Article

Seeing the Error in my "*Bayes*": A Quantified Degree of Belief Change Correlates with Children's Pupillary Surprise Responses Following Explicit Predictions

Joseph Colantonio[1,2,*], Igor Bascandziev[1], Maria Theobald[3], Garvin Brod[3] and Elizabeth Bonawitz[1]

[1] Graduate School of Education, Harvard University, Cambridge, MA 02138, USA
[2] Psychology Department, Rutgers University – Newark, Newark, NJ 07102, USA
[3] DIPF | Leibniz Institute for Research and Information in Education, Rostocker Str. 6, 60323 Frankfurt am Main, Germany
* Correspondence: colantoniojoseph@gmail.com

Abstract: Bayesian models allow us to investigate children's belief revision alongside physiological states like "surprise". Recent work finds that pupil dilation (or the "pupillary surprise response") following expectancy-violations may be predictive of belief revision. How can probabilistic models inform interpretations of "surprise"? Shannon Information considers the likelihood of an observed event, given prior beliefs – suggesting stronger surprise occurs following unlikely events. In contrast, Kullback-Leibler divergence considers the "dissimilarity" between prior beliefs and updated beliefs following observations – with greater surprise indicating more change between belief states to accommodate information. To assess these accounts under different learning contexts, we use Bayesian models that compare these computational measures of "surprise" to contexts where children are asked to either predict or to evaluate the same evidence during a water displacement task. We find correlations between the computed Kullback-Leibler divergence and children's pupillometry responses only when children actively make predictions, and no correlation between Shannon Information and pupillometry. This suggests that when children attend to their beliefs and make predictions, pupillary responses may signal the degree of divergence between a child's current beliefs and updated, more accommodating beliefs.

Keywords: Bayesian Inference; Cognitive Development; Learning; Prediction; Pupil Dilation; Science Learning; Surprise

1. Introduction

To no surprise, understanding the process of belief revision is of great interest and has a rich history in many fields including philosophy, psychology, education, and computer science (e.g. [1-4]). Psychological and philosophical work suggests that two interrelated components of human intelligence are the ability to deploy abstract, causal, "intuitive theories" to support inference and the ability to revise these theories in light of evidence [3,5,6]. Contemporary approaches in the Cognitive Sciences align empirical work with computational implementations, typically finding that Bayesian models can provide a framework with which to understand human inference from, and learning of causal beliefs [7-11]. These models provide an account of how learners can draw rich inferences relatively rapidly even when data is limited or ambiguous and have been extended to account for the ways in which learners form and revise more abstract intuitive theories as well [12-18]. However, until recently, less work has investigated epistemic emotions and physiological expressions as they relate to rational models of human learning, despite the well-established connection between these arousal states and learning [19-20]. In fact, Bayesian models provide a means to not only understand how humans draw rich inferences from limited data and revised intuitive theories, but also to compare human physiological responses to competing computational theories of surprise and learning.

A large body of literature highlights the importance of affective and physiological states for learning and cognition in general. Physiological states, such as pupil dilation, are often accompanied by phenomenological affective states, like surprise [21-22]. This is why many researchers studying the effects of surprise on cognition rely on objective physiological measures – such as pupil dilation – as a proxy for surprise [23-28]. However, *how* these physiological states relate to learning via belief revision remains less well understood. This challenge of determining *what* factors are closely linked to concept learning and *how* they affect learning is critical to address, as understanding these specific factors themselves provides multiple positive outcomes for research. Thus, doing so computationally may improve our understanding of belief revision while also improving our ability to design human-inspired learning agents.

In the current study, we look to extend Bayesian learning models for investigating the potential relationships between the physio-emotional experience of surprise (as indexed by pupil dilation) and learning. Specifically, we contrast two predictive models related to learning: "Shannon Surprise" and "Kullback-Leibler divergence" belief updating. By building specific predictive models and relating them to children's physiological responses (via pupil dilation), we can better understand the mechanisms that underlie learning in different contexts. Specifically, we will investigate correlations between these two models and children's pupillary surprise as they perform belief revision during a water displacement learning task under different conditions. In one condition, children are asked to predict outcomes prior to observing events (engaging their prior beliefs) and in another, children make *post hoc* evaluations of the same evidence. By evaluating these different types of models and their fit to physiological behavior in these two conditions, we can better understand how different contexts might engage the interplay between cognitive and physiological mechanisms that support learning.

In what follows, we discuss the measure of pupil dilation and *what* pupil dilation indicates. Next, we describe scenarios where pupil dilation may most likely be elicited and more strongly linked to belief revision, namely when making predictions. Then, we investigate two candidates for computationally estimating the pupillary surprise response based on empirical findings and their theoretical interpretations. First, Shannon Information as a data-driven surprise; second, Kullback-Leibler divergence as a belief-driven surprise. Thus, we aim to face the specific challenge of understanding *how* the pupil dilation response as a cognitive-behavioral response relates to learning via belief revision in our tasks.

1.1 The Pupil Dilation Response, Attention & Learning

Pupil dilation holds a special status among multiple connected fields such as psychology, cognitive science, neuroscience, biology, and computer science. This is because pupil dilation has for a long time been considered a reliable instrument for identifying the temporal dynamics of arousal [29-32]. More recently, pupil dilation has been considered a physiological response that represents an integrated readout of an attentional network containing multiple contributing factors [33,34]. Within this attentional network, recent work suggests that pupil dilation in this network may occur as a result of an interactive cascade among varied components, including low-level (e.g., light and focal distance; [35,36]), intermediate-level (e.g., alerting and orienting; [37-39]), and high-level factors (e.g., physio-emotional responses, inference, and executive function; [25,33,40]). Overall, accounts of pupil dilation as an attentional indicator highlight that pupillometry can broadly be attributed to either directed attention or higher-level sensory operations for processing the content that the observer is currently perceiving.

However, it remains unclear whether these discussed attentional factors and their related processes are *what* pupil dilation is expressing *specifically* in relation to learning. Further – if so, whether some, none, or all of these factors are being expressed in the same fashion or to the same degree during belief revision. That is, we know quite a bit about what might *elicit* pupil dilation during learning scenarios (e.g., violations of expectations;

[25,26, 28,41]), but less about what the *processes* coupled with pupil dilation actually are and the implications of said processes. Thus, we propose designing computational Bayesian models of learning that can potentially estimate the degree of surprise experienced by learners, relative to their pupil dilation measurements during a learning task.

With these Bayesian models, we will contrast two broader accounts of "surprise" that may help to clarify the relevance of this physiological marker of belief revision. The first candidate, originating from research on Information theory (i.e., Shannon Information; [42]), posits that surprise (and thus, pupil dilation) correlates with objective expectations of the data and how informative it is given the data's likelihood. The second candidate highlights divergence and dissimilarity (i.e., Kullback-Leibler divergence; [43]) between what is believed by a learner and what revised beliefs the learner expects to better accommodate incoming data, quantifying the degree of belief change needed to correctly represent the actual outcome of a given event by transforming the prior belief into the appropriate posterior.

In fact, recent work has looked into disentangling the pupillary surprise response as separable, distinct processes that can be represented computationally by Shannon Information and Kullback-Leibler divergence. One study by O'Reilly and colleagues [44] performed a combined brain imaging and pupillometry study where participants completed a saccadic eye movement response task. Here, participants needed to use their prior knowledge about a spatial distribution to locate a target (a colored dot) before returning to a fixation cross. The findings showed that there were separate, specific neural signals associated with pupil dilation acting as temporal indicators of surprise (within the posterior parietal cortex) and belief revision (within the anterior cingulate cortex). Specifically, less-likely events were considered more surprising via Shannon Information, and elicited pupil dilation. Meanwhile, they found that the Kullback-Leibler divergence related to when pupil diameters decreased on trials when belief updating may be occurring. This work provides important demonstration of the dissociable roles of Shannon Surprise and Kullback-Leibler divergence in computationally capturing surprise and belief updating, respectively, using a Bayesian framework.

Similarly, Kayhan et al. [45] investigates pupillary surprise and learning with 18-month-old infants and 24-month-old toddlers. Here, young children completed a statistical learning task that measured their pupil dilation as they viewed movies where an agent sampled five colored balls from a transparent bin containing multiple balls of two colors. These bins depicted the distribution of ball colors inside of it (e.g., a majority of yellow balls (80%) and minority of green balls (20%)). Critically, 24-month-olds' (but not 18-month-olds') pupillary responses followed a pattern similar to the prediction error of a causal Bayesian model, calculated as the Kullback-Leibler divergence between prior and updated probability distributions.

Thus, inspired by these exciting results, we designed a study that lets us explore further nuances of how different contexts (asking children to predict vs *post hoc* evaluate outcomes) might engage the cognitive mechanisms associated with these two different accounts of surprise. This provides a means to explore the relationship between behavioral results that find differences in learning via different interventions with the physiological response and potential cognitive mechanisms (surprise vs belief updating) that might underlie them.

In what follows, we first describe these two potential mechanisms of pupil dilation, and highlight key theoretical differences between their interpretations and implementations. Then, we will describe specific contexts where these proposed mechanisms of pupil dilation may be most prevalent, via model-based prediction, as highlighted by a significant amount of recent empirical research. Next, we provide a brief description of the probabilistic Bayesian model used and what metrics we are investigating from it. Finally, we will compare the two estimates of surprise – Shannon Information and Kullback-Leibler divergence – based on their correlations with children's pupil dilation during a water displacement learning task.

