases when stimuli are more uncertain or less informative, and decreases when they are more predictable or highly informative (Joshi and Gold, 2020 [↗](#); Lavín et al., 2013 [↗](#); Preuschoff et al., 2011 [↗](#); Zénon, 2019 [↗](#)). Consequently, we expected that infants' pupil size would increase when presented with an uninformative cue, and decrease when presented with an informative cue, indicating that infants can predict whether they will receive information or not. We expected this effect to build up across trials, suggesting that infants gradually develop expectations regarding the informativity of the cues presented to them.

In addition to assessing infants' ability to learn and predict the informativity of upcoming stimuli, our study also tested infants' capacity to generalize this knowledge to novel stimuli. To explore infants' capacity to generalize this knowledge to novel stimuli, cues with novel shapes were introduced after several trials which followed the same rule as infants had seen so far (e.g., a pointy shape predictive of upcoming information and a smooth shape of no information). By analyzing infants' pupil response to the familiar and novel cues, we aimed to gain insights into whether infants were able to extend their acquired knowledge instead of learning from scratch about the new cues. We expected infants to exhibit a similar pattern of pupil dilation in response to both novel and familiar cues, indicating that they generalized their acquired knowledge to new stimuli

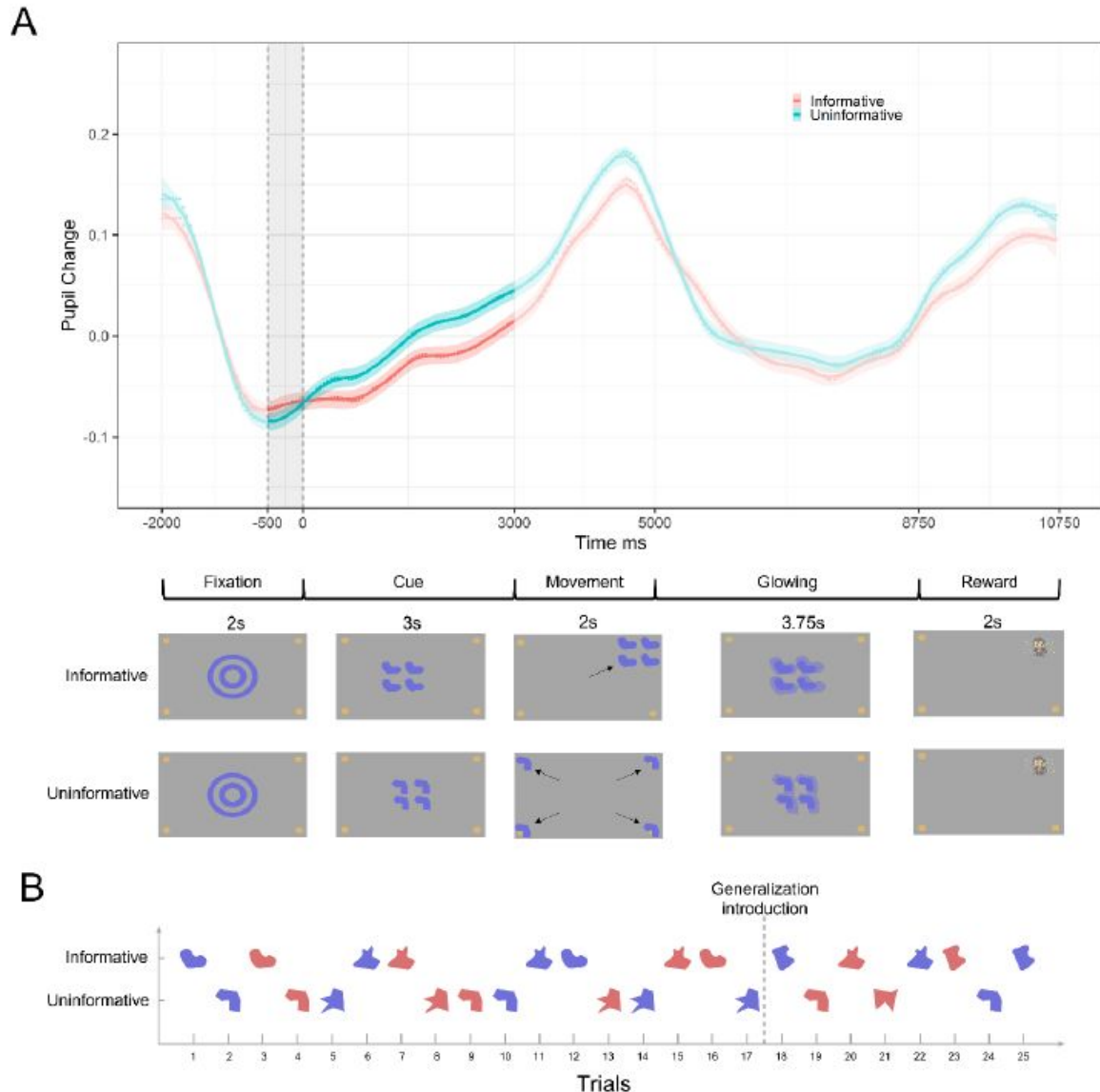
In summary, our study aimed to investigate infants' early information-seeking abilities. Specifically, we examined how infants use domain-general abilities such as statistical learning and generalization to generate expectations about the informativity of upcoming stimuli. This ability is crucial to actively determine where to find information, thus highlighting how infants take an active role in their own learning to maximize information gains while minimizing time and effort.

## Results

### Infants learn the informativity of novel stimuli through statistical learning

We preprocessed infants' pupillometry data (see Methods) and used Bayesian additive models to fit the pupil response time-locked to the cue onsets (see **Fig. 2A** [↗](#)). Additive models are a useful tool for capturing complex relationships between variables as they allow to detect any type of nonlinear relationship between dependent and independent variables. Leveraging Bayesian models' inherent ability to generate predictive distributions for parameters estimates, we investigated multiple aspects of the estimated pupil response. We identified a significant difference in pupil dilation between conditions (beta mean difference = 0.007, 89%HDI = [0.002, 0.01]). When infants were presented with informative cues, their pupil dilation was smaller compared when they were presented with cues that were not informative (see **Fig. 2B** [↗](#)). This is consistent with the idea that pupil dilation decreases when uncertainty is lower (Joshi and Gold, 2020 [↗](#); Lavín et al., 2013 [↗](#); Preuschoff et al., 2011 [↗](#)), and it is indicative of infants' ability to detect what stimuli will lead to information. These findings are further supported by complementary analyses of infants' looking behaviour (see Supplementary Materials).

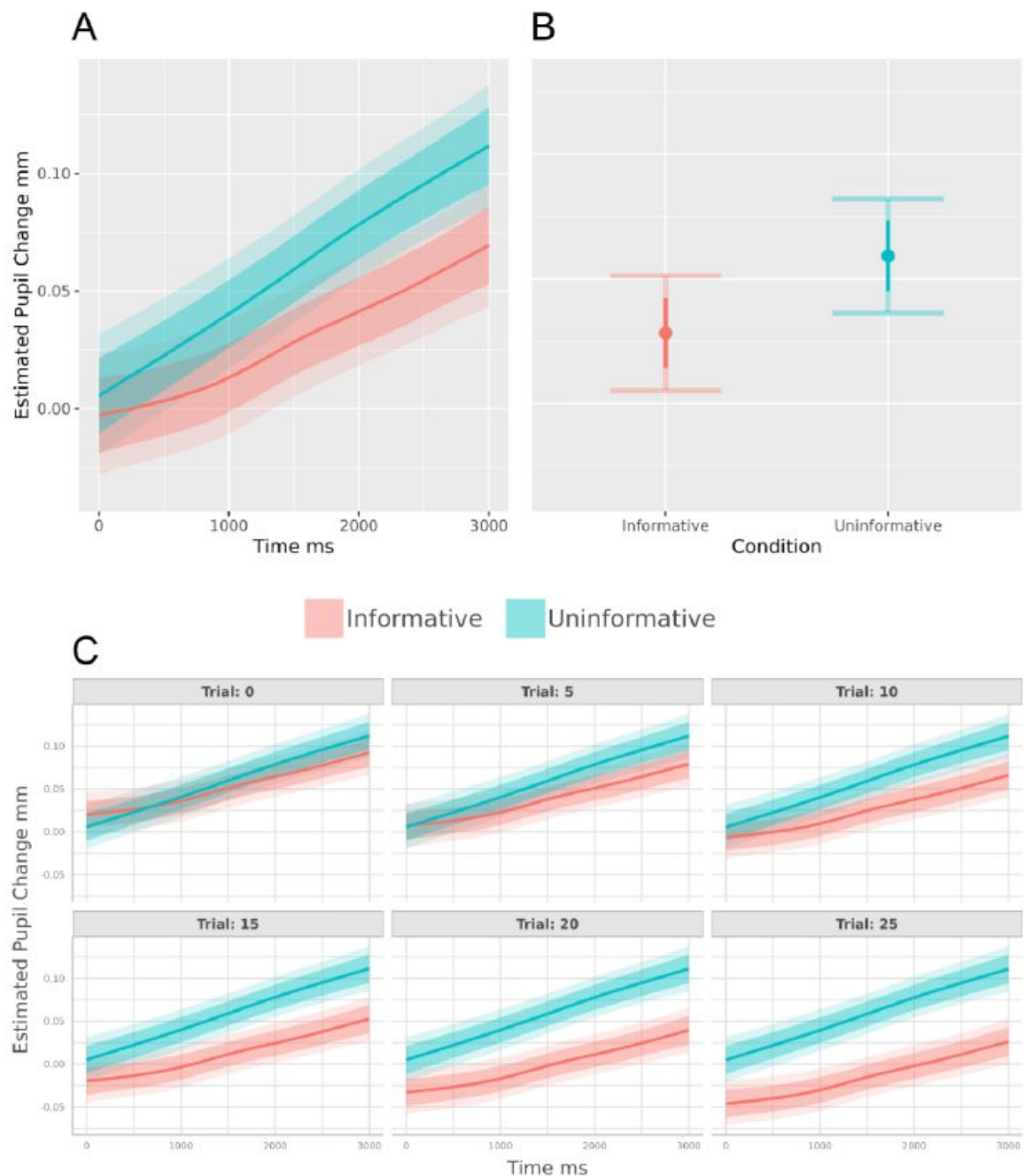
To study infants' learning we focused on the trial-by-trial change in the pupillary response to the cues (see **Fig. 2C** [↗](#)). This showed a significant interaction between condition (informative vs uninformative) and trial number (beta mean difference = 0.002, 89%HDI = [0.002, 0.003]). The interaction was driven by a gradual decrease in pupil size in the informative condition (estimated beta = -0.002, 89%HDI = [-0.002, -0.002]) over the course of multiple trials, while the pupil size in



