19 Number concept development

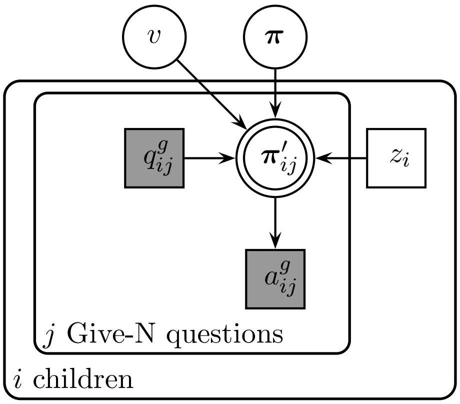
A basic challenge in understanding human cognitive development is to understand how children acquire number concepts. Since the time of Piaget (1952), the concept of number has been one of the most active areas of research in the field. This chapter focuses on one prominent current theory about the origin of integer concepts, called the "knower-level" theory (Carey, 2001; Carey & Sarnecka, 2006; Wynn, 1990, 1992).

The knower-level theory asserts that children learn the exact cardinal meanings of the first three or four number words one-by-one and in order. That is, children begin by learning the meaning of "one" first, then "two," then "three," and then (for some children) "four," at which point they make an inductive leap, and infer the meanings of the rest of the words in their counting list. In the terminology of the theory, children start as PN-knowers (for "Pre-Number"), progress to one-knowers once they understand "one," through the two-knower, three-knower, and (for some children) four-knower levels, until they eventually become CP-knowers (for "Cardinal Principle").

There are at least two common behavioral tasks that are used to assess children's number knowledge. In the "Give-N" task, children are asked to give some number of objects, such as small toys, to the experimenter, or an experimenter substitute, such as a puppet (e.g., Frye, Braisby, Lowe, Maroudas, & Nicholls, 1989; Fuson, 1988; Schaeffer, Eggleston, & Scott, 1974; Wynn, 1990, 1992). In the "Fast-Cards" task, children are asked how many objects, such as pictures of animals, were displayed on a briefly presented card (e.g., Le Corre & Carey, 2007). In both tasks, the behavioral data are just a set of question–answer pairs, recording how many objects were present or asked for, and how many the child gave or answered.

The responses that children give in the Give-N and Fast-Cards tasks naturally depend on what numbers they understand, but also depend on properties of the tasks themselves. For example, if there are 15 toys available in a Give-N task, it is common for a child to give all 15 when asked for a number they do not understand. The same child, however, asked for the same number in a Fast-Cards task, would be very unlikely to answer 15, because there is nothing special about that answer in the different task setting. Thus, a good model needs to combine child-specific and task-specific components to capture the behavioral data.

We consider a knower-level model, developed by Lee and Sarnecka (2010), that incorporates task-specific components. We use a subset of 20 children from the data considered by Lee and Sarnecka (2010), first applying it to Give-N data, then to



$$\boldsymbol{\pi} \sim \text{Dirichlet}(\overbrace{1,\ldots,1}^{15})$$
$$v \sim \text{Uniform}(1, 1000)$$
$$z_i \sim \text{Categorical}(\tfrac{1}{6},\ldots,\tfrac{1}{6})$$
$$\pi'_{ijk} \propto \begin{cases} \pi_k & \text{if } k > z_i \\ v \times \pi_{ijk} & \text{if } k \le z_i \text{ and } k = q^g_{ij} \\ \frac{1}{v} \times \pi_{ijk} & \text{if } k \le z_i \text{ and } k \ne q^g_{ij} \end{cases}$$
$$a^g_{ij} \sim \text{Categorical}(\boldsymbol{\pi}'_{ij})$$

Fig. 19.1 Graphical model for behavior on the Give-N task according to the Lee and Sarnecka (2010) knower-level model.

Fast-Cards data from the same children, and then to both sets of data simultaneously.