2. Competing accounts of Surprise: Shannon and Kullback-Leibler

2.1. Estimating Pupil Dilation as Data-driven via Shannon Information

Shannon Information is a well-known metric in information theory and describes how informative an outcome is [42,46-49]. It is largely found in machine learning literature to describe computational "surprise" - quantifying how meaningful incoming data is relative to a specific target despite other unwanted, noisy interference. When interpreted with respect to learning (via Bayesian inference), Shannon Information can be used to describe the "unexpectedness" of incoming data given the prior beliefs of the learner. Computationally, Shannon Information can be calculated as the negative log-likelihood of some data's probability, $p(d)$, given some beliefs over models of the world (H), where Shannon Information Surprise (Eq. (1)) is

$$Shannon\ Information\ =\ -log(p(d)). \tag{1}$$

Shannon Information for some incoming data given an inferred model is typically quantified as a "signal" of information. Information theory captures this intuition as simply the negative log probability of the data. Note that this is computationally the same as marginalizing out hypotheses by considering the probability of the data given each hypothesis in space H, weighed by the prior probability of each hypothesis, h. One might interpret Shannon surprise psychologically as a violation of expectation, which depends on comparing the observation to a prior prediction of outcome likelihoods given the weighted set of prior beliefs.

If Shannon Information correlates more strongly with children's pupillometry compared to its competitor, the Kullback-Leibler divergence, then we posit that perhaps the pupillary surprise response may be more "objective" or "external-focused", acting as a reaction to acknowledge the unexpectedness of an event that has occurred and draw attention to it. Specifically, "surprise as information" would represent an attentional mechanism homed in on incoming data – emerging as a sign to alert the learner and re-orient (or heighten) their attention; a process of an "intermediate-level" of complexity among cognitive responses (per recent review of pupillometry research [34]). Thus, finding that Shannon Information best fits pupil responses may indicate a response akin to prediction error, as typically associated with surprise during violation of expectation events.

2.2. Estimating Pupil Dilation as Belief-driven via Kullback-Leibler Divergence

In contrast, other computational accounts describe pupil dilation and surprise in regard to how effectively new data "transforms" a learner's prior beliefs into their posterior beliefs [50,51]. Here, the summed Kullback-Leibler divergence is considered the second candidate for estimating surprise, measuring the summed dissimilarity or relative entropy between a learner's distributions of prior and posterior beliefs, given the observation of some new data [43,52]. Computationally, the Kullback-Leibler divergence for models considering multiple, competing hypotheses is calculated (Eq. (2)) as the weighted log-odd ratio between a posterior belief, $p(h\,|\,d)$, and prior belief, $p(d)$, summed across hypotheses within the set of hypotheses considered ($h \epsilon H$)[1]:

$$Kullback - Leibler\ Divergence = \sum_{h \epsilon\ H} p(h|d) \log\left[\frac{p(h|d)}{p(h)}\right]. \tag{2}$$

[1] A symmetric (and finite value) form of Kullback-Leibler divergence (Jensen-Shannon) can also be used to compute distance. In the computational analyses that follow, we apply standard Kullback-Leibler divergence, but results are not qualitatively different if the Jensen-Shannon divergence is used instead.

As mentioned, Kullback-Leibler divergence calculations describe not simply a distance between distributions, but a measure of dissimilarity between them. Thus, when describing belief revision processes, Kullback-Leibler divergence can be considered as how much *'work'* is needed to affect an initial probability distribution (e.g., one's prior beliefs) in a way that changes it into another (e.g., updated posterior beliefs). Here, if we find that Kullback-Leibler divergence relates to learning responses, then we believe that pupil dilation may be a more "subjective" physiological marker of learning that follows from the belief updating process.

Central to our empirical question, this computational approach will allow us to contrast different models of "surprise" when learning. Specifically, Shannon Information will quantify the probability of the data accumulated by learners trial-by-trial. Here, Shannon Information might be depicting pupil dilation as a temporal indicator of when children may be alerted to an unexpected, highly "informative" outcome that the child should orient themselves toward. Meanwhile, Kullback-Leibler divergence will quantify the dissimilarity between a child's prior beliefs and what inferred models of the world would best explain potential outcomes. This means that Kullback-Leibler divergence presents pupil dilation as a physiological signal of the amount of effort needed to update their beliefs (given the learner's current belief distribution and the to-be posterior belief distribution that best explains the new data).

2.3 Model-Based Learning through Prediction

Asking learners to generate predictions is a popular method for improving children's learning. Studies investigating prediction generation (or "hypothesis generation") in children tend to find that explicitly predicting an outcome before seeing it improves learning (e.g., of physics; [53-55]). The benefits of making predictions have been connected to successful activation of prior knowledge when learning new material, but less is known about the specific mechanisms by which predicting affects learning success, in particular when it comes to theory revision [56]. Understanding the cognitive processes that are engaged during prediction generation can help us understand how, why, and when these interventions are likely to be successful.

Experiments on making predictions that investigate pupil dilation and learning find that actively generating a prediction compared to making *post hoc* evaluations increases the degree of pupil dilation, particularly when faced with events that are predicted incorrectly [25,57]. Furthermore, this work has found a positive relationship between the degree of pupil dilation and successful belief revision [27,28,40]. The enhanced pupillary surprise response after a violation of expectations may be due to children activating some task-relevant prior knowledge when they generate a prediction (i.e., leveraging their prior beliefs). Further, if the outcome following a prediction is different from what the learner expects, then conflict awareness may be heightened and increase the subjective value of the outcome's informativeness, which facilitates belief revision.

We believe that with all other things equal, making a prediction may give children an "edge" over their peers and promote their learning by engaging cognitive mechanisms associated with surprise. Assessing this prediction depends on two measures. First, it requires building models for individual learners that predict computationally when surprise is highest given the learners beliefs and the observed evidence. Relating these model predictions to physiological markers like pupillometry helps us understand the computational and potentially mechanistic basis for pupil-marked surprise in learning. It also allows us to contrast competing computational markers of surprise under different learning contexts. Second, we can relate the degree to which individual children's physiological states are correlated with these quantitative models and predict that children who have better "alignment" between physiological and model based surprise may also be more "optimal" learners, in the sense that their learning behavior is better matched to idealized learning models. That is, if a heightened, "rational surprise" response leads to more efficient learning, then children who experience surprise when a rational model (e.g., a probabilistic Bayesian model) would expect them to, may also be better simulated by said rational model, as well. Thus, the second hypothesis in our investigation is that children whose pupillometry measures are better fit by the model estimates are also more strongly represented by the simulated behaviors of an ideal learner, depicted by an Ideal Bayesian learning model.

3. The Current Study

The broad hypothesis of this paper is that children that engage more with a learning task by making predictions will have stronger correlations between their pupil dilation measurements and the model estimates of the pupillary surprise response, compared to peers that are only making *post hoc* evaluations (specific regarding the modeled data to be described, below). However, two alternative hypotheses are also considered here regarding *which* of the model estimates better fits their responses. Recent interpretations of prediction suggest that actively making a prediction entails leveraging one's prior beliefs and extrapolating potential outcomes given these beliefs (e.g., [11,27,28,40]). Both Shannon Information and Kullback-Leibler divergence accounts are consistent with this proposal because they both leverage prior beliefs towards predictions. However, they differ in the mechanism (and potential) implications of leveraging those beliefs. If the pupil measurements for children making predictions are better matched by the Shannon Information metric, then this suggests that pupil dilation may indicate more robust engagement with the feedback they receive. In particular, good performance of the Shannon Information estimate may represent children’s heightened attention to evidence that violates their beliefs (e.g., [47-49]). Such heightened response could support later learning by increasing arousal and thus improve the encoding of surprising data, but the Shannon response does not reflect the learning in the moment. However, if the Kullback-Leibler divergence instead performs better than Shannon Information, we would find support for physiological responses capturing belief-updating in the moment, suggesting that children may be performing an effortful computation that captures degree of belief change. Critically, assessing the performance of these candidate metrics of quantifiable pupillary surprise – both in general and in competition with one another – helps us better understand the role of surprise during belief revision. Does surprise simply serve to guide attention to relevant outcomes? Or does it aid learners by highlighting their beliefs and inform their integration of new information?

We modeled data from an experiment that investigated elementary school (six- to nine-year-old) children’s theories of water displacement for the current model (experimental procedure, data, and empirical results are those found in [40]. Children’s causal beliefs of water displacement were chosen as children frequently have the misconception that water displacement depends on the weight of an object or a combination of weight and size rather than on its size only (e.g., [58]), providing an appropriate domain for the investigation of variability across individual children’s beliefs, as well as their impact on children’s subsequent learning. Furthermore, previous work has modeled this experimental data for investigation of children's learning during a belief revision task [11] and found very strong fits between “optimal” Bayesian learning and children’s performance on the task.

The to-be-modeled experiment’s design in [40] entailed a Pretest phase, a Learning phase, and a Posttest phase. On each trial, children (regardless of assigned condition) were presented with two spheres of varied features (e.g., in size, material, and/or weight) side-by-side (see Figure 1 for a trial example). Then, children stated which sphere they thought would displace the most water (between two identical containers). These judgments were assessed using a 5-point-scale (e.g., (1 = *certainly the left sphere*, 2 = *maybe the left sphere*, 3 = *equal amounts of water for both*, 4 = *maybe the right sphere*, 5 = *certainly the right sphere*). During the Pretest and Posttest phases, no feedback was provided to the children to allow for a clean initial assessment of beliefs (prior to learning) and final learning outcomes. Children were randomly assigned to one of two experimental conditions – a Prediction or a Postdiction Condition, and were provided feedback during the Learning Phase of the experiment according to condition. Children in the Prediction condition were asked to provide a response *prior to seeing the outcome*; responses were values from 1-5 stating their expectation (and confidence) about which sphere displaces more water. In contrast, children in the Postdiction condition first saw the results of the presented trial, then were asked to state what their expectations had been (*prior to the evidence*)[2]. Importantly, children's pupil dilation measurements were collected as outcomes were presented during the Learning phase for both conditions.