**Figure 1.**

### Task Design and Pupil Dilation Signal.

**A. Upper part:** Baseline-corrected change in pupil dilation during a trial (mean signal and 95% confidence intervals estimated with a Local Polynomial regression). The grey-shaded area indicates the baseline and the unshaded area indicates the time-window of interest, in which infants can predict whether they will receive information. **Lower part:** Informative and uninformative trials. After a fixation stimulus, 4 identical static shapes (i.e., the cue) were presented. The border type of the shapes (pointy vs. smooth) predicted whether their following movement was informative. In informative trials, all four shapes moved to one corner of the screen signalling the location of the reward. In uninformative trials, each shape moved to a different corner of the screen. After the shapes moved back to the centre and glowed up twice, a cartoon animal was presented as reward. **B.** Example of the cues presented in the first 25 trials. Border type and colours were counterbalanced across participants. After 17 trials, new shapes were added. From that moment onwards, on each trial either familiar or novel shapes were presented as cues.



**Figure 2.**

### Pupil dilation during informative and uninformative events.

**A.** The Bayesian additive models estimated pupil change during the predictive time window, with informative trials in red and uninformative trials in blue. The shaded areas represent the standard error (darker) and 89% credible interval (lighter) of the estimate. The x-axis displays time in milliseconds and the y-axis shows the estimated pupil change from baseline. **B.** Overall, the estimated pupil change was lower for informative trials compared to uninformative trials. **C.** The difference between the conditions developed over trials, as infants learned which stimuli were informative. Across trials, the pupil constricted more in informative trials while it remained unchanged in uninformative trials.

the uninformative condition remained unchanged (estimated  $\beta = -3e-5$ , 89%HDI =  $[-3e-4, 2e-4]$ ) (see **Supplementary figure 1** [↗](#)). The pupil size difference between conditions thus gradually emerged as evidence accumulated.

## The temporal dynamics of infant learning match the predictions of a reward-based reinforcement-learning model

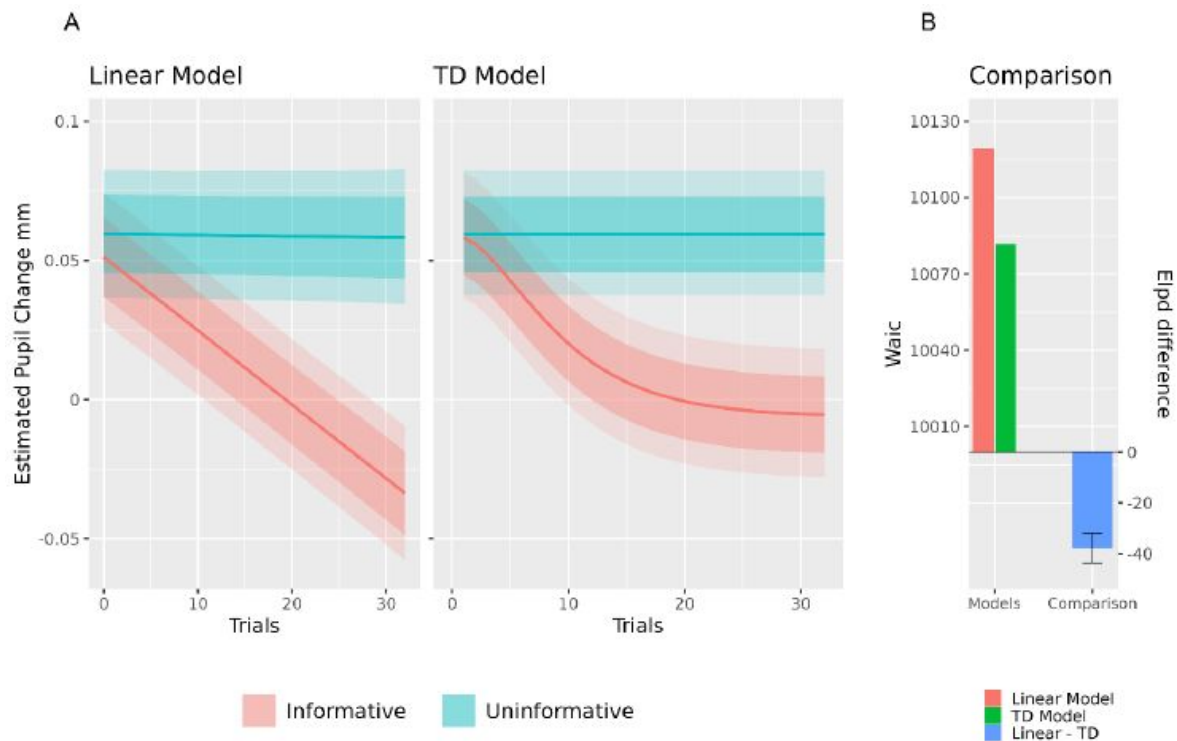
While our data demonstrate that infants can learn where to find information, the mechanism underlying this learning remains unclear. As mentioned previously, statistical learning may enable infants to learn which stimuli are informative and which are not. However, our analysis has assumed a linear relationship between uncertainty and trials, without explicitly modelling this change in uncertainty over the course of the trials. As a result, the model assumed a gradual linear decrease in uncertainty associated with the cues as the number of trials increased. To gain a better understanding of the underlying learning process, we conducted an exploratory analysis to investigate whether the change in uncertainty aligned with the predictions of a temporal-difference (TD) learning model.

Temporal difference (TD) learning is a reinforcement learning algorithm ([Gabriel and Moore, 1990](#) [↗](#); [Sutton and Barto, 2012](#) [↗](#)) that quantifies the temporal dynamics that enable an agent to predict rewards based on environmental cues. This is achieved by continually updating an agent's predictions as new evidence becomes available, thus improving them over time. Recently, TD-learning has been employed as a biologically plausible implementation of statistical learning ([Orpella et al., 2021](#) [↗](#)). In our study, TD-learning predicts that, over time, infants shift the value that they attribute to the information (here the movement indicating the location of a reward stimulus) to the static cue which signals upcoming information. The term “value” refers to the relative importance or usefulness of a particular event or stimulus in achieving a goal, in this case, obtaining information about the location of the reward. Hence, when the movement is informative, the value of the cues should slowly increase over time, following a TD-learning function (see **Fig. 4** [↗](#)). Our results are consistent with these predictions, as we found a high correlation between infants' pupil dilation and the expected reduction of uncertainty (i.e., the information gain) estimated by a TD-learning model ( $\beta$  mean =  $-0.064$ , 89%HDI =  $[0.059, 0.069]$ ). Crucially, this model performed better than the previous statistical model which assumed a linear decrease in pupil dilation over time (waic elpd difference =  $-37.7$ , waic standard error difference =  $5.9$ ). This suggests that infants learn where to expect information by shifting the value of the informative events to the predictive cues.

## Infants generalize their acquired knowledge

After 17 trials, novel cues were introduced, which shared relevant (i.e. the type of shape) and irrelevant features (i.e., the colour) with the familiar cues. We tested whether infants exploited the informative features to quickly generalize the informativity to the novel cues. A Bayesian additive model showed that infants' pupil dilation was reduced for novel cues. This was specific to those novel cues that shared the features of the familiar informative cues (estimated mean difference =  $-0.05$ , 89%HDI =  $[-0.062, -0.038]$ ). The size of this effect approximated the difference between conditions that were observed for familiar stimuli (estimated mean difference =  $-0.067$ , 89% HDI =  $[-0.077, -0.057]$ ). Crucially, this difference was not observable at the start of the task, when the familiar stimuli were first introduced (estimated mean difference =  $-0.007$ , 89%HDI =  $[-0.015, 0.001]$ ). This suggests that two different learning processes were at play: one slower process allowed to learn where to find information, while a faster generalization process allowed to apply what had been learned to novel instances.

