Meta-analysis

The simplest way to run a meta-analysis of the data is to aggregate the results from all experiments and analyze them as a whole. Altogether, children's success rate was 0.7875 with 95% CI: 0.7-0.875. This shows that the overall pattern of data supports our theory. However, this analysis does not capture the possibility that some tasks may be easier than others, and it does not consider the constraints of how the successes and failures are distributed across the four experiments. For example, if every participant succeeded in the first three experiments, and every participant failed on the last experiment, the conclusions from this meta-analysis would be almost identical, but intuitively, the fact that the failures were all concentrated in one experiment matters – in contrast to our actual pattern of data, in which similar proportions of children (around three-quarters) succeed in all experiments. In general, a meta-analysis of the overall distribution of children's responses should also be sensitive to each experiment's local distribution.

Supplemental Figure 1 shows a sketch of the model used for the meta-analysis of our study. Children's responses in each experiment *i* are modeled as a binomial distribution with unobservable bias θ_i. For any given experiment, the value of the bias is assumed to be drawn from a beta distribution with parameters $\alpha=\mu\kappa$ and $\beta=(1-\mu)\kappa$. This parameterization, based on Kruschke (2015), allows for a more intuitive interpretation of the parameters. Here, the beta distribution is determined by its mean value μ, a real number between 0 and 1 which reflects the shared overall response rate across all experiments assumed to be due to the underlying naïve utility calculations, and a parameter κ, varying between 1 and infinity, that reflects the amount of between-experiment variability – how much children's response rate is expected to vary from one experiment to the next. (Greater values of κ indicate lower between-experiment variability.)

Given our complete dataset $D = (D_1, D_2, D_3, D_4, D_5)$[1] we can compute the posterior distribution of μ through Bayes' rule:

$$p(\mu|D;\kappa) \propto L(D|\mu;\kappa)p(\mu) =$$

$$\int_\theta p(D_i|\theta_i)p(\theta_i|\mu;\kappa)p(\mu),$$

where $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ and $p(\mu)$ is a uniform (uninformative) distribution. We approximated the integral above numerically, with μ varying discretely in steps of 0.01. The inferences we make about μ depend on what inferences or assumptions we make about the variability parameter κ. Supplemental Figure 2 shows the expected value of μ for different values of the parameter κ, ranging in integer steps

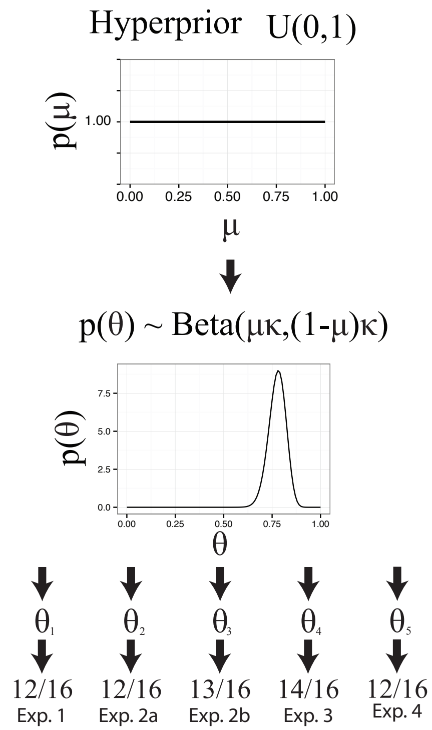
[1] While we presented our work as four experiments, in Experiment 2 the cookie-cracker condition and the clover-daisy condition varied slightly in the methods and, as such, we opted to allow the model to set different biases for each condition. Hence we had effectively five experiments in the meta-analysis. However, treating them as a single experiment has no impact on the conclusions.

between 1 and 150. Supplemental Figure 3a shows some examples of the posterior distributions on μ for different values of κ. Supplemental Figure 3b shows the beta distributions obtained by estimating the expected value of μ for different values of κ.