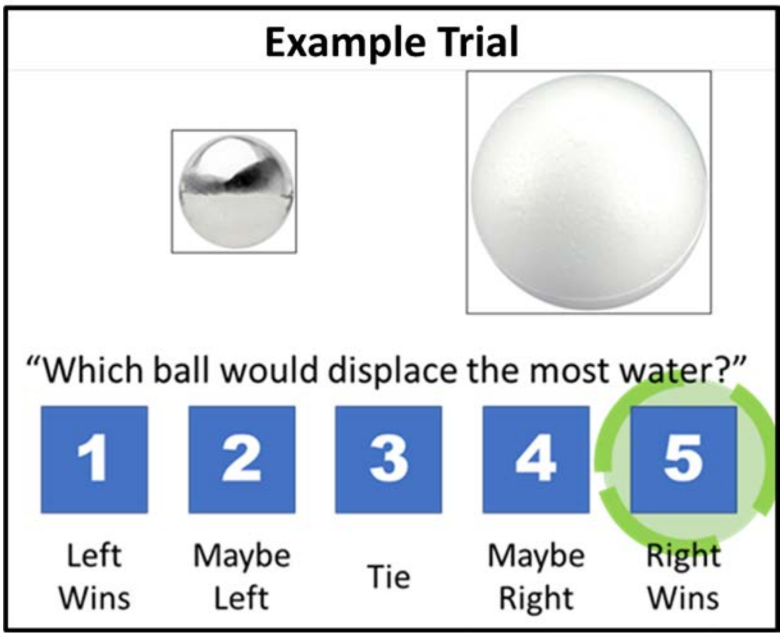


Figure 1. An example of a trial during the original experiment. Here, the correct response for the trial example is option "5 - Right Wins", highlighted by the green dashed circle. This is provided as evidence following children's response (Prediction Condition) or preceding their response (Postdiction Condition). Children with the correct "Size" rule would accurately select "5" (or "4") here and see confirming feedback. However, because in this trial the metal ball is much heavier than the styrofoam ball, despite its smaller size children with the incorrect Material or Mass beliefs may incorrectly respond 1, 2, or 3 in their predictions or postdictions, and potentially be surprised by the evidence (that 5 "wins").

3.1 Bayesian Model of the Pupillary Surprise Response

The Ideal Bayesian learning model that we employ for our investigations builds on a recent investigation of individual differences in children's belief revision (the Optimal Bayesian model described in [11]). Here, the Bayesian model constructed computational representations of children's beliefs based on their task responses. Doing so highlighted the importance of individual differences in prior beliefs during learning, while further demonstrating the impact of multiple, competing beliefs that guide inferences, as the Bayesian model's correlations to children's behavioral responses were significantly stronger than competing frameworks for the entire subject pool (Bayesian Correlations > .8; Directional accuracy > 90%). Additionally, this model found that children in the experiment's Prediction condition were better simulated by the model than children in the Postdiction condition.

[2] Measures in this study and others reveal that children are honest about their responses in these postdiction conditions.

We will build upon the Bayesian model's simulations for estimating children's pupil dilation measurements during the original experiment. Specifically, we will look at the Bayesian trial-by-trial surprise predictions for individual children. Children's estimated beliefs are based on their responses during the pretest and follow Bayesian posterior updating during the test trial observations ($p(h_t | d_t)$ ∀ $h_t \epsilon H_t$). Surprisal (whether Shannon or Kullback-Leibler divergence) for each trial depends on an individual child's expected belief state given the evidence for that trial.

Children's beliefs about how much water will be displaced by different objects have been identified by past literature (e.g., [58] Burbules & Linn, 1978), falling into relatively simple causal rules for predictions: a rule based on the size of the objects, one based on the material of the objects, one on based on mass of the objects (a mixture of size and material), and one reflecting random responding. Thus, in our model children's beliefs were represented computationally as a distribution across these four possible beliefs ("Size" (*S*), "Material" (*M*), "Mass" (*W*), and finally a "Random" (*R*)). Thus, each child's "model" ($p(H_t | d_t)$; Eq. (3)) of water displacement on a given trial (*t*) could be represented as the posterior probability over just four rules (S, M, W, R):

$$p(H_t|d_t) \ = [p(h_{st} = S|d_t), p(h_{mt} = M|d_t), p(h_{wt} = W|d_t), p(h_{rt} = R|d_t)]. \tag{3}$$

3.1.1 Calculating Shannon Information

From Eq. (1), we derive the model's trial-by-trial *SI* surprise estimates in Eq. (4). That is, on some trial (*t*), we determine the likelihood ($p(d_t | H_t)$) of that trial's new data (d_t) observed by the child given their currently inferred model (H_t):

$$SI \ = \ -log(p(d_t)). \tag{4}$$

Here, Eq. (5) describes how our model calculates the probability of the data ($p(d_t)$) on a given trial (*t*), as marginalizing over the four competing beliefs at time *t*, ($h_i = h_{(w,m,w,r)}$), which is the summation over the likelihood and prior for each model:

$$p(d_t) \ = \textstyle\sum_{h_{,i} \, \epsilon \, H_t} p(d_t|h_{t,i})p(h_{t,i}). \tag{5}$$

This calculation entails treatment of each individual hypothesis's (h_i) current state at each trial (h_t). The likelihood is weighed by the strength of belief for each model under this summation. Thus, evidence that is less likely under more strongly held beliefs will contribute more to surprise than when evidence is unlikely under a weakly held belief. (See Figure 2 for illustration.)

3.1.2 Calculating Kullback-Leibler Divergence

From Eq. (2), we derive trial-by-trial Kullback-Leibler divergence as a surprise estimate in Eq. (6). For some trial (*t*), we calculate the relative entropy for each considered belief (hypothesis $h_{t,i}$) of the child's currently held distribution of prior beliefs ($p(h_{t,i} | d_t)$ ∀ $h_{t,i} \epsilon H_t$) with its respective posterior belief, $p(h_{t+1,\, i} | d_{t+1})$. Kullback-Leibler divergence (*KLD*) is taken as the sum of these relative entropies between prior and posterior beliefs capturing the shift in distributions between time (*t*) and after observing the data at time (*t+1*):

$$KLD(H_{t+1}||\, H_t) \ = \sum_{h_{t,i} \epsilon \, H_t} p(h_{t+1,i} \,|d_{t+1}) \, log[\frac{p(h_{t+1,i} \,|d_{t+1})}{p(h_{t,i} \,|d_t)}]. \tag{6}$$

Here, on a trial (t), the data have not yet been observed and capture the distribution of beliefs prior to observing the evidence, where-as trial t+1 captures the posterior distribution. Kullback-Leibler is simply capturing the relative change between prior and posterior given some observation. (See Figure 3 for illustration.)

Assessing the performance of these candidate metrics of quantifiable pupillary surprise – both in general and in competition with one another – provides a means to explore

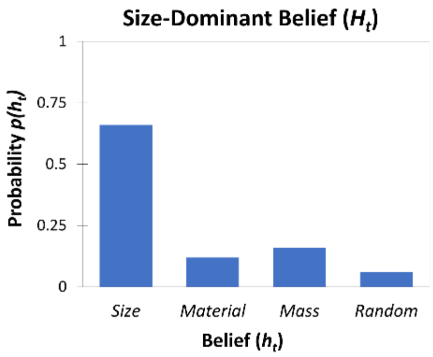
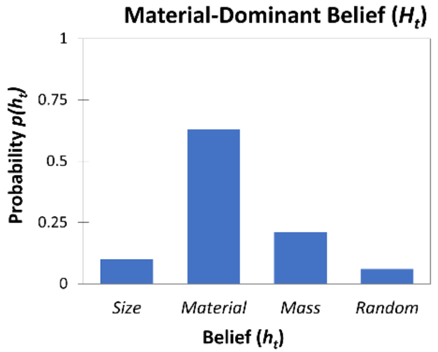
Examples of Hypothetical Belief Distributions (H_t)	Prior Beliefs ($p(h_t)$)	Data Likelihood ($p(d_t \mid h_t)$)	Prior-Weighted Likelihoods ($p(h_t)p(d_t \mid h_t)$)	Shannon Information
(A) 	$p(h_{t,\,Size}) \sim 0.66$ $p(h_{t,\,Material}) \sim 0.12$ $p(h_{t,\,Mass}) \sim 0.16$ $p(h_{t,\,Random}) \sim 0.06$	For the example trial observation of d_t = “5 – Right Wins”, $p(\text{“5”} \mid h_t = \text{“Size”}) \sim 0.65$ $p(\text{“5”} \mid h_t = \text{“Material”}) \sim (9.9)^{-5}$ $p(\text{“5”} \mid h_t = \text{“Mass”}) \sim 0.10$ $p(\text{“5”} \mid h_t = \text{“Random”}) \sim 0.20$	$p(h_{t,\,Size})\ p(\text{“5”} \mid h_{t,\,Size}) \sim 0.43$ $p(h_{t,\,Material})\ p(\text{“5”} \mid h_{t,\,Material}) \sim (1.2)^{-5}$ $p(h_{t,\,Mass})\ p(\text{“5”} \mid h_{t,\,Mass}) \sim 0.02$ $p(h_{t,\,Random})\ p(\text{“5”} \mid h_{t,\,Random}) \sim 0.01$ Summed Total $p(d_t = \text{“5”}) \sim 0.46$	$-\log(p(d_t = \text{“5”})) \sim 0.77$ Higher Probability Data, Lower Surprise
(B) 	$p(h_{t,\,Size}) \sim 0.10$ $p(h_{t,\,Material}) \sim 0.63$ $p(h_{t,\,Mass}) \sim 0.21$ $p(h_{t,\,Random}) \sim 0.06$		$p(h_{t,\,Size})\ p(\text{“5”} \mid h_{t,\,Size}) \sim 0.06$ $p(h_{t,\,Material})\ p(\text{“5”} \mid h_{t,\,Material}) \sim (6.3)^{-5}$ $p(h_{t,\,Mass})\ p(\text{“5”} \mid h_{t,\,Mass}) \sim 0.02$ $p(h_{t,\,Random})\ p(\text{“5”} \mid h_{t,\,Random}) \sim 0.01$ Summed Total $p(d_t = \text{“5”}) \sim 0.10$	$-\log(p(d_t = \text{“5”})) \sim 2.29$ Lower Probability Data, Higher Surprise