19.1 Knower-level model for Give-N

The basic idea behind the Lee and Sarnecka (2010) model is that a behavioral task has a base-rate for responding. This is simply a predisposition towards giving each possible answer, before any question has been asked. For the Give-N task, for example, we might expect one toy, two toys, a small handful of toys, or all of the available toys, to be possible answers with relatively high probabilities in the base-rate. The child's actual response comes from updating that base-rate when they are asked a question. The base-rate is task-specific, and the updating depends on the question and the child's number knowledge.

The nature of the updating in the model can be explained in terms of two cases. In the first case, the child is asked to give a number they understand. This means the probability of giving that number increases, and the probability of giving other numbers they understand but were not asked to give all decrease. The relative probabilities of giving other numbers in the base-rate are not changed, because the child does not know about them. So, for example, if a three-knower is asked to give two, their probability of giving 2 increases, and their probability of giving 1 or 3 decreases, relative to the base-rate. But the numbers 4 and above do not change in their relative probabilities.

In the second case, the child is asked to give a number they do not know. This means the probabilities of giving numbers they do know decrease, but all of the other numbers will retain the same relative probabilities. So, for example, if a three-knower is asked to give five, they become much less likely to give 1, 2, or 3, but equally relatively likely to give 4 and above.

Figure 19.1 presents a graphical model for this knower-level model. The data are the observed q^g_{ij} questions and a^g_{ij} answers for the ith child on their jth question. The task-specific base-rate probabilities are represented by the vector $\boldsymbol{\pi}$, so that π_k is the probability for giving k as an answer. The base-rate is updated to $\boldsymbol{\pi}'$. The updating occurs using the number asked for, the knower level z_i of the child, and an evidence value v that measures the strength of the updating. The updating logic explained by the two cases earlier can be formalized as

$$\pi'_{ijk} \propto \begin{cases} \pi_k & \text{if } k > z_i \\ v \times \pi_{ijk} & \text{if } k \leq z_i \text{ and } k = q^g_{ij} \\ \frac{1}{v} \times \pi_{ijk} & \text{if } k \leq z_i \text{ and } k \neq q^g_{ij}. \end{cases}$$

The actual answer produced by the child is then assumed to be sampled according to the probabilities for each possibility in the updated base-rate $\boldsymbol{\pi}'$. The categorical distribution, which can be thought of as the extension of the Bernoulli distribution to cases where there are more than two possible choices, expresses this naturally as

$$a^g_{ij} \sim \text{Categorical}\big(\boldsymbol{\pi}'\big).$$

The knower-level parameter z_i is given a categorical prior so that the six possibilities—PN-knower, one-knower, two-knower, three-knower, four-knower, and CP-knower—are given equal prior probability. The base-rate prior is given by a Dirichlet distribution that makes all possible distributions over the 15 toys equally likely. The Dirichlet distribution can be thought of as an extension of a beta distribution to cases where there are more than two possible alternatives, so $\text{Dirichlet}\big(1,\ldots,1\big)$ is naturally conceived as the extension of the uniform distribution $\text{Beta}\big(1,1\big)$.