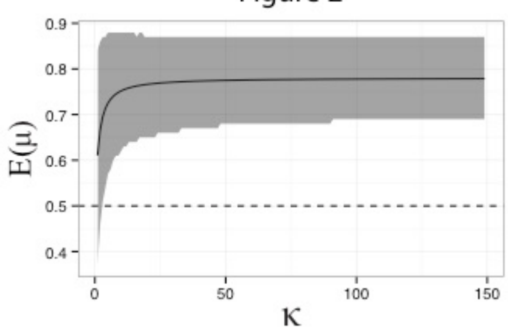
We can summarize these results in several ways. As κ increases, the expected posterior value of μ converges to 0.78, with the 95% highest-density interval (HDI) 0.69-0.87 (i.e., the highest posterior probability of μ is in this interval, containing 95% of the total probability mass). The likelihood of κ is strictly increasing over the range 1-150, so this estimate of μ = 0.78, with 95% HDI between 0.69 and 0.87 also reflects the posterior on μ at the maximum likelihood value of κ. Note that this approach returns an estimate of μ and 95% HDIs almost identical to the simplest analysis we began with, aggregating the results of all experiments into a single large sample. In this case, that should not be surprising. The fact that the most likely value of κ is so high reflects the inference that the underlying response rates in each of our experiments were essentially indistinguishable, intuitively because they were likely measuring the same underlying response process in similar ways. We can also place a prior on κ and compute a joint posterior over κ and μ. Placing a uniform prior on κ for values between 1 and 150 produces a posterior distribution with expected values $\mu = 0.78$ and $\kappa = 88$. Under this distribution, the probability that the mean μ is higher than chance level ($p(\mu$>0.5)) is greater than 0.99999. The corresponding beta distribution resulting from these parameters is shown in the model sketch in Supplemental Figure 1. Placing other reasonable priors on κ , such as an uninformative prior ($1/\kappa$) or a gamma prior, yields extremely similar results. The evidence that μ is well above 0.5, and most likely within 10% of 0.78 (as determined by the 95% HDI) is so strong that it is supported under the model for every individual value of κ tested except the two most extreme values ($\kappa = 1,2$) on the bottom end (see Supplemental Figure 2). However, these two values had the lowest likelihoods of all values tested (p($D|\kappa = 1$)<0.0001, and p($D|\kappa = 2$)<0.0005), and reflect implausibly high levels of between-experiment variability.

In sum, our inferences about the most probable values of μ are the same whether we look at the most likely values of κ, place a prior on κ and integrate it out in a joint Bayesian analysis of μ and κ, or look at special cases of the model for any plausible specific value of κ. Taken together, the individual analyses of each of our experiments reported in the main text, the simple meta-analysis aggregating all data into one experiment, and this meta-analysis of the underlying response rate across all our experiments all support the conclusion that children's behavior follows the predictions of the naïve utility calculus.

Supplemental
Figure 1

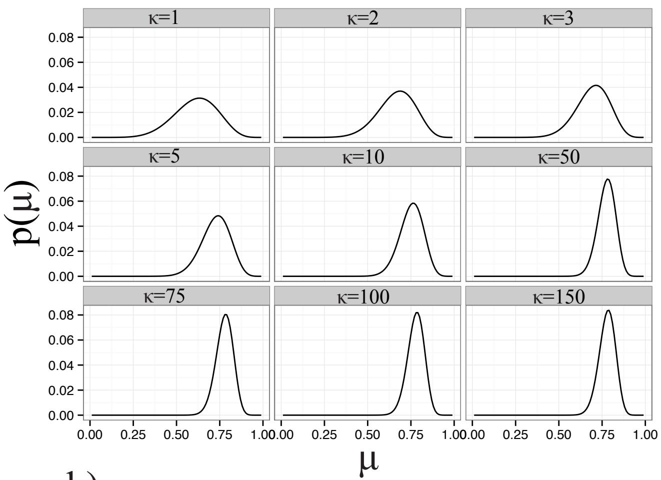


Supplemental
Figure 2



Supplemental
Figure 3

a)



b)