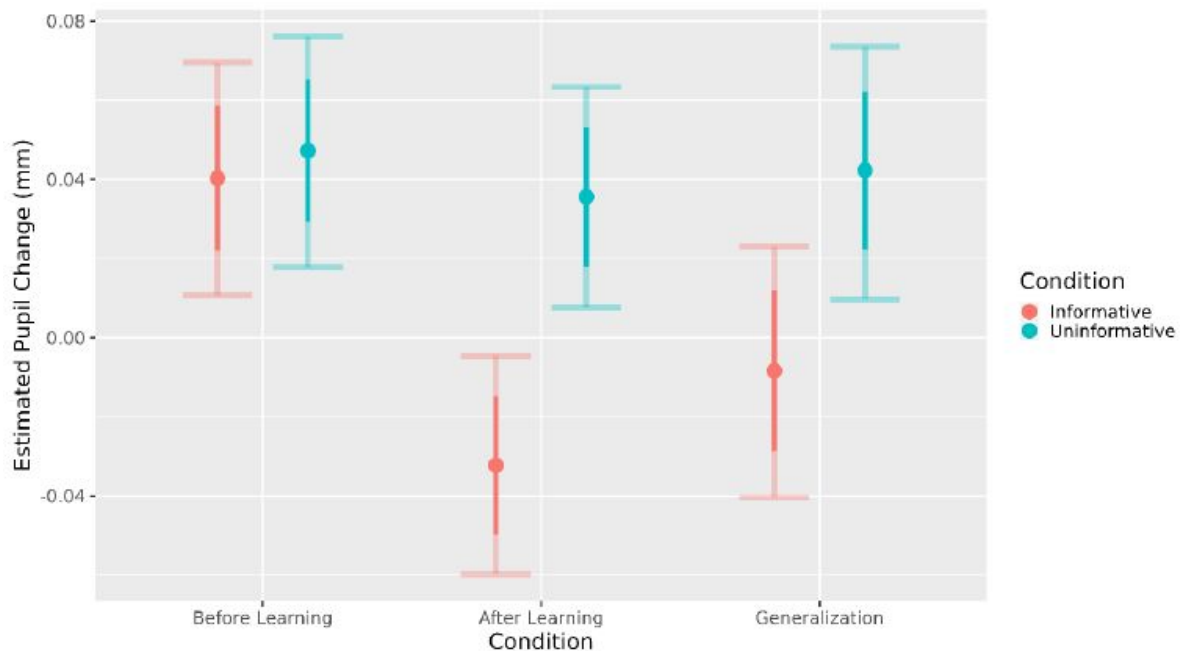




**Figure 3.**

### Model comparison.

**A.** Estimated pupil change over trials as predicted by the linear and TD-learning models. The linear model is a purely statistical model, while the TD model also makes assumptions about the underlying cognitive mechanisms. The shaded areas represent the standard error (darker) and 89% credible interval (lighter) of the estimate. **B.** Model comparison was performed comparing waic scores. The TD model (in green) had a lower WAIC score, indicating better performance. The elpd difference (in blue) offers a direct comparison between the two models, showing that the TD-learning model was significantly better than the linear model. The errorbar represents the standard deviation of the elpd difference.



**Figure 4.**

#### **Pupil dilation during different learning moments.**

Change in the estimated pupil change is displayed as a function of cue type (informative/uninformative) and learning (before learning/after learning/generalization). As expected, before learning (i.e., trials 1 to 4), there was no difference in pupil size between the informative and uninformative trials. After learning (i.e., trials 12 to 15), infants showed a more constricted pupil in informative trials compared to uninformative ones. This pattern was also shown for the generalization trials (i.e., trials 18 to 21), suggesting that infants were able to generalize their knowledge to novel, unseen stimuli. The shaded areas represent the standard error (darker) and 89% credible interval (lighter) of the estimates.



## Discussion

How do we learn to find information that helps us to successfully navigate this world? We approached this question from a developmental perspective, examining the cognitive roots that underlie our remarkable learning skills from the first year of life. This study offers novel insights into the mechanisms by which infants detect and predict informativity and thereby unravels the fundamental learning mechanisms that support information-seeking. Infants have exceptional learning abilities, allowing them to acquire vast amounts of knowledge in a short time (James, 2010 [↗](#); Westermann et al., 2010 [↗](#)). Previous research suggested that infants preferentially deploy their cognitive resources on informative stimuli (Gottlieb et al., 2013 [↗](#); Poli et al., 2020 [↗](#)). Specifically, infants can detect whether stimuli provide new information and become more likely to disengage when information content decreases. While this study demonstrates that infants respond to the level of information provided by a stimulus, it remains unclear whether infants can predict the sources of new information. This would enable infants to strategically allocate attentional resources, focusing on stimuli likely to provide information and prepare for the acquisition of new knowledge which would make learning more effective.

In this study, we investigated the ability of infants to distill the informativity of upcoming stimuli. We found a steady reduction in pupil size over trials, indicating that infants learned that specific cues predicted whether they would later receive information about the location of a reward. This discovery supports the growing body of evidence indicating that infants are proactive in shaping their learning environment by searching for and focusing on information-rich stimuli (Poli et al., 2023 [↗](#), 2020 [↗](#)). For the first time, this demonstrates that infants do not only react to information but can learn about the informativity of a stimulus and anticipate information before it is available. This offers an explanation for other behavioural patterns previously observed in infants, like their enhanced attention towards social cues. By repeated exposure to social cues, such as eye contact (Csibra and Gergely, 2009 [↗](#)), the mouth while speaking (Hunnus and Geuze, 2004 [↗](#); Lewkowicz and Hansen-Tift, 2012 [↗](#)), or pointing gestures (Sodian and Thoermer, 2004 [↗](#)), infants may learn to expect that they carry relevant information (Begus et al., 2016 [↗](#); Tummeltshammer et al., 2014 [↗](#); Zmyj et al., 2010 [↗](#)).

We did not only show that infants can learn where to find information, but using a reinforcement learning (TD-learning) model, we also demonstrated that the changes in pupil dilation over trials were compatible with a shift of value from the informative event itself to the static cues predictive of the informativity. The TD-learning model performed better than a linear model, but it also comes with greater explanatory power, as it captures the specific learning mechanism behind this ability to build expectations of informativity. Our model assumes that information is valuable, but it remains agnostic as to why this is the case. One possibility is that information has an intrinsically positive value from birth, and this comes as a fundamental bias of the human brain to support information-seeking behaviours (Burda et al., 2018 [↗](#); Houthoofd et al., 2016 [↗](#)). Another compelling hypothesis is that the same TD-learning mechanism observed in our study is involved in shaping the value of information over time. This would predict that early in life, only rewards are valuable, not information. As information is often instrumental in obtaining rewards, the value of the reward may be progressively transferred onto information via TD-learning. For example, an infant may initially only value the immediate reward of food when presented with a spoonful of pureed fruit. However, over time, the infant may begin to associate certain sounds or movements with the delivery of food, such as the sound of a caregiver opening a jar of baby food or the sight of the spoon being brought to their mouth. Following the TD-learning approach, the infant may gradually begin to place value on these cues that predict the delivery of food, rather than solely on the immediate reward of the food itself.

Future studies are needed to investigate whether this mechanism is indeed at play very early in life, which would provide insight into the origins of the value placed on information. Future research is also needed to explore alternative computational models that may capture different

learning strategies. By comparing these models to infant learning, we may gain a more comprehensive understanding of the specific mechanisms involved in early information-seeking abilities.

Finally, our study not only demonstrates that infants can learn to predict where to find information, but also that they can generalize this knowledge to novel stimuli. Their remarkable generalization abilities<sup>17,18</sup> allow them to extend expectations about the informativity of familiar stimuli to novel stimuli which share relevant features. These findings suggest the presence of multiple learning processes in infants, with one being slower and more data-hungry, and the other being faster and relying on generalization.

This combination of fast and slow learning processes is key for effective learning, and a similar implementation in artificial agents may be essential to develop machines that learn and explore as humans do. Only recently, machines have been endowed with a bias towards information that makes them experience novel information as rewarding in itself leading to an improvement in learning speed and performance (Pathak et al., 2017 [↗](#)). Yet, the abilities displayed by infants in our study - slow learning of informativity and fast generalizations to novel stimuli - are still lacking in artificial agents. Given that these mechanisms are active from early on in humans, they may be fundamental to a successful implementation of efficient and flexible human-like learning.

In conclusion, our study sheds new light on the cognitive processes that underlie infants' remarkable learning skills. Specifically, it identifies the ability to form expectations about the informativity of new stimuli as a fundamental aspect of their learning. This study offers an explanation of how infants can process and make sense of the vast amount of sensory information they are exposed to. Hence, it contributes to a mechanistic understanding of how infants develop sophisticated models of the social and physical world around them at such a breathtaking rate.

## Materials and Methods

### Participants

Forty-four full-term infants (mean age: 8.2 months, SD: 0.24; 22 males) were recruited from a database of interested families in the Nijmegen region, a middle-sized city in the Netherlands (ECSW-2021-096). Based on the number of trials left after trial rejection, 14 infants were excluded from the subsequent data analysis (see Preprocessing) resulting in a final sample of 30 infants (mean age = 8.3 months, SD = 0.19; 12 males).

### Stimuli and Procedure

The aim of this study was to investigate infants' ability to predict the informativity of a stimulus. To accomplish this we created a set of four different shapes. Two shapes were pointy and two shapes were smooth. One of the border types (e.g., smooth) was used to signal informative trials and the other (e.g., pointy) to signal uninformative trials. The border types linked to informative and uninformative trials were counterbalanced between participants. Each shape could either be red or blue. The colour was pseudorandomized and it had no influence on the task.

Each trial of the experiment began with a fixation stimulus made of two concentric circles. Its colour and overall size matched the colour and size of the shapes presented later during the trial. Following a 2-second presentation of the fixation stimulus, four identical shapes were presented in the centre of the screen for 3 seconds. The identity of the shapes was sampled pseudorandomly from the four available shapes. This defined the "cue" which was predictive of whether or not infants would later receive information about where a cartoon animal would appear.