Figure 2. Example of the procedure for calculating Shannon Information given the current model’s simulations, formally described by Eq. (4) and Eq. (5). Row A and B display two examples of the different profiles of children’s prior beliefs captured in graph (Column 1) and numeric (Column 2) form. Given some incoming data (e.g., the example trial from Figure 1; a Small Metal ball vs a Large Styrofoam ball), the likelihood of the observation (that event “5 - Right Wins” occurs) is estimated for all four models (Column 3). Then, a posterior probability is calculated by weighing the individual child’s prior beliefs against the likelihood (Column 4). Shannon Information is calculated by summing over (marginalizing out $h_{t,i}$) these posteriors and taking the negative log likelihood of the final summed total. Thus, there is an inferred negative relationship between data likelihood ($p(d_t)$) and model surprise according to the Shannon Information account (Column 5). That is, when the weighted likelihood of data is low, model surprise is high; similarly, when the likelihood of data is high, model surprise is low.

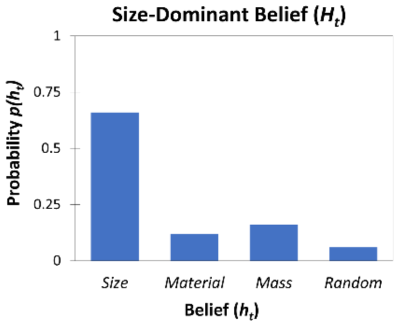
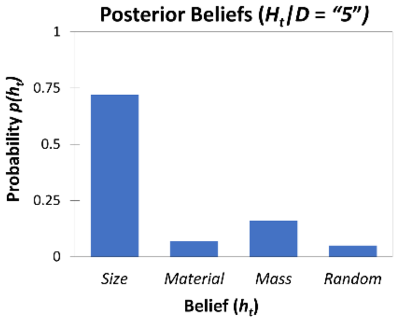
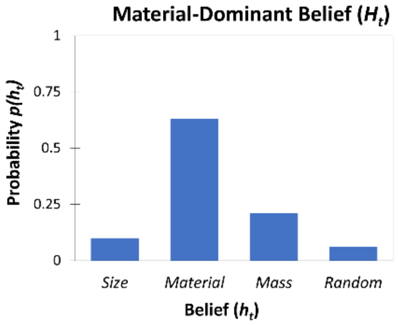
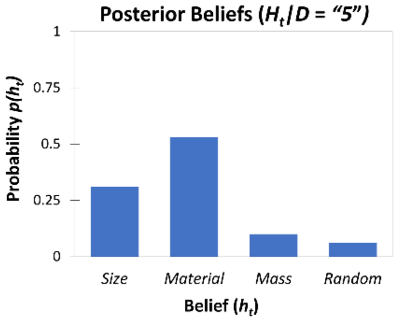
	Examples of Hypothetical Belief Distributions (H_t)	Prior Beliefs ($p(h_t)$)	Normalized Posterior Beliefs Distributions ($H_t \mid d_t$)	Normalized Posterior Beliefs ($p(h_t \mid d_t)$)	Kullback-Leibler Divergence
(A)		$p(h_{t,\ Size}) \sim 0.66$ $p(h_{t,\ Material}) \sim 0.12$ $p(h_{t,\ Mass}) \sim 0.16$ $p(h_{t,\ Random}) \sim 0.06$		$p(h_{t,\ Size} \mid d_t) \sim 0.93$ $p(h_{t,\ Material} \mid d_t) \sim (2.5)^{-5}$ $p(h_{t,\ Mass} \mid d_t) \sim (3.7)^{-5}$ $p(h_{t,\ Random} \mid d_t) \sim (2.6)^{-2}$	$\sum_{h_t \epsilon H_t} p(h_t \mid d_t) \log\left[\frac{p(h_t \mid d_t)}{p(h_t)}\right] \sim 0.018$ Similar Priors & Posteriors, Lower Surprise
(B)		$p(h_{t,\ Size}) \sim 0.10$ $p(h_{t,\ Material}) \sim 0.63$ $p(h_{t,\ Mass}) \sim 0.21$ $p(h_{t,\ Random}) \sim 0.06$		$p(h_{t,\ Size} \mid d_t) \sim 0.65$ $p(h_{t,\ Material} \mid d_t) \sim (6.2)^{-4}$ $p(h_{t,\ Mass} \mid d_t) \sim 0.23$ $p(h_{t,\ Random} \mid d_t) \sim 0.19$	$\sum_{h_t \epsilon H_t} p(h_t \mid d_t) \log\left[\frac{p(h_t \mid d_t)}{p(h_t)}\right] \sim 0.151$ Dissimilar Priors & Posteriors, Higher Surprise

Figure 3. Example of the procedure for calculating Kullback-Leibler Divergence (KLD) given the current model's simulations, as formally described by Eq. (6). Row A and B display two examples of the different profiles of children's prior beliefs captured in graph (Column 1) and numeric (Column 2) form. Given some incoming data (e.g., observing option "5" = right side wins for the example trial in Figure 1; a Small Metal ball vs a Large Styrofoam ball) and the prior beliefs of the learner (Belief Distributions, H_t), we consider the posterior belief distribution that best accommodates the observed data (e.g., $p(H_t \mid$"5"$)$), again, captured in graph (Column 5) and numeric (Column 4) form. Then, the Kullback-Leibler Divergence, is calculated as the sum of relative entropies between the prior probability and posterior probability between each of the individual competing beliefs ($h_{t,i}$). Thus, there is an inferred positive relationship between the degree of dissimilarity between distributions (divergence between the prior and posterior) and model surprise according to the Kullback-Leibler Divergence account. That is, when the prior and posterior are dissimilar, model surprise is high; conversely, when the prior and posterior are similar, model surprise is low.

doi:10.20944/preprints202211.0142.v1

the implications of different learning responses to data at the individual level, trial-by-trial. If the Shannon Information (SI) estimates better correlate with children's pupil dilation, then this may suggest that pupil dilation is an indicator of robust engagement with incoming data – particularly when it is of low likelihood and highly "informative". If the Kullback-Leibler divergence (KLD) correlates more strongly with pupil dilation, then this may suggest that pupil dilation is an indicator of belief updating "in-the-moment". Assessing these correlations under different contexts (prediction vs postdiction) allows exploration of potentially different mechanisms engaged by different types of learning interventions.

4. Results

4.1. Assessing Fit of Model-Estimates

The analyses performed for assessing each of the surprise estimates, Shannon Information and Kullback-Leibler divergence, use direct correlations between model predictions of and children's pupil dilation responses recording during the experiment. Bonferroni correction is performed where needed for conservative analyses and interpretation, with correlation p-values tested against a Bonferroni-corrected alpha (Condition [Prediction, Postdiction] × Estimate [*SI, KLD*], $\alpha = 0.05/4 = 0.0125$. All correlations discussed in the Results section are additionally compiled in a table found in Appendix A for ease of comparison.

4.1.1 Condition-combined analyses

When looking at the full dataset (2890 trials across 94 children), we found no significant correlation for either the *Shannon Information* ($r(2889) = 0.01$, $p = 0.49$) or the *Kullback-Leibler divergence model* ($r(2889) = 0.02$, $p = 0.12$) to children's pupillometry measurements. As noted, our primary question involves assessing the models accounting for two different response modalities (prediction *and* postdiction) to assess the potential differences between these interventions.

4.1.2 Condition-separate analyses

We first explored condition differences of children's pupillometry response as related to Shannon Information. The *Shannon Information* estimate did not correlate with the pupillary response for either the Prediction ($r(1437) = 0.03$, $p = 0.20$) or Postdiction ($r(1461) = 0.01$, $p = 0.62$) condition. In contrast, exploring condition differences of children's pupillometry response as related to Kullback-Leibler divergence did reveal differences. The *Kullback-Leibler divergence* estimate was significantly correlated with children's pupillary response within the prediction condition ($r(1437) = 0.07$, $p = 0.004 < \alpha$). There was no correlation between pupillary response and Kullback-Leibler divergence for children in the Postdiction condition ($r(1461) = -0.003$, $p = 0.90$). The difference between the strength of the *Kullback-Leibler divergence* and *Shannon Information* correlations within the Prediction condition was also significant, ($z = 2.98$, $p = 0.0014$)[3]. Correlations between *Kullback-Leibler divergence* and pupillary response were also significantly different between the Prediction and Postdiction conditions (Fisher's r-to-z transformation; $z = 2.11$, $p = 0.0174$).

