How to Make a Proceedings Paper Submission

Anonymous CogSci submission

Abstract

Include no author information in the initial submission, to facilitate blind review. The abstract should be one paragraph, indented 1/8 inch on both sides, in 9~point font with single spacing. The heading 'Abstract' should be 10~point, bold, centered, with one line of space below it. This one-paragraph abstract section is required only for standard six page proceedings papers. Following the abstract should be a blank line, followed by the header 'Keywords' and a list of descriptive keywords separated by semicolons, all in 9~point font, as shown below.

Keywords: Add your choice of indexing terms or keywords; kindly use a semi-colon; between each term.

Introduction

Experiment

Methods

Participants We recruited 449' participants (Age range: M = ;) on Prolific. They were randomly assigned to one of the three conditions of the experiment (Curiosity: N =; Memory: N = ; Math: N =;). Participants were excluded if they showed irregular reaction times (N = ???) or their responses in the filler tasks indicates low engagement with the experiment (Curiosity: N =; Memory: N = ; Math: N =;). All exclusion criteria were pre-registered. The final sample included N participants (Curiosity N = ; Memory: N =; Math: N =).

Procedure This is a web-based self-paced visual presentation task. Participants were instructed to look at a sequence of animated creatures at their own pace and answer some questions throughout. At the end of the experiment, participants were asked to rate the similarity between pairs of creatures and complexity of creatures they encountered on a 7-point Likert Scale. Each participant saw eight blocks in total, half of which used creatures with high perceptual complexity, and half of which used creatures with low perceptual complexity. On each trial, an animated creature showed up on the screen. participants can press the down arrow to go to the next trial whenever they want after a minimum viewing time of 500 ms.

Each block consisted of six trials. A trial can be either a background trial (B) or a deviant trial (D). A background trial presented a creature repeatedly, and the deviant trial presented a different creature from the background trial in the block. Two creatures in the blocks were matched for visual complexity. There were four sequences of background trials and deviant trials. Each sequence appeared twice, once with high complexity stimuli and once with low complexity stimuli. The deviant trial can appear at either the second (BDBBBB), the fourth (BBBDBB), or the sixth trial (BBBBBD) in the block. Two blocks do not have deviant trials (BBBBBB). The creatures presented in the deviant trials and background trials were matched for complexity.

Participants were randomly assigned to one of the three conditions: Curiosity, Memory, and Math The three conditions only differed in the type of questions asked following each block. In Curiosity condition, participants were asked to rate "How curious are you about the creature?" on a 5-point Likert scale. In Memory condition, a forced-choice recognition question followed each block ("Have you seen this creature before?"). The creature used in the question in both conditions was either a creature presented in the preceding block or a novel creature matched in complexity. In Math condition, the participants were asked a simple arithmetic question ("What is 5 + 7?") in multiple-choice format.

Stimuli The animated creatures (Fig 1) were created using Spore (a game developed by Maxis in 2008). There were forty creatures in total, half of which have low perceptual complexity (e.g. the creatures do not have limbs, additional body parts, facial features, or textured skin), and half of which have high perceptual complexity (i.e. they do have the aforementioned features). We used the "animated avatar" function in Spore to capture the creatures in motion.

Results

Analytic Approach The sample size, methods, and main analyses were all pre-registered and are available at [LINK]. Data and analysis scripts are available at [LINK].

Manipulation Check The complex animated creatures were rated as more perceptually complex (M = ; SD =) than the simple animated creatures (M = ; SD =). Pairs of background creature and deviant creature were rated as moderately dissimilar to one another (M = ; SD =).