After the 3-second presentation of the cue, the shapes moved. Their movement pattern was dependent on the trial type. In informative trials, all four shapes moved to one corner of the screen (pseudorandomly selected on each trial). In uninformative trials, each shape moved to a different corner of the screen. After moving to the corners, the shapes returned to the centre of the screen. In total, this lasted 2 seconds. The shapes then remained static for 0.5 seconds, glowed (i.e. expanded and contracted) twice to capture back the attention of the infant for a total of 2 additional seconds, remained static for another 0.5 seconds, and then disappeared. Finally, 0.75 seconds after the disappearance of the shapes, a video of a cartoon animal appearing from a present box was displayed for 2 seconds. On each trial, the specific cartoon animal was selected pseudorandomly from a pool of four different animals. In the informative condition, the location where the cartoon animal appeared was the same as the one cued by the movement of the shapes, while in the uninformative condition, it was selected pseudorandomly. This design allowed us to separate the moment in which infants could predict that they would receive information (i.e. cue of informativity) from the moment the information was actually provided.

After 17 trials, two shapes were replaced. Specifically, we replaced one pointy shape with a new pointy shape, and we replaced one smooth shape with a new smooth shape. Thus, from the 18th trial on, infants were presented with a combination of familiar shapes that they had seen before and a novel shape that they had not encountered before, while still maintaining the overall manipulation of the study. The purpose of this condition was to investigate if infants can generalize the rule they learned to novel stimuli, thereby examining their ability to generalize their expectations of informativity to new concepts and objects.

## Preprocessing

PupillometryR software ([Forbes, 2020](#)) was used to preprocess the pupil size data. We first identified the relevant trials from the continuous data and removed segments with excessive noise. Specifically, trials that had fewer than 99% valid samples were excluded, resulting in the rejection of 1.34% of the total trials (13 trials were rejected). We then regressed the left and right pupils against each other and calculated the mean of the two pupils as the final indicator of pupil size. The data was downsampled to a rate of 20Hz by taking the median of each 50ms time-bin ([Mathôt and Vilotijević, 2022](#)).

Additionally, to further ensure the quality of the data we rejected any trials that had less than 75% of the data, and participants who kept less than 80% of the trials. This resulted in a trial rejection of 46.9% (453 trials) in relation to the total trial number and a rejection of 13 participants. The final sample on which further analysis was conducted contained 30 participants. To reduce noise in the data, we applied a Hanning window with a degree of 11 samples. Linear interpolation was used to estimate missing pupil size samples to ensure data continuity ([Jackson and Sirois, 2009](#)). Samples were then baseline-corrected by subtracting the average of the final 500 ms of the fixation stimulus thus representing pupil size change in comparison to the baseline. Finally, we selected only the 3000ms of the cue time window for further analysis ([Colizoli et al., 2018](#)).

## Analysis

To investigate the changes in pupil size, we used Bayesian mixed-effect additive models to analyze the preprocessed data. This was accomplished using the brms package ([Bürkner, 2021](#), [2018](#), [2017](#)). Additive models have been widely utilized in statistical literature as a powerful tool for data analysis, as they allow for flexible modelling of complex relationships between variables while maintaining interpretability and ease of estimation. To flexibly model the relationships between variables, additive models allow to include smooth terms. Smooth terms are flexible functions that allow for non-linear relationships between the predictor and the response variable. Furthermore, the utilization of a Bayesian framework allows for the expression of model uncertainty and the incorporation of prior information about the parameters. The application of

additive models to pupillometry data is particularly advantageous as it allows for the capture of nonlinear and complex relationships between different predictors and pupil dilation response (Hershman et al., 2023 [↗](#); van Rij et al., 2019 [↗](#)).

To properly model the fluctuation of pupil size over the time course of the trials, we included two smooth terms in each model. The first term modelled the differences in pupil fluctuation between conditions (informative and uninformative), while the second term modelled differences between subjects. Additionally, to account for residual variability between participants, we included a random intercept term in the models.

To ensure robust model fitting, each model was estimated using Markov Chain Monte Carlo (MCMC) sampling with 4 chains of 6000 iterations. We discarded 4000 iterations as warm-up and kept the remaining 2000 iterations for analysis. We specified weakly informative priors for the models to allow for flexibility in the estimation process. The detailed specification of the priors used can be found in the supplementary materials. By using this approach, we were able to model the complex relationships between the predictors and the pupil dilation response and to account for the random variation in the data. Convergence and stability of the Bayesian sampling were assessed using R-hat, which should be below 1.01 (Vehtari et al., 2020 [↗](#)), and effective sample size (ESS), which should be greater than 1000 (Bürkner, 2017 [↗](#)).

## Effect of informativity

To investigate the impact of informativity on pupil size, we conducted a statistical analysis on all trials except the generalization trials. This was done to ensure that the results of the analysis were not confounded by the additional manipulation of the generalization trials.

To investigate the effect of informativity on pupil size change, we used an additive model to analyze the preprocessed pupillometry data (Thurman et al., 2019 [↗](#)). Condition (informative vs uninformative), trial number, and their interaction were fixed factors. This approach allowed us to examine how baseline-corrected pupil size changed in relation to the predictability of the upcoming information and the trial number. Furthermore, by including the interaction term in the model, we were able to assess the development of the effect of condition on pupil size changes over time.

To further analyze the interaction between trials and the two conditions, we made use of the `estimate_slopes` function from the `modelbased` package (Makowski et al., 2020 [↗](#)). This function enabled us to determine if the slopes of the informative and uninformative trials were significantly different from zero. Thus, showing whether there were significant differences in the change of pupil dilation over trials between the two conditions.

## Ideal Learner

We employed a Temporal Difference (TD) learning algorithm to explore the mechanism underlying infants' learning. TD learning is a reinforcement learning method that enables an agent to predict whether a reward will be delivered based on previous experience. When new evidence is observed, the model can update its predictions, hence improving over time (Gabriel and Moore, 1990 [↗](#); Sutton and Barto, 2012 [↗](#)). Here, we used TD learning to predict whether information will be delivered. The model learns to predict the delivery of information based on the preceding cue stimulus. At the start of the task, the cues have not been associated with the information yet, and the initial value  $V$  of the cues is thus zero. On every timepoint  $t$  of every trial  $n$ , a prediction error (i.e., temporal-difference error, TDE) is computed as follows:

$$TDE_{n,t} = I_{n,t} + V_{n,t+1} - V_{n,t}$$

Note that TDE increases not only when information ( $I_{n,t}$ ) is delivered unexpectedly, but also when information is expected to be delivered ( $V_{n,t+1}$ ). This implies that cues preceding information increase in value across learning. This learning is implemented by updating the value associated with the cue with the prediction error:

$$V_{n+1,t} = V_{n,t} + \alpha TDE_{n,t}$$

Where  $\alpha$  indicates the learning rate, which is a free parameter that determines how much weight is assigned to the prediction error when updating the predictions. We expected  $V_{n,t}$  to correlate with the baseline-corrected changes in pupil dilation during the presentation of the static cue.

To test the value of alpha that best fitted the infant data, we used a grid-search approach with 30 different values of alpha (from 0.01 to 0.30). Specifically, we analyzed the pupillary response data in relation to the  $V_{n,t}$  values obtained with the different learning rate values using additive models (MGCV package, (Anderson-Cook, 2007 [↗](#); Pedersen et al., 2018 [↗](#))). Compared to brms, MGCV allowed for a quicker evaluation of which model would perform best. The mean pupil was modelled based on the uncertainty values extracted from the TD-learning model. To account for the fluctuations of pupil size, two smooth terms were included in each model, one to model differences in pupil fluctuation between conditions (informative and uninformative), and another to model differences between subjects.

The Akaike information criterion (AIC) was used to compare and evaluate the best-performing model. The lower the AIC, the better the model fits the data. The model with  $\alpha = 0.19$  performed best (see **Supplementary figure 3** [↗](#)). Hence, we ran a Bayesian additive mixed-effects model using a similar approach as described in the previous paragraph, with the only difference being that the mean pupil was modelled using the  $V_{n,t}$  values extracted from the TD-learning model instead of using Condition and Trial number.


Finally, we compared the performance of the TD-learning model with the linear model using the waic function from the brms package (Bürkner, 2017 [↗](#)). The function allows to compare the theoretical expected log pointwise predictive density (elpd) (Vehtari et al., 2017 [↗](#)) of different models by returning the elpd difference and the elpd standard error.

## Generalization

To investigate infants' ability to generalize the association learned between the border type and the upcoming information, we selected trials: we compared pupil dilation at the presentation of the static cue during generalization trials (trials 18 to 21) with pupil dilation at the beginning of the task, before learning has occurred (trials 1 to 4) and with pupil dilation later in the task, after learning has occurred (trials 12 to 15). To avoid possible confounds related to stimulus novelty and surprise, the very first trial of the study (trial 0) and the very first generalization trial (trial 17) were excluded. Participants that did not watch generalization trials were excluded from this analysis, resulting in a sample of 19 infants.

We modelled baseline-corrected pupil size by including condition (informative vs uninformative), learning (before learning, after learning, generalization), their interaction and trial number as fixed factors. This approach allowed us to examine the extent to which infants were able to generalize the rule they had learned to new, previously unseen trials. By comparing the baseline-corrected pupil size before and after learning the association between border type and informativity, as well as during the generalization trials, we could investigate whether infants were able to transfer the rule to novel situations and whether this transfer was reflected in the pupillary response. After fitting the model to the data we explored the contrast of the interaction between condition and generalization by using the estimate\_contrast function (Makowski et al., 2020 [↗](#)). This enabled us to explore the difference between each contrast of this interaction.