4.1.3 Exploratory analysis with data subsets

Sources of noise, such as individual differences in prior beliefs and an identified critical learning period (both highlighted in previous modeling work; Colantonio et al., *in review*) may have affected the correlation between the model estimates and pupillary surprise. Therefore, we looked to control for two additional sources of noise in our data via follow-up analyses. First, not all of the children in the study were still "learners", as a

[3] Similar results are found for a bounded version of the Kullback-Leibler divergence measure, the Jensen-Shannon divergence [59,60]. These results can be found in Appendix A.

subset of the participants began the Learning Phase with the correct Size-belief. Applying the same method as above, we looked at just the children who did not have beliefs based on the correct theory of water displacement at the beginning of the experiment (19 children had the correct theory already, leaving n = 75 of 94 children who began with an incorrect theory, approximately equally between conditions). Re-analyzing the data with this subset replicated the results above. There was no significant correlation between *Shannon Information* and children's pupillometry for the "learners" subset (overall $r(2259) = 0.02$, $p = 0.29$; Prediction, $r(1142) = 0.04$, $p = 0.11$; Postdiction, $r(1116) = 0.01$, $p = 0.54$). Meanwhile, while the *Kullback-Leibler divergence* had no significant correlation with the entire "learner" subset ($r(2259) = 0.036$, $p = 0.08$), there were significantly stronger correlations between *Kullback-Leibler divergence* and the pupil dilation response for learners within the Prediction condition ($r(1142) = 0.08$, $p = 0.002 < \alpha$) compared to the Postdiction condition ($r(1116) = -0.002$, $p = 0.93$; comparing conditions: Fisher's r-to-z transformation; $z = 2.15$, $p = 0.0158$). The *Kullback-Leibler divergence* did not have a significantly stronger correlation than *Shannon Information* for "learners" in the Prediction condition ($z = 0.99$, $p = 0.16$) for this subset[4].

Our second subset analysis explored only trials where "learning" was likely to take place. Previous modeling of children's learning over the course of the study revealed that most children converged onto the correct Size belief by trial 19 based on their choice behavior (where the 19th trial was the 75th percentile of when children in the study seemed to have "learned" the Size belief according to the model; discussed in more detail in Colantonio et al., *in review*). The sharp-then-plateaued learning rate was likely because the initial trials (n = 9) provided in the Learning Phase provided no differentiation between the competing belief models (Size, Material, Mass). They were selected to be "congruent" with all theories and thus offered no "surprise" for any model or opportunity for learning. Following a handful of incongruent evidence (trials 10-19) the majority of children revised their beliefs and began responding consistently with the correct Size belief. This design (no conflicting evidence to support learning initially, nor learning after the correct beliefs are settled) may have artificially created "noise" in our pupillometry correlations. This is because variability of responses in pupillometry measures caused by other artifacts could temper correlations due to a relatively large number of trials where *Shannon Information* and *Kullback-Leibler divergence* estimates were both very low. Thus, we also looked at "critical learning trials" – those that started with the first incongruent trial (trial 10, where data would be differentiated by the competing beliefs) and extended to trial 19 where almost all children (n = 74 of 94 children) had learned the correct belief (size dictates water displacement) as measured by Bayes Posterior Odds. For these "critical learning trials", we again replicated the overall pattern of results. *Shannon Information* did not correlate overall during these critical trials ($r(858) = 0.04$, $p = 0.24$), nor did it correlate within either condition (Prediction condition: $r(431) = 0.05$, $p = 0.26$; Postdiction condition: $r(426) = 0.06$, $p = 0.18$). Again, the *Kullback-Leibler divergence* did not correlate for all children across all of the "critical" trials, ($r(858) = 0.04$, $p = 0.24$). However, (replicating the other analyses) there was a significant correlation between the *Kullback-Leibler divergence* estimate and pupillary response within the Prediction condition ($r(431) = 0.12$, $p = 0.013 < \alpha$); while no correlation was found in the Postdiction condition ($r(426) = -0.003$, $p = 0.90$). These correlations are significantly different between Prediction and Postdiction conditions for *Kullback-Leibler divergence* (Fisher's r-to-z transformation; $z = 2.11$, $p = 0.0174$). . The difference between

[4] As would be expected by small sample size and the fact that children's with the correct theory would have predicted low surprise for trials across the full study, none of these correlations are significant when looking at the subset of "already-knowers" (overall for *SI*, $r(639) = -0.05$, $p = 0.20$; for *KLD*, $r(639) = -0.02$, $p = 0.58$), even when looking between the Prediction (for *SI*, $r(294) = -0.07$, $p = 0.19$; for *KLD*, $r(294) = 0.01$, $p = 0.79$) and Postdiction condition (for *SI*, $r(344) = -0.03$, $p = 0.52$; for *KLD*, $r(344) < 0.01.$, $p = 0.99$).

Kullback-Leibler divergence and *Shannon Information* for the Prediction condition yielded a significant difference, as well (z = 2.98, p = 0.0014) during these "critical" trials. This suggests that the pupillary surprise response reflects something like belief-updating, but only in conditions when children are actively engaged in prediction (a point we return to in the Discussion).

4.2. Modeling Individual Differences

We were also interested in relating pupillary response and modeled surprise to learning. Thus, we looked at how, at the individual level, the degree of fit between physiological response and model response related to the degree to which children's responses reflected Bayesian "optimal" learning. That is, we are correlating two correlations. Specifically, for this investigation, we looked at the correlation of children's answer behavior (1-5) to Bayesian model predictions of those answers as one set of correlations, and children's pupillary response performance and our models of surprise as the second set of correlations. If pupillary response relates to learning, we might expect to see that those children whose pupillary responses are more aligned with model predictions are also the same children who learn more "optimally". Indeed, we found that the correlation of individual children's pupil response to *Kullback-Leibler divergence* correlated significantly to the correlation of those children's answers and ideal Bayesian learning (r(88) = 0.27, p = 0.007). In contrast, correlations based on children's pupil response and *Shannon Information* did not correlate to this learning measure (r(88) = 0.05, p = 0.58). The difference between the correlation coefficients was marginally significant (z = 1.49, p = 0.06).

5. Discussion

This paper describes one of the first computational investigations of the links between children's pupillary surprise response and their science concept learning, as related to the contextual effects of engaging in an explicit prediction or postdiction. We modeled data, including pupillometry responses, collected from elementary school children who provided predictions or predictions in a water-displacement learning task. By modeling individual children's beliefs and learning over trials, we could capture two different forms of "surprise": Shannon Information and Kullback-Leibler divergence. Overall, we find that the children's pupillary surprise response is related to Kullback-Leibler divergence – but only in cases where children have generated an explicit prediction prior to observing the potentially surprising events. Furthermore, we found that children whose pupillometry data was best estimated by the Kullback-Leibler divergence also tended to be the children whose behavioral response data (from an experiment on learning water displacement via belief revision) was best fit by an ideal Bayesian learning model.

Our findings fit well with the theory described at the intersection of cognitive, emotional, and physiological research (e.g., [33,34]), with particular links to recent work investigating the role of prediction in belief revision (e.g., [25,27,40,57,61,62]). Our findings also converge with other related research. Like Kayhan and colleagues [45], we found a relationship between pupil dilation and the Kullback-Leibler divergence. Both this previous work and the current investigation find that the calculated divergence may affect belief revision in regard to the amount of updating needed to adjust current beliefs. However, there are two key differences between our modeling work and that of Kayhan et al. [45] which are important to note. First, the current paper investigates children's pupillary surprise under different contextual conditions. The current results find that the relationship between modeled surprise (via Kullback-Leibler divergence) and children's pupillary surprise response may *only* occur when children are actively making predictions – but not when they are passively observing and evaluating. This highlights that there are instances where pupillary surprise might be more likely to occur when making predictions – as proposed by other recent empirical work (e.g., [28]). Second, in line with the original paper that we draw our model from [11], the current model accounts for individual differences among children's prior beliefs and the processes by which they update. In of Kayhan et al. [45], children's behaviors are modeled to all follow the same inferred computational model[5]. In contrast, we formalize the prior beliefs that children may have at the individual level, as informed by their past behavior.

Like other work investigating surprise during learning, we found a relationship between the Kullback-Leibler divergence and the pupillary surprise response (e.g., [44,45]). However, unlike O'Reilly and colleagues [44], we did not find a relationship between likelihood-based Shannon Information and pupil dilation. One potential reason for this divergence is that there are differences in the degree of complexity of the learned "concept" of each study and in the number of hypotheses considered. Specifically, the previous work entailed a task that only required reasoning about one variable (the angle that the target appeared at on a screen; [44]), however the angle of the target may have taken many different values. In contrast, the currently modeled task may require reasoning about more complex, causal beliefs (e.g., whether an object's size, material, or weight determines the amount of water displaced and how each of these features generates displacement; [40]), but only considered a few possible hypotheses[6]. Thus, one particular reason for the significant relationship in past work between pupil dilation and Shannon Information (or likelihood-based prediction error), and the poorer fit with children's pupillometry in the Prediction condition of the current dataset may relate to either differences in the complexity of the concept being inferred or differences in the size of hypothesis space being considered.