The script `NumberConcept_1.txt` implements the graphical model in WinBUGS. Notice that the script collects posterior predictions for each individual child, as well as for each knower level as a group of children:

```
# Knower Level Model Applied to Give-N Data
model{
  # Data
  for (i in 1:ns){
    for (j in 1:gnq[i]){
      # Probability a z[i]-Knower Will Answer ga[i,j] to Question gq[i,j]
      # is a Categorical Draw From Their Distribution over the 1:gn Toys
      ga[i,j] ~ dcat(npiprime[z[i],gq[i,j],1:gn])
    }
    # Posterior Predictive
    for (j in 1:gn){
      predga[i,j] ~ dcat(npiprime[z[i],j,1:gn])
    }
  }
  # Model
  for (i in 1:nz){
    for (j in 1:gn){
```

```
      for (k in 1:gn){
        piprimetmp[i,j,k,1] <- pi[k]
        piprimetmp[i,j,k,2] <- 1/v*pi[k]
        piprimetmp[i,j,k,3] <- v*pi[k]
        # Will be 1 if Knower-Level (i.e, i-1) is Same or Greater than Answer
        ind1[i,j,k] <- step((i-1)-k)
        # Will be 1 for the Possible Answer that Matches the Question
        ind2[i,j,k] <- equals(k,j)
        # Will be 1 for 0-Knowers
        ind3[i,j,k] <- equals(i,1)
        # Will be 1 for HN-Knowers
        ind4[i,j,k] <- equals(i,nz)
        ind5[i,j,k] <- ind3[i,j,k]+ind4[i,j,k]*(2+ind2[i,j,k])
                       + (1-ind4[i,j,k])*(1-ind3[i,j,k])
                       * (ind1[i,j,k]+ind1[i,j,k]*ind2[i,j,k]+1)
        piprime[i,j,k] <- piprimetmp[i,j,k,ind5[i,j,k]]
        npiprime[i,j,k] <- piprime[i,j,k]/sum(piprime[i,j,1:gn])
      }
    }
  }
  # Posterior Prediction For Knower Levels
  for (i in 1:nz){
    for (j in 1:gn){
      predz[i,j] ~ dcat(npiprime[i,j,1:gn])
    }
  }
  # Base rate
  for (i in 1:gn){
    pitmp[i] ~ dunif(0,1)
    pi[i] <- pitmp[i]/sum(pitmp[1:gn])
  }
  predpi ~ dcat(pi[1:gn])
  # Priors
  v ~ dunif(1,1000)
  for (i in 1:ns) {
    z[i] ~ dcat(priorz[])
  }
  for (i in 1:nz){
    priorz[i] <- 1/6
  }
}
```

The code `NumberConcept_1.m` or `NumberConcept_1.R` applies the model to the Give-N data used by Lee and Sarnecka (2011), and produces several analyses.

Figure 19.2 shows the base-rate that is inferred, by showing the posterior distribution it generates over the 15 toys. It shows an intuitively reasonable result, giving high probability to small numbers of toys, as well as to the whole set of 15 toys. It is important to understand that this is a latent 15-dimensional parameter that was inferred from the data, and not something that was assumed to explain the data. This is a powerful piece of inference that is straightforward using the Bayesian approach, and makes it theoretically possible to propose a complicated construct like a base-rate as an important component of determining the observed data.

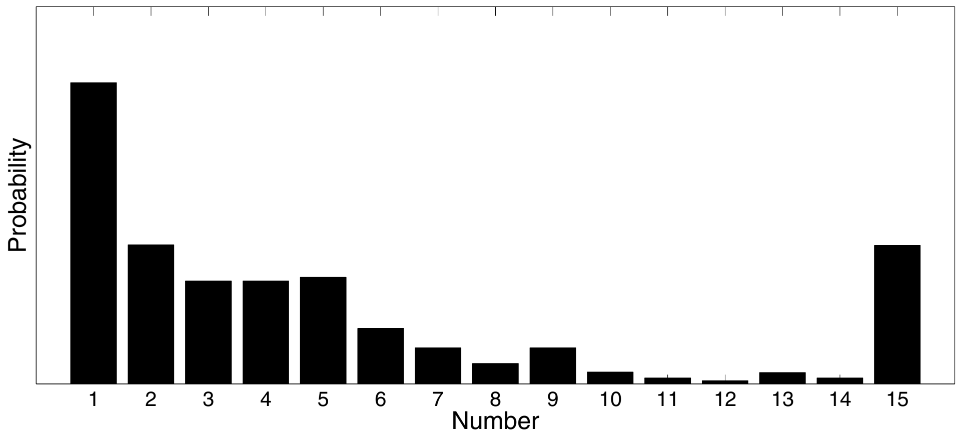


Fig. 19.2 Posterior base-rate for giving $1, \ldots, 15$ toys, inferred by applying the knower-level model to the Give-N data.