## Data Availability



The data and the analysis scripts are publicly available on OSF via: [https://osf.io/tkzf9/?view\\_only=458ffe27e0b344c8ae519259b1ef630d](https://osf.io/tkzf9/?view_only=458ffe27e0b344c8ae519259b1ef630d) 

## Supplementary material for: “Early roots of information-seeking: Infants predict and generalize the value of information”



### Pupil Size Trends

### Latency to reward location

#### Analysis

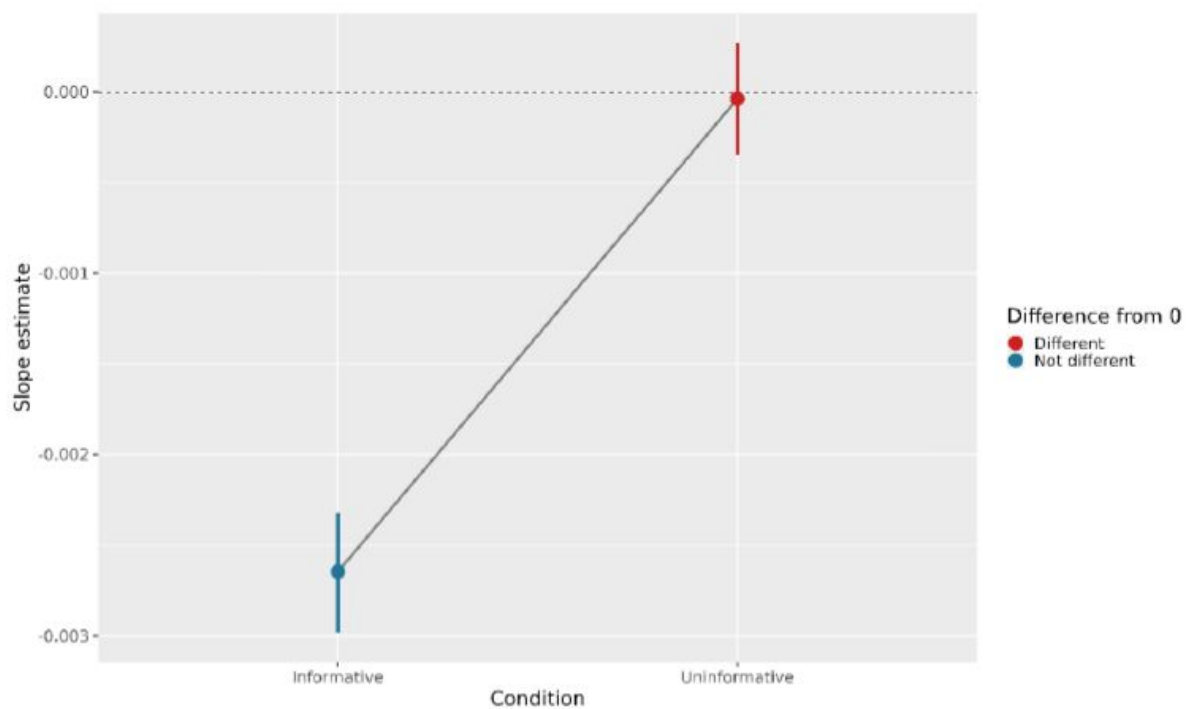
In addition to analyzing the pupillary response to the cues, we also examined infants’ latency to look at the reward location. This analysis aimed to determine if infants were using the cues’ movements to predict the reward location. Raw eye-tracking data were processed using 2-means clustering (I2MC) (Hessels et al., 2017 ) , that allows for a robust identification of fixations especially in noisy data as the one from infants. After running I2MC we delineated four areas of interest of  $400 \times 400$  pixels around the target locations and extracted latencies using a Python script. Latencies were extracted between 1250ms before the reward presentation and 1000ms after the presentation of the reward. These latency values were then standardized for each individual participant by computing z-scores using their mean and standard deviation. Finally, we employed a Bayesian generalized linear mixed-effects model to investigate whether infants relied on the movement of the cues to guide their gaze to the location of the reward. Due to the response variable being limited to 1000ms after the reward appearance, the response distribution of the model was truncated using the trunc function (Bürkner, 2017 ) . The model included condition (Informative vs Uninformative) and trial number as fixed factors. We also included a random intercept term in the model to account for individual subject variability. Similarly to the models run for the main analysis, the model was estimated using MCMC sampling with 4 chains of 6000 iterations, with 4000 iterations as a warm-up. The detailed specification of the priors used can be found in the *Models’ priors* paragraph.

#### Results

Following preprocessing of infants’ eye-tracking data, we fitted Bayesian additive models to analyze latency to the reward location time-locked to the reward onsets (**Fig. 2A** ) . The analysis revealed a significant difference in latency between conditions (mean = 0.16485723, 89%CI = [-0.014, 0.347]). Infants were faster to look at the reward when they received information about its location compared to when they did not receive any information (**Supplementary figure 2** ) . These findings confirm infants’ ability to associate the movement of the cues with the location of the forthcoming rewards and utilize this information to attend to the correct location.

#### Learning rate grid search analysis

As mentioned in the main manuscript the TD-learning estimates were extracted using different learning rate values. Specifically, we extracted the TD-learning estimated using learning rates ranging from 0.1 to 0.3. To determine the learning rate that best approximated that of our infant sample, we conducted a grid-search analysis. We applied additive models using the MGCV package

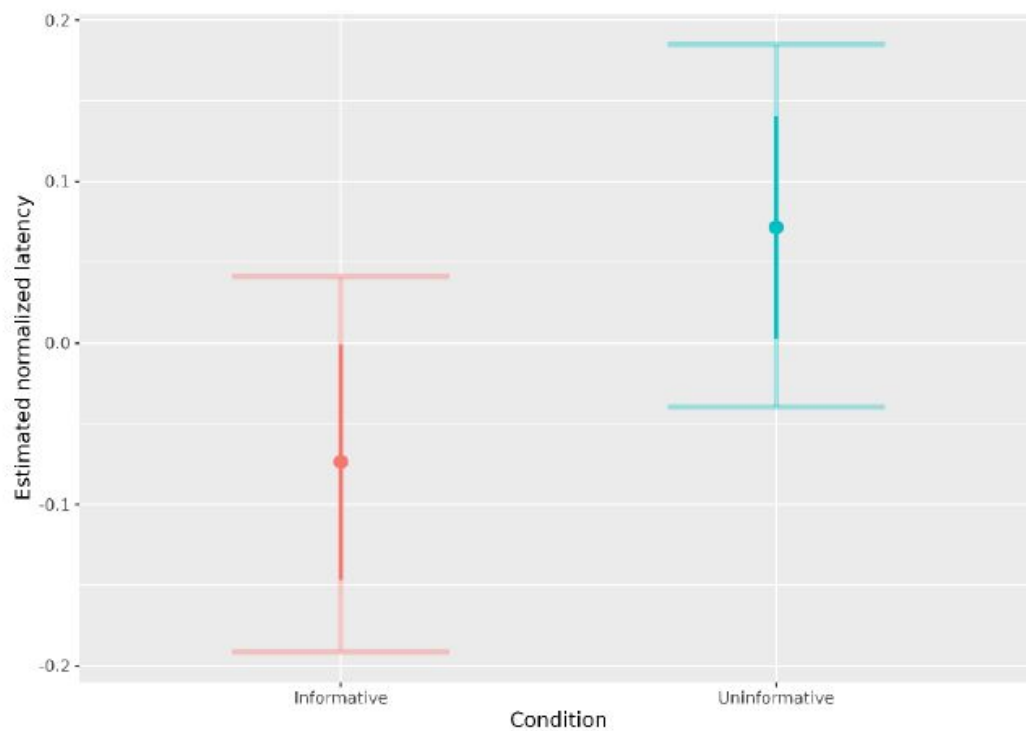


### Supplementary figure 1

#### Pupil size trends across conditions.

Slope estimates by condition derived from the Bayesian additive model applied to the predictive time window. The graph shows a negative slope in the informative condition, indicating a consistent decrease in pupil size over multiple trials. In the uninformative condition, the slope is not different from zero, indicating no significant change in pupil size across trials.





### Supplementary figure 2

#### Latency to the reward location.

The estimated normalized latency to the reward location is shown as a function of cue type (informative/uninformative). The informative trials are plotted in red, while the uninformative trials are plotted in blue. The shaded areas represent the standard error (darker) and 89% credible interval (lighter) of the estimate.

(Anderson-Cook, 2007 [↗](#); Pedersen et al., 2018 [↗](#)) to analyze the pupillary response data in relation to each set of TD-learning estimates. To account for the fluctuations of pupil size over the course of trials, two smooth terms were included in each model, one to model differences in pupil fluctuation between conditions (informative and uninformative), and another to model differences between subjects. The Akaike information criterion (AIC) was used to compare and evaluate the best-performing model. Our results revealed that the model with the highest performance was the one based on the TD-learning estimates extracted using a learning rate of 0.19 (**Supplementary figure 3** [↗](#), **Supplementary table 1** [↗](#)). Consequently, the analyses presented in the main manuscript were conducted using the TD-learning estimates obtained with this learning rate.

## Models' priors

### Pupil dilation

In the additive Bayesian models used to explore the pupil changes fluctuation, we have specified weakly informative priors for the fixed effects, smooth terms, and the intercept. The choice of these priors was based on our knowledge of the response variable (pupil dilation) and insights gained from data visualization.