Evaluating the Paradigm Three criteria were selected to evaluate whether the paradigms successfully captured the characteristic looking time patterns observed in infant literature: habituation (the decrease in looking time for a stimulus with repeated presentations), dishabituation (the increase in looking time to a new stimulus after habituated to one stimulus), and complexity effect (longer looking time for perceptually more complex stimuli). To evaluate the phenomenon quantitatively, we ran a linear mixed effects model with maximal random effect structure. [DESCRIBE THE MODEL]. [REPORT THE MODEL RESULTS]

Order Effect While visualizing the data, we unexpectedly found that the position in which the deviant trial appeared in the sequence had an effect on the shape of the habituation and dishabituation curves. To explore this phenomenon quantitatively, we operationalized the magnitude of dishabituation as the difference between the looking time at the deviant trial minus the background trial at the same position. Then, we fit a mixed effect model with the position of deviant as fixed effect and [???] as a random effect. We found that the position was a significant predictor of the magnitude of dishabituation (looking time at the deviant trial minus the background trial at the same position). Deviant trials that appeared last elicited the strongest dishabituation effect (M = ; SD:,), followed by the deviant trials appeared fourth (M, SD), with the deviant trial on the second showing the smallest amount of dishabituation (M, SD).

Discussion

Model

We formalized the learning problem that participants face in our experiments as a form of Bayesian concept learning (Tenenbaum, 1999; Goodman, 2006), represented graphically in Fig. X. The goal is to learn a concept *theta*, which is a set of probabilities for independent binary features $\theta_{1,2,..,n}$, where n is the number of features. Over the course of a block, the learner receives information about θ by observing exemplars y: instantiations of $\bar{\theta}$, where each feature $y_{1,2,..,n}$ is either on or off. Each feature θ_i and its corresponding exemplar y_i form a Beta-Bernoulli process:

$$p(\theta_i) \sim Beta(\alpha_i, \beta_i) \tag{1}$$

$$p(y_i|\theta_i) \sim Bernoulli(\theta_i) \tag{2}$$

Since the features are independent, this relationship holds for the entire concept θ. However, to model the time course of attention, we do not want to assume that information is encoded perfectly and instantaneously. Instead, we suggest that participants gather repeated noisy samples $\bar{z}$ from the exemplars. For any sample z from an exemplar y there is a small probability ε to misperceive the feature as off when it was actually on, and vice versa. Therefore, by making noisy observations $\bar{z}$, the learner obtains information about the true identity of the exemplar y, and by extension, about the concept $\bar{theta}$. By Bayes' rule:

$$P(\theta|\bar{z}) \quad = p(\bar{z}|y)p(y|\theta)p(\theta)/p(\bar{z}) \tag{3}$$

where $p(\bar{z}|y_i)$ is fully described by ε, and $p(y|\theta)$ by Bernoulli processes as in Eq. 2.

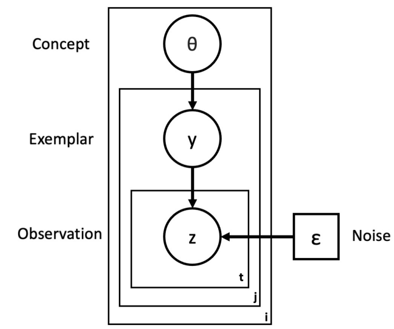

Figure 1: Graphical representation of our model. Circles indicate random variables. The squares indicate fixed model parameters.

Like in our experiment, the learner's task is to decide when to stop sampling. If they do so rationally, then they should anchor their sampling behavior to the expected information gain (EIG) of the next sample. We compute EIG by weighing the information gain from each possible next observation by the probability of that observation. We defined information gain as the KL-divergence between the hypothetical posterior after observing a sample z_{t+1} and the current posterior:

$$EIG(z_{t+1}) = \sum_{z_{t+1} \in [0,1]} p(z_{t+1}|\theta_t) * KL(\theta_{t+1}, p(\theta_t)) \tag{4}$$

Finally, to get actual sampling behavior from the model, it has to convert EIG into a binary decision about whether continue looking at the current sample, or to advance to the next trial. The model does so using a luce choice between the EIG from the next sample and a constant EIG from looking away.

$$p(look) = \frac{EIG(z_{t+1})}{EIG(z_{t+1}) + EIG(world)} \tag{5}$$

We also studied the behavior of the model when replacing EIG with other linking hypotheses, such as surprisal (the probability of a given z under the $P(\theta_t)$) and KL-divergence between the posterior $p(\theta_t)$ and the prior $p(\theta_{t-1})$.

General Discussion

References

References