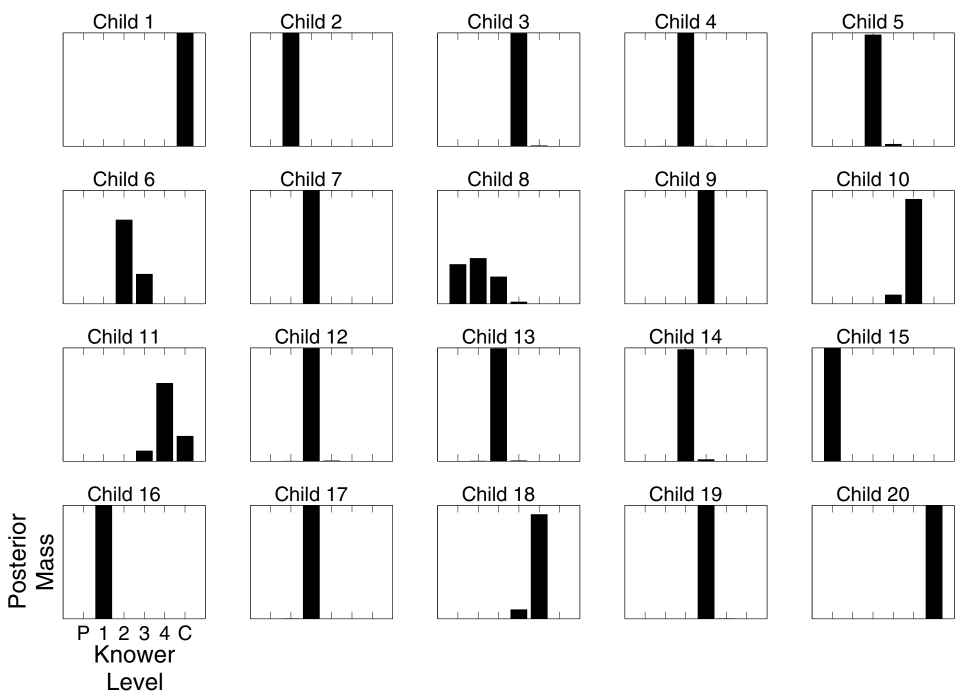


Fig. 19.3 Posterior over knower levels for the 20 children, based on the Give-N data. P="Pre-Number Knower" and C="Cardinality-Principle Knower".

Figure 19.3 shows the posterior distribution over the six possible knower levels for each child. The noteworthy feature of this result is that most of the children are classified with high certainty into a single knower level. There are exceptions, such as children 6, 8, and 11, but, for the most part, there is confidence in a