The fixed effects (class 'b') have been assigned normal priors with a mean of 0 and a standard deviation of 0.2, reflecting our expectation that these effects are centred around zero, with a modest degree of uncertainty. This choice of prior represents a relatively weak constraint on the fixed effects, allowing the data to play a more substantial role in updating the posterior distribution.

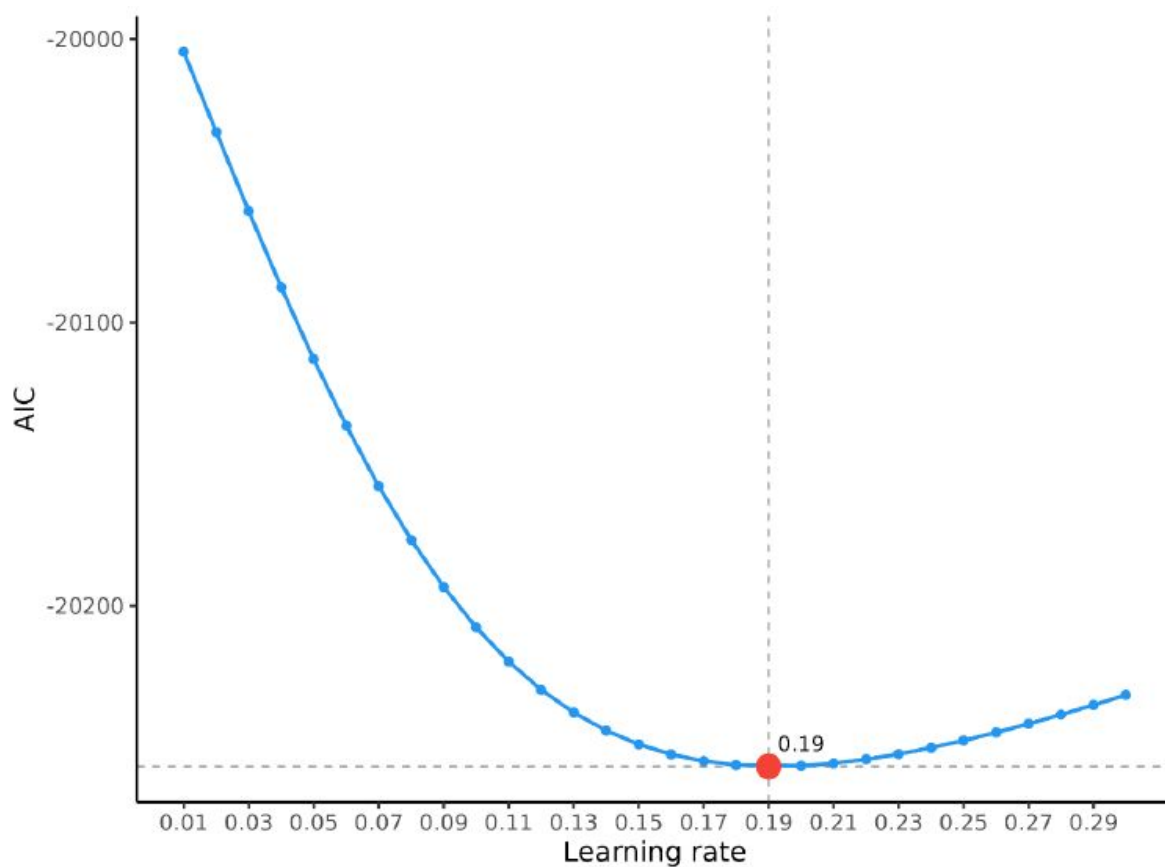
For the standard deviations of the smooth terms (class 'sds'), we employed Student's t-distribution priors with 5 degrees of freedom, a mean of 0, and a scale parameter of 0.2. The use of a Student's t-distribution with a small number of degrees of freedom allows for heavier tails than the normal distribution, providing a more robust estimate against potential outliers. Additionally, the choice of a scale parameter of 0.2 represents a relatively weak constraint on the smooth terms' standard deviations, enabling the data to inform the uncertainty associated with these terms.

Lastly, the intercept (class 'Intercept') was assigned a Student's t-distribution prior with 5 degrees of freedom, a mean of 0, and a scale parameter of 0.3. This choice of prior distribution for the intercept is similar to the one used for the smooth terms' standard deviations, with the heavier tails providing robustness against potential outliers. The scale parameter of 0.3 reflects a modest degree of uncertainty about the intercept's true value, allowing the data to update the prior distribution accordingly.

In addition to these specified priors, we retained the default priors provided by the brms package for other parameters of the model, such as the residual standard deviation (sigma), group-level standard deviations (sd), and the degrees of freedom parameter (nu) for the Student's t family. The default priors utilized are as follows: a gamma distribution with shape and rate parameters set at 2 and 0.1 respectively for the nu parameter, and a Student's t-distribution with 3 degrees of freedom, a mean of 0, and a scale parameter of 2.5 for both the sd and sigma parameters.

### Latency

In the Bayesian model employed to investigate infants' latency to look to the reward location, we utilized the default priors provided by the brms package for all the parameters, including fixed effects, group-level effects, and residual standard deviation.



### Supplementary figure 3

#### AIC values.

AIC values of the additive models fitted on TD-learning estimates extracted using different learning rates. The figure shows that the model with the lowest AIC was obtained using a learning rate of 0.19.

LearningRate	AIC
0.01	-20004.53
0.02	-20032.84
0.03	-20060.75
0.04	-20087.63
0.05	-20113
0.06	-20136.46
0.07	-20157.75
0.08	-20176.73
0.09	-20193.32
0.1	-20207.59
0.11	-20219.63
0.12	-20229.59
0.13	-20237.66
0.14	-20244.02
0.15	-20248.86
0.16	-20252.38
0.17	-20254.74
0.18	-20256.11
0.19	-20256.61
0.2	-20256.38
0.21	-20255.53
0.22	-20254.14
0.23	-20252.3
0.24	-20250.07
0.25	-20247.52
0.26	-20244.7
0.27	-20241.65
0.28	-20238.41
0.29	-20235
0.3	-20231.46

### Supplementary table 1

#### AIC values in relation to the learning rate.

AIC values of models fitted on TD-learning estimates extracted using different learning rates

For the fixed effects (class 'b'), flat priors were used, which implies uniform probability distribution over a wide range of values. The intercept (class 'Intercept') was assigned a default Student's t-distribution prior with 3 degrees of freedom, a mean of 0.1, and a scale parameter of 2.5. For the group-level standard deviations (class 'sd'), a Student's t-distribution prior with 3 degrees of freedom, a mean of 0, and a scale parameter of 2.5 was employed. Lastly, the residual standard deviation (class 'sigma') was assigned a Student's t-distribution prior with 3 degrees of freedom, a mean of 0, and a scale parameter of 2.5.

## References

- Anderson-Cook CM (2007) **Generalized Additive Models: An Introduction With R** *J Am Stat Assoc* **102**:760–761
- Aslin RN (2017) **Statistical learning: a powerful mechanism that operates by mere exposure: Statistical learning** *Wiley Interdiscip Rev Cogn Sci* **8**
- Baram AB, Muller TH, Nili H, Garvert MM, Behrens TEJ (2021) **Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems** *Neuron* **109**:713–723
- Begus K, Gliga T, Southgate V (2016) **Infants' preferences for native speakers are associated with an expectation of information** *Proc Natl Acad Sci U S A* **113**:12397–12402
- Burda Y, Edwards H, Pathak D, Storkey A, Darrell T, Efros AA (2018) **Large-Scale Study of Curiosity-Driven Learning** *arXiv [csLG]*
- Bürkner P-C (2021) **Bayesian Item Response Modeling in R with brms and Stan** *J Stat Softw* **100**:1–54
- Bürkner P-C (2018) **Advanced Bayesian Multilevel Modeling with the R Package brms** *R J* **10**
- Bürkner P-C (2017) **brms: An R Package for Bayesian Multilevel Models Using Stan** *J Stat Softw* **80**
- Case DO (2007) **Looking for information: A survey of research on information seeking, needs, and behavior**
- Colizoli O, de Gee JW, Urai AE, Donner TH (2018) **Task-evoked pupil responses reflect internal belief states** *Sci Rep* **8**
- Csibra G, Gergely G (2009) **Natural pedagogy** *Trends Cogn Sci* **13**:148–153
- Forbes S (2020) **PupillometryR: An R package for preparing and analysing pupillometry data** *J Open Source Softw* **5**
- Gabriel MR, Moore JW (1990) **Learning and Computational Neuroscience: Foundations of Adaptive Networks**
- Gottlieb J, Oudeyer P-Y, Lopes M, Baranes A (2013) **Information-seeking, curiosity, and attention: computational and neural mechanisms** *Trends Cogn Sci* **17**:585–593
- Hershman R, Milshtein D, Henik A (2023) **The contribution of temporal analysis of pupillometry measurements to cognitive research** *Psychol Res* **87**:28–42
- Hessels RS, Niehorster DC, Kemner C, Hooge ITC (2017) **Noise-robust fixation detection in eye movement data: Identification by two-means clustering (I2MC)** *Behav Res Methods* **49**:1802–1823

