Supporting Information

Gweon et al. 10.1073/pnas.1003095107

Bayesian Model

As discussed in the main text, the joint dependence between S (sampling process) and T (property extension) can be described as a simple Bayesian network (Fig. 1). The learner's goal is to predict Y, which depends directly on T, not S or D. However, inferences about T from D must take into account the different possible values of S; formally, our Bayesian analysis must integrate out S in scoring each value of T. Because the data are inconsistent with hypothesis t_2, only two hypotheses for T are relevant; t_3 predicts that yellow balls squeak whereas t_1 predicts that they do not. Following Tenenbaum and Griffiths (1), the evidence for one of these hypotheses over the other can be measured by the likelihood ratio

$$L = \frac{P(D|t_3)}{P(D|t_1)} = \frac{P(n|t_3, \beta)}{P(n|t_1, \beta)}.$$

We posit that children's exploratory behavior—how much they squeeze the yellow ball, expecting a squeak—will be monotonically related to L (Fig. 3*A*). This analysis makes predictions that are independent of the prior probabilities children assign to t_1 or t_3, removing a degree of freedom that would otherwise need to be measured or fit empirically to their behavior. These likelihoods can be computed by integrating out the sampling process:

$$P(n|t, \beta) = \sum_{s_i \in S} P(n|t, s, \beta)P(s).$$

To evaluate these likelihoods we need the following four conditional probabilities*:

$$\begin{aligned} P(n|t_1, s_1, \beta) &= 1 \\ P(n|t_1, s_2, \beta) &= \beta^n \\ P(n|t_3, s_1, \beta) &= \beta^n \\ P(n|t_3, s_2, \beta) &= \beta^n. \end{aligned}$$

Let α denote the prior probability $P(s_1)$ that the experimenter is sampling from just the squeaky balls: $P(s_2) = 1 - \alpha$. We then have

$$\begin{aligned} P(n|t_1, \beta) &= \sum_{s_i \in S} P(n|t_1, s, \beta)P(s) \\ &= P(n|t_1, s_1, \beta)P(s_1) + P(n|t_1, s_2, \beta)P(s_2) \\ &= \alpha + \beta^n(1-\alpha). \\ P(n|t_3, \beta) &= \sum_{s_i \in S} P(n|t_3, s, \beta)P(s) \\ &= P(n|t_3, s_1, \beta)P(s_1) + P(n|t_3, s_2, \beta)P(s_2) \\ &= \beta^n\alpha + \beta^n(1-\alpha) \\ &= \beta^n. \end{aligned}$$

The likelihood ratio, measuring the evidence in favor of the proposition that yellow balls squeak, is then

$$\begin{aligned} L &= \frac{P(n|t_3, \beta)}{P(n|t_1, \beta)} \\ &= \frac{\beta^n}{\alpha + \beta^n(1-\alpha)}. \end{aligned}$$

By setting the parameter α to 0, we can model the possibility that infants expect that evidence is sampled randomly; by setting the parameter α to 1, we can model the possibility that infants expect that evidence is sampled selectively (Fig. 3 *B* and *C*).

1. Tenenbaum JB, Griffiths TL (2001) Generalization, similarity, and Bayesian inference. *Behav Brain Sci* 24:629–640.

*In the behavioral experiment, the balls were sampled without replacement. However, they were sampled from a hidden compartment of a box with a false front, such that the apparent ratio of blue to yellow balls did not change. Because the estimated differences between the model with and without replacement are minor, we use a fixed β to avoid making arbitrary assumptions about the contents of the box.