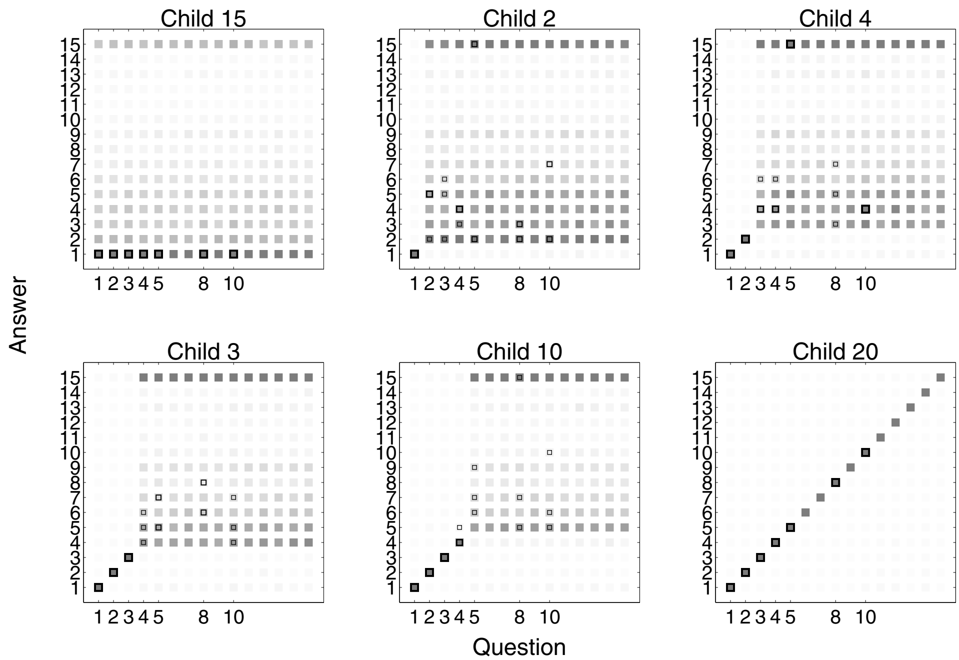


Fig. 19.4 Posterior prediction for six selected children on the Give-N data. The posterior predictive mass of each question and answer combination is shown by shading. The distribution of the child's observed data is shown by squares.

single classification. When inferring a discrete latent variable representing a class, highly-peaked posteriors like this are a good sign that the model is a useful one. When models are badly mis-specified, Bayesian inference naturally does "model averaging" blending over a mixture of possibilities to try and account for the data, making interpretation difficult.

Figure 19.4 shows the relationship between the posterior predictions of the model and the observed behaviors for six children, chosen to span a range of knower levels. Each panel corresponds to a child, with the x-axis giving the question asked and the y-axis giving the answer given. The darkness of the shading in each cell corresponds to the posterior probability that the child will give that many toys when asked that question. The squares show the distribution of the child's observed behavior.

Figure 19.5 shows the same sort of analysis for the posterior predictive for each knower level. The observed data now include every child inferred, using their maximum a posteriori knower level from Figure 19.3.

Exercises

Exercise 19.1.1 Report the posterior for the evidence parameter v.

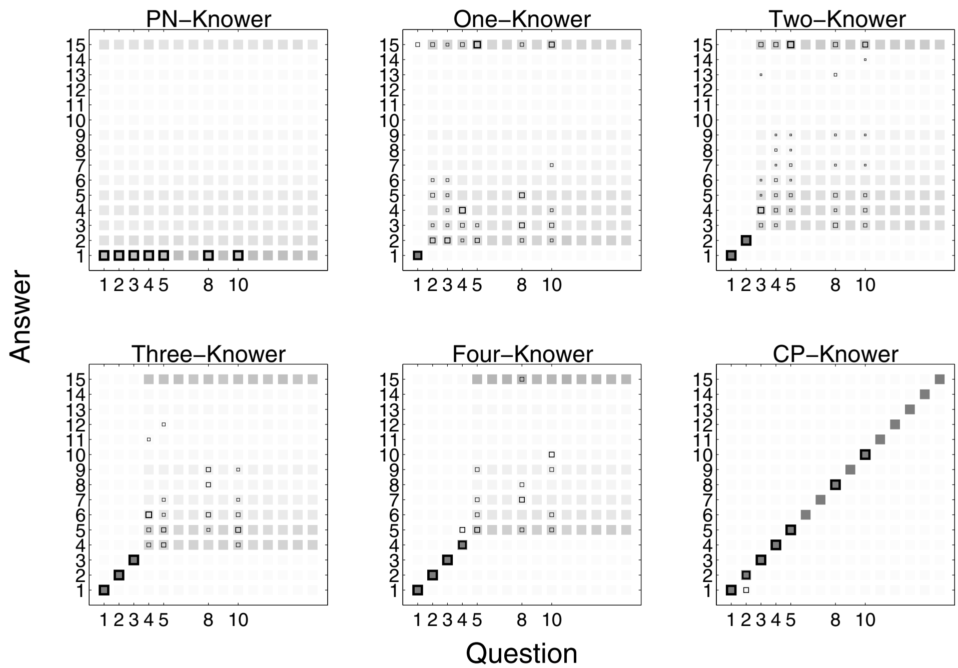


Fig. 19.5 Posterior prediction for the six knower levels on the Give-N data. The posterior predictive mass of each question-and-answer combination is shown by shading. The distribution of the observed data aggregated over every child classified as belonging to that knower level is shown by squares.