- Houthoofd R, Chen X, Duan Y, Schulman J, De Turck F, Abbeel P (2016) **VIME: Variational Information Maximizing Exploration** *arXiv [csLG]*
- Hunnius S (2022) **Early cognitive development: Five lessons from infant learning** *Oxford Research Encyclopedia of Psychology* <https://doi.org/10.1093/acrefore/9780190236557.013.821>
- Hunnius S, Geuze RH (2004) **Developmental Changes in Visual Scanning of Dynamic Faces and Abstract Stimuli in Infants: A Longitudinal Study** *Infancy* **6**:231–255
- Hwang Y, Jeong S-H (2023) **Misinformation Exposure and Acceptance: The Role of Information Seeking and Processing** *Health Commun* **38**:585–593
- Jackson I, Sirois S (2009) **Infant cognition: going full factorial with pupil dilation** *Dev Sci* **12**:670–679
- James DK (2010) **Fetal learning: a critical review** *Infant Child Dev* **19**:45–54
- Johnson SP, Hannon EE (2015) **Perceptual Development** *Handbook of Child Psychology and Developmental Science* <https://doi.org/10.1002/9781118963418.childpsy203>
- Joshi S, Gold JI (2020) **Pupil Size as a Window on Neural Substrates of Cognition** *Trends Cogn Sci* **24**:466–480
- Kemp C, Perfors A, Tenenbaum JB (2007) **Learning overhypotheses with hierarchical Bayesian models** *Dev Sci* **10**:307–321
- Kirkham NZ, Slemmer JA, Johnson SP (2002) **Visual statistical learning in infancy: evidence for a domain general learning mechanism** *Cognition* **83**:B35–42
- Krogh L, Vlach HA, Johnson SP (2012) **Statistical learning across development: flexible yet constrained** *Front Psychol* **3**
- Lake BM, Salakhutdinov R, Tenenbaum JB (2015) **Human-level concept learning through probabilistic program induction** *Science* **350**:1332–1338
- Lavín C, San Martín R, Rosales Jubal E (2013) **Pupil dilation signals uncertainty and surprise in a learning gambling task** *Front Behav Neurosci* **7**
- Lewkowicz DJ, Hansen-Tift AM (2012) **Infants deploy selective attention to the mouth of a talking face when learning speech** *Proc Natl Acad Sci U S A* **109**:1431–1436
- Makowski D, Lüdecke D, Ben-Shachar MS (2020) **Modelbased: Estimation of model-based predictions, contrasts and means**
- Mathot S, Vilotijevic A (2022) **Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis** *Behav Res Methods* <https://doi.org/10.3758/s13428-022-01957-7>
- Orpella J, Mas-Herrero E, Ripollés P, Marco-Pallarés J, de Diego-Balaguer R (2021) **Language statistical learning responds to reinforcement learning principles rooted in the striatum** *PLoS Biol* **19**
- Pathak D, Agrawal P, Efros AA, Darrell T (2017) **Curiosity-Driven Exploration by Self-Supervised Prediction** **2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)** :488–489



- Pedersen EJ, Miller DL, Simpson GL, Ross N (2018) **Hierarchical generalized additive models: an introduction with mgcv** *PeerJ* <https://doi.org/10.7287/peerj.preprints.27320>
- Poli F, Ghilardi T, Mars RB, Hinne M, Hunnius S (2023) **Eight-month-old infants meta-learn by downweighting irrelevant evidence** *Open Mind* **7**:141–155
- Poli F, Serino G, Mars RB, Hunnius S (2020) **Infants tailor their attention to maximize learning** *Sci Adv* **6**
- Pollard JW (1984) **Evolutionary Theory: Paths Into the Future**
- Preuschoff K, 't Hart BM, Einhäuser W (2011) **Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making** *Front Neurosci* **5**
- Saffran JR, Aslin RN, Newport EL (1996) **Statistical learning by 8-month-old infants** *Science* **274**:1926–1928
- Sodian B, Thoermer C (2004) **Infants' Understanding of Looking, Pointing, and Reaching as Cues to Goal-Directed Action** *J Cogn Dev* **5**:289–316
- Sutton RS, Barto AG (2012) **An Reinforcement Learning: Introduction**
- Thurman SM, Cohen Hoffing RA, Lauharatanahirum N, Forster DE, Bansal K, Grafton ST, Giesbrecht B, Vettel JM (2019) **Applying linear additive models to isolate component processes in task-evoked pupil responses** *J Vis* **19**:305–305
- Tummeltshammer KS, Wu R, Sobel DM, Kirkham NZ (2014) **Infants track the reliability of potential informants** *Psychol Sci* **25**:1730–1738
- van Rij J, Hendriks P, van Rijn H, Baayen RH, Wood SN (2019) **Analyzing the Time Course of Pupillometric Data** *Trends Hear* **23**
- Vehtari A, Gelman A, Gabry J (2017) **Practical Bayesian model evaluation using leave-one-out crossvalidation and WAIC** *Stat Comput* **27**:1413–1432
- Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P (2020) **Rank-normalization, folding, and localization: An improved R-hat for assessing convergence of MCMC** *Bayesian Anal*
- Werchan DM, Collins AGE, Frank MJ, Amso D (2016) **Role of Prefrontal Cortex in Learning and Generalizing Hierarchical Rules in 8-Month-Old Infants** *J Neurosci* **36**:10314–10322
- Werchan DM, Collins AGE, Frank MJ, Amso D (2015) **8-month-old infants spontaneously learn and generalize hierarchical rules** *Psychol Sci* **26**:805–815
- Westermann G, Thomas MSC, Karmiloff-Smith A (2010) **Neuroconstructivism** *The Wiley-Blackwell Handbook of Childhood Cognitive Development* <https://doi.org/10.1002/9781444325485.ch28>
- Yuan L, Xiang V, Crandall D, Smith L (2020) **Learning the generative principles of a symbol system from limited examples** *Cognition* **200**
- Zenon A (2019) **Eye pupil signals information gain** *Proc Biol Sci* **286**

Zmyj N, Buttelmann D, Carpenter M, Daum MM (2010) **The reliability of a model influences 14-month-olds' imitation** *J Exp Child Psychol* **106**:208–220

## Article and author information

### Tommaso Ghilardi

Centre for Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London, London, United Kingdom

**For correspondence:** t.ghilardi@bbk.ac.uk

### Francesco Poli

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

### Marlene Meyer

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

### Olympia Colizoli

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

### Sabine Hunnius

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

## Copyright

© 2024, Ghilardi et al.

This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Editors

Reviewing Editor

### Miriam Spering

The University of British Columbia, Vancouver, Canada

Senior Editor

### Michael Frank

Brown University, Providence, United States of America

### Reviewer #1 (Public Review):

This study investigates the underlying mechanisms of information-seeking in infancy. Eight-month-old Dutch infants were tested in a screen-based eye-tracking task in which one of two geometrical shape cues (differing in their shape and motion) either announced the location of an upcoming reward cartoon (informative) or not (non-informative). The authors measured the infants' pupil size before the cartoon appeared. Infants showed smaller pupil sizes when

presented with the informative cue as compared to the noninformative cue. The decrease in pupil size in the informative condition emerged over the course of trials whereas infants' pupil size remained unchanged in the noninformative condition. The authors interpret their findings as supportive evidence of statistical learning and generalization processes organizing infants' information-seeking.

It was a pleasure to read the paper and I think the study makes a valuable contribution to our understanding of information-seeking in infancy. The manuscript is very well written and the study is cleverly designed. My following comments are based on my reading of the manuscript and the supplemental materials. It should be noted that evaluating the details of the statistical procedure the authors used lies outside my expertise. The same applies to some decisions of the authors related to pre-processing and filtering the pupil data. I very much appreciate that the authors shared all their raw data and analysis scripts openly accessible on the Open Science Framework. The study was unfortunately not preregistered, making it difficult to trace when in the study process certain decisions or assumptions were made.

My two main concerns relate to the conceptualization and definition of information-seeking and the proposed speed of the mechanisms explaining infants' behavior. I outline my general comments below before listing some more concrete issues.

1. While reading the manuscript, I was sometimes confused about what the authors refer to when talking about information-seeking - both in terms of the broader conceptualization of the phenomenon as well as when referring to their own study. What information are infants seeking? The informative value of the cue shape in terms of their motion (because it carries information about the location of a rewarding animation)? Or is the target (the rewarding video) the information being sought? From how the study is set up, I assume the authors refer mainly to the first aspect, but I think the manuscript would benefit from some clearer distinctions and definitions of terms.

More specifically, I think it could help if the authors would specify the different aspects involved in information-seeking in the introduction (e.g., seeking information "directly", seeking cues guiding them towards information, etc.). Secondly, it would help if they would sharpen their (already in some parts existing) definitions for their study and then keep consistent with their definitions throughout the methods, results, and discussion. Is the cue the information being sought or the "behavior" (motion) of the cue? Or is the target animation the information being sought and guided via the cueing?

1. Speed of the generalization process:

From my understanding of the study design, the shape of the geometrical shape gains informative value over time (serving as an informative cue) and the \*motion\* of the shape is the actual informative or non-informative visual cue in that it either reliably highlights the actual target region (or all regions). In the generalization trials, only the shape was manipulated while the motion aspect remained consistent with the previous trials. Based on infants' behavior across learning and generalization trials, the authors make an argument about two distinct processes taking place: a slower allowing to learn where to find info and a faster generalization process. Apologies if I missed something, but given that the motion remains consistent, it's maybe not surprising that the generalization trials are "faster"? Maybe the generalization process would have been slower if not only the shape had changed but if also a novel informative motion had been introduced. Also, it would be helpful if the authors could clarify what they mean by the statistical learning process being more "data-hungry" (line 274).

2. I would find it very helpful if the authors would discuss statistical learning and information-seeking processes from other possible mechanisms such as reward learning mechanisms. For example, the authors use a "rewarding" (not informative) stimulus as the target-wouldn't it be possible that the results can be also explained by reinforcement learning processes? Relatedly, in line 396 they write that they used TD learning to predict whether "information will be delivered" and contrast this with the approach being used to predict whether a reward will be delivered. But in their study reward was being delivered, too (in the form of the target), in addition to the informative motion of the cue.