A second difference between our results and O'Reilly and colleagues' [44] was that we found a *positive* correlation between pupil dilation and the Kullback-Leibler divergence during prediction, whereas a *negative* correlation was found in this past work. Our task differed in both the types of beliefs being considered, and whether children were actively engaged in prediction. If beliefs are already engaged in this process (as they likely were for our participants following the explicit prediction prompt), then a relatively instantaneous pupillary growth response to the observed outcome is feasible. In our task, the number of options being considered and "simulated" by children is bounded[7], with children only deciding among five options (really three directional outcomes). One possibility is that the positive dilation we observed in the prediction condition captured the amount of mental effort generated by explicitly considering outcomes over more complex hypotheses. It has been suggested that when the necessary *'work'* appears unexpectedly "large", more mental effort may be exerted to accommodate the new information (e.g., to reduce the '*work*'; Friston et al., 2006; Friston, 2010), and be reflected by increases in children's pupil dilation – similar to findings linking reduction of uncertainty to the presence of signals from neuromodulators (e.g., acetylcholine and norepinephrine; [65,66]). Of

[5] Understandably, we acknowledge limitations of Kayhan et al.'s [45] investigation given the population being studied. Specifically, Kayhan and colleagues faced the challenge of investigating this domain in18-month-old infants and 24-month-old toddlers. Thus, acquiring explicit measures to inform computational representations of prior beliefs may have been difficult or not plausible.

[6] It is of course likely that children were entertaining a more varied set of potential causal beliefs about displacement than the four considered here. Responses in the pretest aligned well across these four and past work has focused on these, but we are open to there being a more complex space of beliefs in this domain as well. Indeed, as learners consider more complex interactions (like buoyancy, water permeable materials like sponges, etc.) the space will balloon.

[7] Additional analyses investigating a bounded divergence measure, Jensen-Shannon divergence [59,60], is also performed and described in Appendix A. Importantly, the Jensen-Shannon divergence performs almost identically to the Kullback-Leibler divergence in terms of its correlations with children's pupillometry.

course, we do not have enough evidence that confirms that pupil dilation actually accompanies a more "effortful" mental process (e.g., like those found by [29,67]), only that the found correlations indicate a relationship between pupil dilation and the amount of *'work'* needed to update beliefs.

5.1. Understanding potential cognitive mechanisms

Both Shannon Information and Kullback-Leibler divergence accounts of pupillary surprise have support in the literature exploring cognitive mechanisms. Specifically, these proposed computational interpretations align with the mentioned attentional network described in past work and are not necessarily exclusive. Shannon Information has been suggested to relate more to the "intermediate-level" factors, addressing what it is externally a learner might be trying to process when pupil dilation occurs (e.g., [37-39]). Similarly, the Kullback-Leibler divergence has been suggested to represent "higher-level" factors relating to internal processes and state-like fluctuations that the learner might be experiencing (e.g., [25,33,40]). Thus, support for either the Shannon Information or the Kullback-Leibler divergence (or potentially both) estimating children's pupillometry would have fit with various findings and interpretations of pupil dilation as some form of attentional network activation (see a thorough review in [34]).

If these accounts of Shannon capturing "intermediate-level" factors and Kullback-Leibler divergence capturing "higher-level" features are correct, our results provide support for "higher-level" factors being engaged in our task – at least when children are explicitly making predictions. Perhaps when making predictions, children are orienting their attention toward their beliefs. That is, pupil dilation in our task may be an indicator of children's online assessment of their current models of the world *and* what the implications would be (how much effort is needed to change these models) given the potential outcomes of an upcoming event.

Why might Kullback-Leibler divergence capture greater attention or cognitive effort? As described earlier, Shannon Information quantifies a single signal of data informativeness against only the current hypothesis space [42,46-49]. In contrast Kullback-Leibler requires a computation over two hypothesis spaces – the prior and the posterior. In this way, Kullback-Leibler might reflect more effortful cognitive processes.

5.2. Limitations & Future Work

The implications of this work highlight key investigations that future work should pursue. Specifically, one such avenue entails empirically and computationally capturing a "construct" of surprise that accounts for its emotional, cognitive, and physiological components. Next, future work may also be interested in further refining our understanding of the "higher-level" processes that our results suggest being associated with surprise – that is, interactions among prediction, planning, and other executive functioning.

5.2.1 The Noisiness of Pupillometry Measurements

We acknowledge the impact of noise within the original experiment's pupillometry data, which could be due to many possible reasons. First, both the children and the model seemed to "quickly" learn the scientific concept (that size determines the amount of displaced water). Thus, opportunities for experiencing pupillary surprise may have been in short supply as misconceptions of water displacement were not held onto for long. In response to this, we also analyzed subsets of the data to account for potential noise due to learning dynamics: whether children had already "known" the size principle at Pretest, and the "critical" trials where learning would be most likely to happen. Doing so did lead to improvements in the fit between the Kullback-Leibler divergence when estimating surprise, and did not affect the lack of fit with Shannon Information.

The second reason that noise may have been prevalent was that despite best efforts for careful task administration and data collection, there do exist drawbacks when collecting pupillometry measures. For example, careful preparation of the study's location is needed[8], as low-level issues like light levels and focal distance do affect fluctuations in pupil size [34-36]. This is important to acknowledge, as many interpretations of pupillometry entail an assessment of the average change in pupil size within a timeframe. Additionally, work investigating the influence of low-level factors like light levels finds that pupil dilation can be oscillatory with respect to fluctuations in the luminance of objects and their environments [68]. This may lead to a pupillary surprise response with a short latency (relative to the measured timeframe), but particularly strong amplitude being washed out by constriction of the pupil (whether by nervous system relaxation or slight light level variance) during the timeframe when measures are averaged.

Finally, following the acknowledgment of the potential sources of noise, we also acknowledge the relative strength of the found correlations (e.g., in order of the *Results* section, the significant correlations held Pearson's correlation coefficients of $r = 0.07$, 0.08. 0.12). However, these correlations were found to be significant even when performing analyses conservatively (via Bonferroni correction). To the best of our knowledge, this work seems to be the first to find significant correlations between pupillometry and a computational model estimate during science concept learning.

5.2.2 Capturing Pupillary Surprise across Modalities

Notably, we found no correlations between either Shannon Information or Kullback-Leibler divergence and the pupillometry measures of children in the Postdiction condition. As described in the previous section, this may be partially due to noise leading to underpowered detection. However, it may also suggest that perhaps another mechanism (and thus another model surprise metric) needs to be considered and investigated in future work regarding when (or even, if) pupillary surprise occurs in different response modalities. The current work highlights that when making predictions, pupil dilation may be indicating the performance of a higher-level, learning-effort estimates. However, we did not find significant correlations between pupillary response and model predictions in the postdiction condition despite the fact that over the course of the experiment, these children also learned. Indeed, pupil dilation did occur at times during the original study for children in the Postdiction condition – just not in a way that correlated with models of surprise. Thus, future work should investigate whether other response modalities indicate that processes are being performed when pupil dilation is elicited with theory-based metrics for estimating said pupillometry computationally.

5.2.3 Empirically Measuring Surprise

In contemporary work on surprise, the physiological measure of pupil dilation is commonly collected as a proxy or marker that signals an individual's experience of surprise (e.g., [22]). This tends to be proposed due to the occurrence of pupil dilation following a violation of one's expectations – often inducing heightened attention, physiological arousal (e.g., the release of noradrenaline and norepinephrine), and increased activity in brain areas (e.g., within the brainstem) related to monitoring uncertainty [23,30,69]. But, as with most emotions, special care needs to be taken when discussing measures and expressions of affective states. In particular, surprise has received considerable attention since the mid-20th century that still informs theoretical concerns regarding what surprise actually is and connecting the (less-so recently) disparate fields that investigate surprise (see [19,20,70]). Importantly, these conceptualizations and implementations of surprise only relate to physiological instances of surprise's attentional capacities. Thus, future research that looks to finely define surprise not only in terms of its proposed physiological markers but also subjective experiential phenomena, could also collect self-reported

[8] In collecting the modeled data (Theobald & Brod, 2021), great efforts were made to prepare the study location at a local science museum. For example, the experimenters used a room with no windows, allowing only for artificial light to keep the light levels as consistent as possible.

measures of experienced surprise as an additional correlate to further substantiate claims surrounding physiological measures of surprise.

5.2.4 Investigating Modalities that Potentially Leverage Prior Beliefs

Future work might consider investigating science concept learning by revisiting interview methods of past studies to further understand children's subjective prior beliefs and what processes children (propose that they) may have employed to revise them (e.g., as in earlier water displacement studies; [58]). In fact, recent work highlights that thought experiments – imagining outcomes of an event and revising assumptions – can be beneficial for learning in both adults [71] and young children (six-year-olds; [72]). Thus, future work may tackle the integration of key experimental design aspects from the currently modeled data (the role of prediction and pupillometry) and research on other learning-by-thinking methods like thought experiments. Doing so may help determine whether such planning is being implemented by children. However, such approaches should be done carefully and interpreted cautiously, as such meta-cognitive awareness and performance of thought experiments may be difficult to do, and work explicitly on whether people (especially children) typically benefit from thought experiments (compared to origins in allusions to scientific revolutionaries like Galileo, Kepler, and Einstein) is relatively new to the field [73].

5.2.5 Potential Roles of Executive Function

Strides in research on attention highlight that top-down regulation and executive control are vital for processing and awareness of relevant information in the environment (extensively reviewed in [33,34]). Specifically, executive function is important for the guidance of intermediate-level attentional processes (e.g., alerting and orienting) for sensory operations. Here, we propose that future work should perform further computational investigations centered on incorporating measures of executive function. Modeling any relationships among theory change, prediction, pupillary, and executive function skills (such as inhibition and cognitive flexibility; [74,75]) may provide further insight into other relevant mechanisms that support science concept learning. Such modeling would highlight whether executive function affects model performance straightforwardly, where higher executive function measures might correlate with better model performance. Additionally, future work may entail the design of Bayesian models that account for various executive function skills. For example, would a model that has the ability to inhibit incorrect prior beliefs perform better? Or perhaps, would a model that flexibly switches focus towards updated, "more correct" theories be plausible and sufficiently capture children's behavior?