Exercise 19.1.2 Interpret the distinctive visual patterns of posterior prediction in Figure 19.4 for each child, explaining how they combine knower-level knowledge, and the task base-rate.

Exercise 19.1.3 What do you think of relying on the maximum a posteriori summary of the posterior uncertainty about knower levels to classify children? What might be a justifiable alternative?

Exercise 19.1.4 The model currently assumes that the final decision is sampled from the distribution over all possible responses, in proportion to the mass associated with each response. Does this probability-matching strategy seem psychologically plausible? What is an alternative model of this part of the decision-making process?

Exercise 19.1.5 Explain why the distribution shown in Figure 19.2 is not exactly the posterior distribution of the base-rate $\boldsymbol{\pi}$. What distribution is shown, how is it related to the posterior of $\boldsymbol{\pi}$, and what are the advantages of presenting the distribution in Figure 19.2?

Box 19.1 Bayesian statistical analysis of Bayesian cognitive models

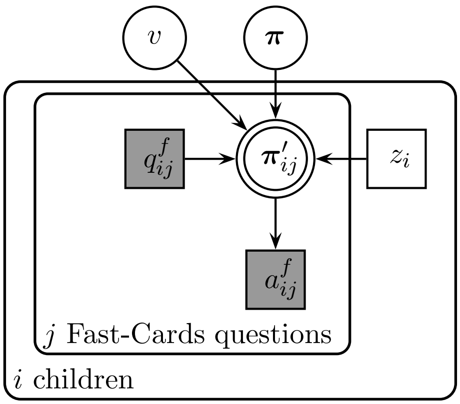
One way to think about Bayesian methods is that they address the problem of drawing inferences over structured models—hierarchical models, mixture models, and so on—from sparse and noisy data. That is obviously a basic challenge for any method of statistical inference. It also, however, seems like a central challenge faced by the mind. We have structured mental representations, and we must deal with incomplete and inherently uncertain information. This analogy suggests that Bayesian statistics can be used not just as a method of analyzing models and data in the cognitive sciences, but also as a theoretical metaphor for developing models of cognition in the first place.

In fact, the "Bayes in the head" theoretical position, which assumes that the mind does Bayesian inference, is an important, and controversial, one in cognitive science (e.g., Chater et al., 2006; Griffiths et al., 2008; Jones & Love, 2011; Tenenbaum et al., 2011). The Bayesian cognitive models developed within this approach usually focus on providing "rational" accounts of psychological phenomena. This means they are mostly (see Sanborn et al., 2010, for an exception) pitched at the computational level within the three-level hierarchy described by Marr (1982), giving an account of why people behave as they do, without trying to account for the mechanisms, processes, or algorithms that produce the behavior, nor how those processes are implemented in neural hardware.

The model of number-development is probably unique among those considered in this book as being interpretable as a Bayesian model of cognition. The base-rate over responses is naturally interpreted as a prior, and the updating that takes place depending on the information presented and the child's knowledge defines a likelihood function. The logic of the model corresponds to applying Bayes' rule to produce a posterior distribution, and observed behavior is sampled from this posterior. This means we are making Bayesian inferences about a Bayesian model of cognition. Perhaps surprisingly, this is rarely done (see Kruschke, 2010b; Lee, 2011b, for critiques), even though all of the advantages Bayesian inference has for analyzing cognitive models still apply to Bayesian cognitive models. For example, it would be difficult to infer the base-rate distribution—that is, the prior the children bring to the task—from behavioral data without using Bayesian methods.