- <https://doi.org/10.7554/eLife.92388.1.sa2>

## Reviewer #2 (Public Review):

### Summary

The study used eye tracking with a focus on pupillometry to examine how infants can learn to distinguish between informative and uninformative visual cues. Infants ( $n = 30$ , mean age = 8.2-months-old) viewed displays consisting of a sequence of stimuli: a fixation point, a central cue that predicted a subsequent informative or uninformative signal, the signal itself, and the target event (a cartoon animal, referred to as the reward). The key results are that: (1) pupil size differs depending on whether the infants anticipated an informative or uninformative signal, (2) this difference develops across trials, consistent with a slow learning process, and (3) there is rapid generalization when new shapes were introduced that shared features with the informative vs uninformative cues. The study complements a rich literature, including from this same group, showing that children are sensitive to information gains, and is interesting and important in revealing that pupil size is a physiological marker of information anticipation. We have several comments and concerns and believe that addressing them would substantially strengthen the manuscript.

Major points are related to interpretation, statistical robustness, and clarity

1. There is a tendency to overinterpret the findings.

a. Throughout, the authors interpret the findings as meaning that pupil size tracks the "value" of information; however, the results do not demonstrate conclusively whether, or what kind of value information has in this task. A natural hypothesis is that infants are intrinsically motivated to predict - i.e., value the ability to predict the target event as early as possible. In a supplementary figure, the authors present evidence that infants indeed fixate on the target event sooner after seeing informative vs uninformative cues, consistent with the idea that they use the information for improving predictions. However, those results are not fully convincing, as we detail in point 2. Most importantly, the analysis is not integrated or even mentioned in the main analyses analysis. Making the link between the pupil reaction and the use of the information would greatly strengthen the paper (whether or not the supplementary findings hold up to more thorough scrutiny). Either this link should be made and discussed, or the authors should soften their conclusions about the utility of the informative cues.

b. On line 236, the text states that the evidence "...supports the growing body of evidence indicating that infants are proactive in shaping their learning environment by searching for and focusing on information-rich stimuli". The results do not show that the infants search for information, only that they have a pupil reaction that differentiates between informative and uninformative stimuli.

c. On lines 248-249, it seems a stretch to relate the changes in pupil dilation to a shift in information value onto the cue. Without some other measure (e.g., EEG), this remains speculative. While I believe the suggestion is plausible, the language should be softened to

highlight this as a follow-up research question that the present research cannot directly speak to.

## 2. Several findings are statistically weak and several analyses are insufficiently controlled.

a. The analysis in Supplementary Figure 2, which shows that the latencies of target fixations are shorter after informative vs uninformative cues, raises several questions.

- i. We were unable to fully test these analyses as the OSF project seems to only contain latency data for 33 participants (including 22 of the 30 that remain in the final sample).
- ii. The results are described as revealing a significant difference, but the 89% confidence interval of the difference contains 0. How did the authors establish significance here?
- iii. How do the authors distinguish incidental fixations (which just happened to land near the target) from true predictive gaze shifts? Fixations were pooled if they occurred from 1.25 seconds before to 1 second after target onset. This is sufficient time for the eye to move in and out of the window several times. The authors should analyse the distributions of fixation durations to rule out various artifacts unrelated to target prediction.
- iv. Latencies to fixation were standardized, bringing the mean across each participant to 0, and yet the statistical model includes a random intercept; is there a justification for this?
- v. Standardizing removes information about whether fixations were proactive or reactive. It would be very interesting to see if/how information affects these two differently.
- vi. Since informativeness was learned across trials, it seems desirable that the model should include as random effects a trial number and an interaction between trial number and informativeness. This would allow a comparison between learning to predict and the pupil reaction. Are infants who have a stronger (or earlier) pupil reaction also more likely to show stronger learning to anticipate?

b. The main finding that pupil size differs between informative and uninformative cues is based on a 3-second analysis window. This long window most likely spans many saccades, which can affect pupil size on its own or by bringing the eye on or off visual stimuli. There is no analysis to show that the statistics of saccades or fixation locations are equivalent between the two trial types - but this is necessary to convincingly rule out a spurious artifact.

c. The second main finding that the effect of informativeness grows across trials seems statistically weak. The text (line 138) states that the interaction had a beta of 0.002, which was equal to the lower border of the 89%HDI ([0.002, 0.003]). For the second claim that pupil size decreased across informative trials, the beta is -0.002, and 89% HID is non-existent - i.e., [-0.002, -0.002]. (In general, the authors should check their numbers more carefully and make sure they are presented with a degree of precision that allows the reader to interpret them meaningfully.

d. The analyses do not indicate how well the TD model fits; we are shown only that it fits better than a linear model. On line 177 a correlation analysis is mentioned between the data and model, but the statistic cited for this test on line 179 is a mean beta coefficient, so it is impossible to know what this means. An analysis of goodness of fit or, at the very least, a figure superimposing the model and data, would be much more convincing.

## 3. The descriptions are very unclear in some key parts of the paper

a. The description of the TD model applied to pupil learning (starting on line 391) is very unclear. The model has to include some measure of informativeness - i.e., the match between the cued and true target location - but it is unclear how this was formalized. It is also very unclear how time within the trial is incorporated (the meaning of the TDE equation).

b. The description of the generalization analysis (Fig. 5) is also very unclear. Every single sentence in it evoked some confusion, so I will go through them one by one. "A Bayesian additive model showed that infants' pupil dilation was reduced for novel cues." Reduced relative to what? "This was specific to those novel cues that shared the features of the

familiar informative cues (estimated mean difference = -0.05, 89%HDI = [-0.062, -0.038])." All the novel cues shared features with the informative cues; do the authors mean the novel cues that had the critical feature indicative of the informative cue? "The size of this effect approximated the difference between conditions that were observed for familiar stimuli (estimated mean difference = -0.067, 89% HDI = [-0.201, -0.077])." What is "this effect"? "Crucially, this difference was not observable at the start of the task, when the familiar stimuli were first introduced (estimated mean difference = -0.007, 89%HDI = [-0.015, 0.001])." At the start of the task, the stimuli were novel, and not familiar.

- <https://doi.org/10.7554/eLife.92388.1.sa1>

### Reviewer #3 (Public Review):

#### Summary:

The study attempts to shed light on the mechanisms underlying information-seeking in infants by investigating whether infants distinguish between informative and uninformative stimuli to resourcefully allocate their attention. The authors show that 8-month-old infants can learn whether a visual stimulus is informative or uninformative about the location of a later appearing rewarding stimulus by employing statistical regularities from the input. Specifically, infants showed decreased pupil dilation for informative over uninformative cues, which developed over the course of trials as more and more information was gathered from the input. The pattern of learning was in line with a reinforcement learning model which employed a steep learning curve in the beginning followed by a more shallow but steady learning growth over trials. After 17 trials, the authors presented novel cues that shared certain visual features with the previous stimuli and showed that pupil dilation was reduced for novel cues that shared features with the previous informative stimuli, suggesting that infants were able to generalize their acquired knowledge about the informativeness of certain features to novel stimuli. The present study adds to the existing literature about the underlying mechanisms of learning by showing that infants cannot only predict an upcoming stimulus based on statistical regularities of a preceding cue but also the informativeness of the cue itself.

#### Strengths:

The authors use a suitable method to test the highly relevant question of whether and how infants infer the informativeness of stimuli from experience and whether they can generalize this knowledge to new stimuli. Their experiment is carefully designed and well controlled with conditions closely matched (e.g., the shape and color of objects and the structure of each trial). Their measure of interest (i.e., pupil dilation) is also examined at a time point in each trial when the conditions are the most similar, which further points to a thought-through and careful design. This empirical data is backed up with a computational approach (using a Bayesian model and training a reinforcement learning algorithm) to elucidate the learning mechanisms at play. This approach is explained concisely to readers not familiar with the models.

The results are convincing showing a clear difference between informative and uninformative condition and development over trials. Specifically, this difference is not apparent in the first trial (Fig. 2c) but develops over time which supports a learning trajectory. The data support the authors' conclusion that infants learn about the informativeness of the object cue from the input, and the employed learning algorithms give further insights into the learning trajectory of the infants. Overall, the statistical analyses seem solid and the priors for the Bayesian models are well reported.

Data and scripts are openly available fostering transparency.

Overall, the manuscript is very well and concisely written.

#### Weaknesses:

The authors' conclusion that infants can generalize the acquired knowledge to similar but novel stimuli is weakened by methodological concerns regarding the analysis. It is not fully clear which trials the authors excluded and analyzed as they do not consistently report the trials in the manuscript (e.g., it is stated that after trial 17 the first generalization trial started, but also that trial 17 was excluded as the first trial of the generalization phase). As there are only a few novel trials and novel and familiar trials alternated, the inclusion or exclusion of trial analyses might have a significant impact on the results. Thus, this needs further clarification. The authors also mentioned that the novel stimuli shared relevant as well as irrelevant features, but it was not clear to me whether the authors could establish that only the relevant features contributed to the observed generalization effect.

Some methodological decisions were not explained and need justification, in particular, as the study is not preregistered. This includes, for example, the exclusion criteria and the choice not to analyze all generalization trials. Further, the authors did not perform model comparison (e.g., their model against a null model) and therefore do not report the strength of evidence for a difference in conditions.

Another weakness is that the sample sizes of 30 infants for the initial part and 19 infants for the generalization part of the experiment are rather small (especially with regard to the chosen weakly informative priors).

- <https://doi.org/10.7554/eLife.92388.1.sa0>