5.3. Conclusions

Here, we have identified a candidate computational measure that may capture the pupillary surprise response in a quantifiable way when children are making predictions during science learning. Specifically, we found that when children make predictions, their pupil dilation in response to observed outcomes may be a temporal indicator of the child leveraging their initial prior beliefs and extrapolating the implications of those outcomes given said prior beliefs. The current work contributes to our knowledge of *what* pupil dilation may be an expression of during the learning process. Specifically, by identifying contexts where pupillometry can be estimated computationally via the Kullback-Leibler divergence, we have also identified candidate mechanisms and processes that children may be performing when pupil dilation is elicited. That is, since the Kullback-Leibler divergence typically describes dissimilarity, or the amount of "work" needed to transform one probability distribution into another, the current findings have highlighted that explicit prediction may elicit the pupil dilation response as a physiological marker of children's belief revision – estimating how much "work" is needed to move from prior to posterior. This behavior was not associated for children who were only *post hoc* evaluat-

doi:10.20944/preprints202211.0142.v1

ing, suggesting a privileged role for prediction in engaging learning-relevant physiological responses. This computational modeling investigation, alongside the recent experiments centered on prediction, provides some initial insight into why engaging children to generate predictions may support learning more effectively than other interventions. Such a simple manipulation may differently engage affective states and impact children's learning; that is perhaps most surprising of all.

Author Contributions: Conceptualization, J.C., I.B., M.T., E.B. and G.B.; formal analysis, J.C. and E.B; methodology, J.C. and E.B.; software, J.C..; investigation, M.T. and G.B.; writing—original draft preparation, J.C. and E.B.; writing—review and editing, I.B., M.T. and G.B.; visualization, J.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The experimental data and code used in this study are available online at https://osf.io/wqnf2/?view_only=fba98d3304bc4a759e4346eaa7df3332

Acknowledgments: We would like to thank the members of the Computational Cognitive Development labs for their support and discussion of the theory and results presented in this paper.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Table A. Correlations between children's pupillometry measurements and each of three computational estimates of pupillary surprise: Shannon Information, Kullback-Leibler Divergence, and Jensen-Shannon Divergence. Values in boldface formatting highlight significant correlations following Bonferroni correction among the three measures ($\alpha = 0.05 / 3 = 0.1\overline{66}7$).

	Shannon Information		**Kullback-Leibler Divergence**		**Jensen-Shannon Divergence**	
Condition	**Prediction**	**Postdiction**	**Prediction**	**Postdiction**	**Prediction**	**Postdiction**
All Trials	*Combined* *r*(2889) = 0.01, *p* = 0.49		*Combined* *r*(2889) = 0.028, *p* = 0.12		*Combined* *r*(2886) = 0.029, *p* = 0.11	
	r(1437) = 0.03 *p* = 0.20	*r*(1461) = 0.01 *p* = 0.62	***r*(1437) = 0.07** ***p* = 0.004**	*r*(1461) = -0.003 *p* = 0.90	***r*(1435) = 0.07** ***p* = 0.005**	*r*(1459) = 0.01 *p* = 0.59
Learners Only	*Combined* r(2259) = 0.02, *p* = 0.29		*Combined* r(2259) = 0.036, *p* = 0.082		*Combined* r(2259) = 0.04, *p* = 0.056	
	r(1142) = 0.04 *p* = 0.11	*r*(1116) = 0.01 *p* = 0.54	***r*(1142) = 0.08** ***p* = 0.002**	*r*(1116) = -0.002 *p* = 0.93	***r*(1142) = 0.09** ***p* = 0.002**	*r*(1116) = 0.01 *p* = 0.70
"Already Knowers"	*Combined* r(639) = -0.05, *p* = 0.20		*Combined* r(639) = -0.02, *p* = 0.58		*Combined* r(636) = -0.01, *p* = 0.76	
	r(294) = -0.07 *p* = 0.19	*r*(344) = -0.03 *p* = 0.52	*r*(239) = 0.01 *p* = 0.79	*r*(344) = 0.01 *p* = 0.99	*r*(292) = -0.01 *p* = 0.82	*r*(342) = 0.10 *p* = 0.054
Critical Trials	*Combined* r(858) = 0.04, *p* = 0.17		*Combined* r(858) = 0.039, *p* = 0.24		*Combined* r(858) = 0.04, *p* = 0.18	
	r(431) = 0.05 *p* = 0.26	*r*(426) = 0.06 *p* = 0.18	***r*(431) = 0.11** ***p* = 0.013**	*r*(426) = -0.01 *p* = 0.72	***r*(431) = 0.10** ***p* = 0.03**	*r*(426) = 0.02 *p* = 0.60

References

1. Gärdenfors, P. Conditionals and Changes of Belief. *Acta Philosophica Fennica* **1978**, *30*, 381–404.
2. Siegler, R.S. Three Aspects of Cognitive Development. *Cognitive psychology* **1976**, *8*, 481–520.
3. Carey, S. *Conceptual Change in Childhood*; MIT press Cambridge, MA, 1985; Vol. 460.
4. Shapiro, E.Y. *Inductive Inference of Theories from Facts*; Yale University, Department of Computer Science, 1981.
5. Gopnik, A.; Meltzoff, A.N. *Words, Thoughts, and Theories*; MIT Press, 1997.
6. Wellman, H.M.; Gelman, S.A. Knowledge Acquisition in Foundational Domains. **1998**.
7. Schulz, L.E.; Bonawitz, E.B.; Griffiths, T.L. Can Being Scared Cause Tummy Aches? Naive Theories, Ambiguous Evidence, and Preschoolers' Causal Inferences. *Developmental psychology* **2007**, *43*, 1124.
8. Gopnik, A.; Bonawitz, E. Bayesian Models of Child Development. *Wiley interdisciplinary reviews: cognitive science* **2015**, *6*, 75–86.
9. Griffiths, T.L.; Tenenbaum, J.B. Theory-Based Causal Induction. *Psychological review* **2009**, *116*, 661.
10. Kemp, C.; Tenenbaum, J.B.; Niyogi, S.; Griffiths, T.L. A Probabilistic Model of Theory Formation. *Cognition* **2010**, *114*, 165–196.

11. Colantonio, J., Bascandziev, I., Theobald, M., Brod, G., & Bonawitz, E. Priors, Progressions, and Predictions in Science Learning: Theory-Based Bayesian Models of Children's Revising Beliefs of Water Displacement. *IEEE TCDS* 2022, *revised and resubmitted*.
12. Mansinghka, V.; Kemp, C.; Griffiths, T.; Tenenbaum, J. Structured Priors for Structure Learning. *arXiv preprint arXiv:1206.6852* **2012**.
13. Tenenbaum, J.B.; Griffiths, T.L.; Kemp, C. Theory-Based Bayesian Models of Inductive Learning and Reasoning. *Trends in cognitive sciences* **2006**, *10*, 309–318.
14. Tenenbaum, J.B.; Kemp, C.; Griffiths, T.L.; Goodman, N.D. How to Grow a Mind: Statistics, Structure, and Abstraction. *science* **2011**, *331*, 1279–1285.
15. Ullman, T.D.; Goodman, N.D.; Tenenbaum, J.B. Theory Learning as Stochastic Search in the Language of Thought. *Cognitive Development* **2012**, *27*, 455–480.
16. Bonawitz, E.; Ullman, T.D.; Bridgers, S.; Gopnik, A.; Tenenbaum, J.B. Sticking to the Evidence? A Behavioral and Computational Case Study of Micro-Theory Change in the Domain of Magnetism. *Cognitive Science* **2019**, *43*, e12765, doi:10.1111/cogs.12765.
17. Ullman, T.D.; Tenenbaum, J.B. Bayesian Models of Conceptual Development: Learning as Building Models of the World. **2020**.
18. Goodman, N.D.; Ullman, T.D.; Tenenbaum, J.B. Learning a Theory of Causality. *Psychological review* **2011**, *118*, 110.
19. Reisenzein, R.; Horstmann, G.; Schützwohl, A. The Cognitive-Evolutionary Model of Surprise: A Review of the Evidence. *Topics in Cognitive Science* **2019**, *11*, 50–74.
20. Stahl, A.E.; Feigenson, L. Violations of Core Knowledge Shape Early Learning. *Topics in cognitive science* **2019**, *11*, 136–153.
21. Ekman, P. Are There Basic Emotions? *Psychological Review* **1992**, *99*, 550–553, doi:10.1037/0033-295X.99.3.550.
22. Reisenzein, R.; Bördgen, S.; Holtbernd, T.; Matz, D. Evidence for Strong Dissociation between Emotion and Facial Displays: The Case of Surprise. *Journal of personality and social psychology* **2006**, *91*, 295.
23. Preuschoff, K.; 't Hart, B.M.; Einhäuser, W. Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making. *Frontiers in neuroscience* **2011**, *5*, 115.
24. Kloosterman, N.A.; Meindertsma, T.; van Loon, A.M.; Lamme, V.A.; Bonneh, Y.S.; Donner, T.H. Pupil Size Tracks Perceptual Content and Surprise. *European Journal of Neuroscience* **2015**, *41*, 1068–1078.
25. Brod, G.; Hasselhorn, M.; Bunge, S.A. When Generating a Prediction Boosts Learning: The Element of Surprise. *Learning and Instruction* **2018**, *55*, 22–31, doi:10.1016/j.learninstruc.2018.01.013.
26. Krüger, M.; Bartels, W.; Krist, H. Illuminating the Dark Ages: Pupil Dilation as a Measure of Expectancy Violation across the Life Span. *Child Development* **2020**, *91*, 2221–2236.
27. Breitwieser, J.; Brod, G. Cognitive Prerequisites for Generative Learning: Why Some Learning Strategies Are More Effective Than Others. *Child Development* **2021**, *92*, 258–272, doi:10.1111/cdev.13393.
28. Brod, G.; Greve, A.; Jolles, D.; Theobald, M.; Galeano-Keiner, E.M. Explicitly Predicting Outcomes Enhances Learning of Expectancy-Violating Information. *Psychon Bull Rev* **2022**, doi:10.3758/s13423-022-02124-x.
29. Kahneman, D.; Beatty, J. Pupil Diameter and Load on Memory. *Science* **1966**, *154*, 1583–1585.
30. Aston-Jones, G.; Cohen, J.D. An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience* **2005**, *28*, 403–450, doi:10.1146/annurev.neuro.28.061604.135709.
31. Laeng, B.; Sirois, S.; Gredebäck, G. Pupillometry: A Window to the Preconscious? *Perspectives on psychological science* **2012**, *7*, 18–27.
32. Sirois, S.; Brisson, J. Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science* **2014**, *5*, 679–692.
33. Petersen, S.E.; Posner, M.I. The Attention System of the Human Brain: 20 Years After. *Annual review of neuroscience* **2012**, *35*, 73.
34. Strauch, C.; Wang, C.-A.; Einhäuser, W.; Van der Stigchel, S.; Naber, M. Pupillometry as an Integrated Readout of Distinct Attentional Networks. *Trends in Neurosciences* **2022**, *45*, 635–647, doi:10.1016/j.tins.2022.05.003.
35. Mathôt, S. Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition* **2018**, *1*.
36. Strauch, C.; Koniakowsky, I.; Huckauf, A. Decision Making and Oddball Effects on Pupil Size: Evidence for a Sequential Process. *Journal of cognition* **2020**, *3*.
37. Sokolov, E.N. Higher Nervous Functions: The Orienting Reflex. *Annual review of physiology* **1963**, *25*, 545–580.
38. Corneil, B.D.; Munoz, D.P. Overt Responses during Covert Orienting. *Neuron* **2014**, *82*, 1230–1243, doi:10.1016/j.neuron.2014.05.040.
39. Wetzel, N.; Buttelmann, D.; Schieler, A.; Widmann, A. Infant and Adult Pupil Dilation in Response to Unexpected Sounds. *Developmental psychobiology* **2016**, *58*, 382–392.
40. Theobald, M.; Brod, G. Tackling Scientific Misconceptions: The Element of Surprise. *Child Development* **2021**, *92*, 2128–2141.
41. Dörrenberg, S.; Rakoczy, H.; Liszkowski, U. How (Not) to Measure Infant Theory of Mind: Testing the Replicability and Validity of Four Non-Verbal Measures. *Cognitive Development* **2018**, *46*, 12–30, doi:10.1016/j.cogdev.2018.01.001.
42. Shannon, C.E. A Mathematical Theory of Communication. *The Bell system technical journal* **1948**, *27*, 379–423.
43. Kullback, S.; Leibler, R.A. On Information and Sufficiency. *The annals of mathematical statistics* **1951**, *22*, 79–86.
44. O'Reilly, J.X.; Schüffelgen, U.; Cuell, S.F.; Behrens, T.E.; Mars, R.B.; Rushworth, M.F. Dissociable Effects of Surprise and Model Update in Parietal and Anterior Cingulate Cortex. *Proceedings of the National Academy of Sciences* **2013**, *110*, E3660–E3669.
45. Kayhan, E.; Heil, L.; Kwisthout, J.; Rooij, I. van; Hunnius, S.; Bekkering, H. Young Children Integrate Current Observations, Priors and Agent Information to Predict Others' Actions. *PLOS ONE* **2019**, *14*, e0200976, doi:10.1371/journal.pone.0200976.

46. Good, I.J. The Surprise Index for the Multivariate Normal Distribution. *The Annals of Mathematical Statistics* **1956**, *27*, 1130–1135.
47. Greenland, S. Valid P-Values Behave Exactly as They Should: Some Misleading Criticisms of p-Values and Their Resolution with s-Values. *The American Statistician* **2019**, *73*, 106–114.
48. Cole, S.R.; Edwards, J.K.; Greenland, S. Surprise! *American Journal of Epidemiology* **2021**, *190*, 191–193, doi:10.1093/aje/kwaa136.
49. Rafi, Z.; Greenland, S. Semantic and Cognitive Tools to Aid Statistical Science: Replace Confidence and Significance by Compatibility and Surprise. *BMC medical research methodology* **2020**, *20*, 1–13.
50. Itti, L.; Baldi, P. Bayesian Surprise Attracts Human Attention. *Vision research* **2009**, *49*, 1295–1306..
51. Baldi, P.; Itti, L. Of Bits and Wows: A Bayesian Theory of Surprise with Applications to Attention. *Neural Networks* **2010**, *23*, 649–666, doi:10.1016/j.neunet.2009.12.007.
52. Kullback, S. Probability Densities with given Marginals. *The Annals of Mathematical Statistics* **1968**, *39*, 1236–1243.
53. Champagne, A.B.; Klopfer, L.E.; Gunstone, R.F. Cognitive Research and the Design of Science Instruction. *Educational Psychologist* **1982**, *17*, 31–53, doi:10.1080/00461528209529242.
54. Crouch, C.; Fagen, A.P.; Callan, J.P.; Mazur, E. Classroom Demonstrations: Learning Tools or Entertainment? *American Journal of Physics* **2004**, *72*, 835–838, doi:10.1119/1.1707018.
55. Inagaki, K.; Hatano, G. Amplification of Cognitive Motivation and Its Effects on Epistemic Observation. *American Educational Research Journal* **1977**, *14*, 485–491.
56. Brod, G. Predicting as a Learning Strategy. *Psychonomic Bulletin & Review* **2021**, 1–9.
57. Brod, G.; Breitwieser, J.; Hasselhorn, M.; Bunge, S.A. Being Proven Wrong Elicits Learning in Children – but Only in Those with Higher Executive Function Skills. *Developmental Science* **2020**, *23*, e12916, doi:10.1111/desc.12916.
58. Burbules, N.C.; Linn, M.C. Response to Contradiction: Scientific Reasoning during Adolescence. *Journal of Educational Psychology* **1988**, *80*, 67–75, doi:10.1037/0022-0663.80.1.67.
59. Lin, J. Divergence Measures Based on the Shannon Entropy. *IEEE Transactions on Information theory* **1991**, *37*, 145–151.
60. Wong, A.K.; You, M. Entropy and Distance of Random Graphs with Application to Structural Pattern Recognition. *IEEE transactions on pattern analysis and machine intelligence* **1985**, 599–609.
61. Noordewier, M.K.; Topolinski, S.; Van Dijk, E. The Temporal Dynamics of Surprise. *Social and Personality Psychology Compass* **2016**, *10*, 136–149.
62. Brod, G.; Breitwieser, J. Lighting the Wick in the Candle of Learning: Generating a Prediction Stimulates Curiosity. *npj Sci. Learn.* **2019**, *4*, 1–7, doi:10.1038/s41539-019-0056-y.
63. Friston, K.; Kilner, J.; Harrison, L. A Free Energy Principle for the Brain. *Journal of Physiology-Paris* **2006**, *100*, 70–87, doi:10.1016/j.jphysparis.2006.10.001.
64. Friston, K. The Free-Energy Principle: A Unified Brain Theory? *Nat Rev Neurosci* **2010**, *11*, 127–138, doi:10.1038/nrn2787.
65. Yu, A.J.; Dayan, P. Uncertainty, Neuromodulation, and Attention. *Neuron* **2005**, *46*, 681–692.
66. Parr, T.; Friston, K.J. Uncertainty, Epistemics and Active Inference. *Journal of the Royal Society Interface* **2017**, *14*, 20170376.
67. Just, M.A.; Carpenter, P.A. The Intensity Dimension of Thought: Pupillometric Indices of Sentence Processing. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* **1993**, *47*, 310.
68. Naber, M.; Alvarez, G.A.; Nakayama, K. Tracking the Allocation of Attention Using Human Pupillary Oscillations. *Frontiers in psychology* **2013**, *4*, 919.
69. Lavín, C.; San Martín, R.; Rosales Jubal, E. Pupil Dilation Signals Uncertainty and Surprise in a Learning Gambling Task. *Frontiers in behavioral neuroscience* **2014**, *7*, 218.
70. Munnich, E.L.; Foster, M.I.; Keane, M.T. Editors' Introduction and Review: An Appraisal of Surprise: Tracing the Threads That Stitch It Together. *Topics in Cognitive Science* **2019**, *11*, 37–49.
71. Bascandziev, I. Inconsistencies Among Beliefs as a Basis for Learning via Thought Experiments. In Proceedings of the CogSci; 2020.
72. Bascandziev, I.; Carey, S. Young Children Learn Equally from Real and Thought Experiments. *Proceedings of the Annual Meeting of the Cognitive Science Society* **2022**, *44*.
73. Bascandziev, I.; Harris, P.L. Can Children Benefit from Thought Experiments. *The scientific imagination* **2019**, 262.
74. Miyake, A.; Friedman, N.P. The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current directions in psychological science* **2012**, *21*, 8–14.
75. Diamond, A. Activities and Programs That Improve Children's Executive Functions. *Current directions in psychological science* **2012**, *21*, 335–341.