19.2 Knower-level model for Fast-Cards

Applying the graphical model in Figure 19.1 to Fast Cards data only requires changing the base-rate parameter to allow for answers greater than 15. In the data set we are considering, which now consists of the observed q^f_{ij} questions and a^f_{ij} answers for the ith child on their jth question, the largest number given as an answer by any child is 50.



$$
\begin{aligned}
\boldsymbol{\pi} &\sim \text{Dirichlet}(\overbrace{1,\ldots,1}^{50})\\
v &\sim \text{Uniform}(1, 1000)\\
z_i &\sim \text{Categorical}(\tfrac{1}{6},\ldots,\tfrac{1}{6})\\
\pi'_{ijk} &\propto \begin{cases} \pi_k & \text{if } k > z_i\\ v \times \pi_{ijk} & \text{if } k \leq z_i \text{ and } k = q^f_{ij}\\ \frac{1}{v} \times \pi_{ijk} & \text{if } k \leq z_i \text{ and } k \neq q^f_{ij} \end{cases}\\
a^f_{ij} &\sim \text{Categorical}(\boldsymbol{\pi}'_{ij})
\end{aligned}
$$

Fig. 19.6 Graphical model for behavior on the Fast-Cards task according to the Lee and Sarnecka (2010) knower-level model.

The revised graphical model is shown in Figure 19.6, and differs only in terms of the dimensionality of the base-rate parameter $\boldsymbol{\pi}$.

The script `NumberConcept_2.txt` makes this change to the original script. The code `NumberConcept_2.m` or `NumberConcept_2.R` applies the model to the Fast-Cards data. Note that the 50-element base-rate is implemented in an approximate but computationally efficient way, given its sparseness, by considering only those answers actually observed in the data.

Figure 19.7 shows the base-rate that is inferred. It is again intuitively reasonable, with most the greatest posterior mass given to numbers $1,\ldots,5$, which have high frequency in the child's environment. Relatively large mass is given to the remaining single-digit numbers, and most of the remaining mass occurs at so-called "prominent" numbers like 10, 20, 30, and 50 (Albers, 2001).

Figure 19.8 shows the posterior distribution over the six knower levels for each child.

Figure 19.9 shows the analysis for the posterior predictive for each knower level, based on the Fast-Cards data. As before, children have been classified into knower levels based on the mode of their posterior membership shown in Figure 19.8. Note that no child in this data set, for this task, is classified as being a PN-knower, but the model naturally still makes predictions about the behavior of children at this developmental stage.

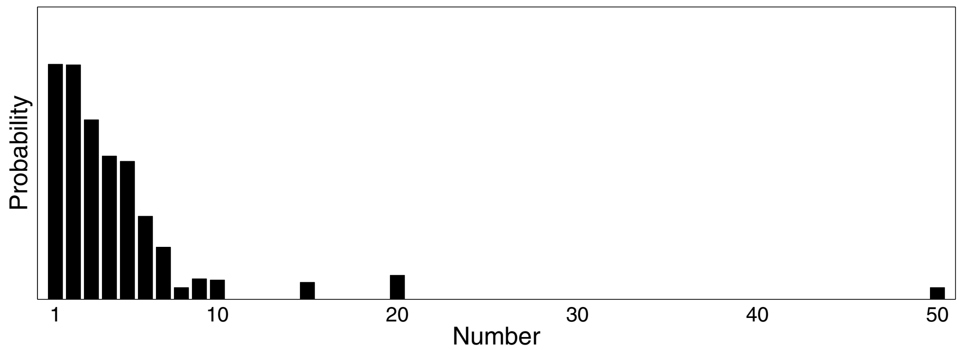


Fig. 19.7 Posterior base-rate for giving the answer $1, \ldots, 50$ toys, inferred by applying the knower-level model to the Fast-Cards data.

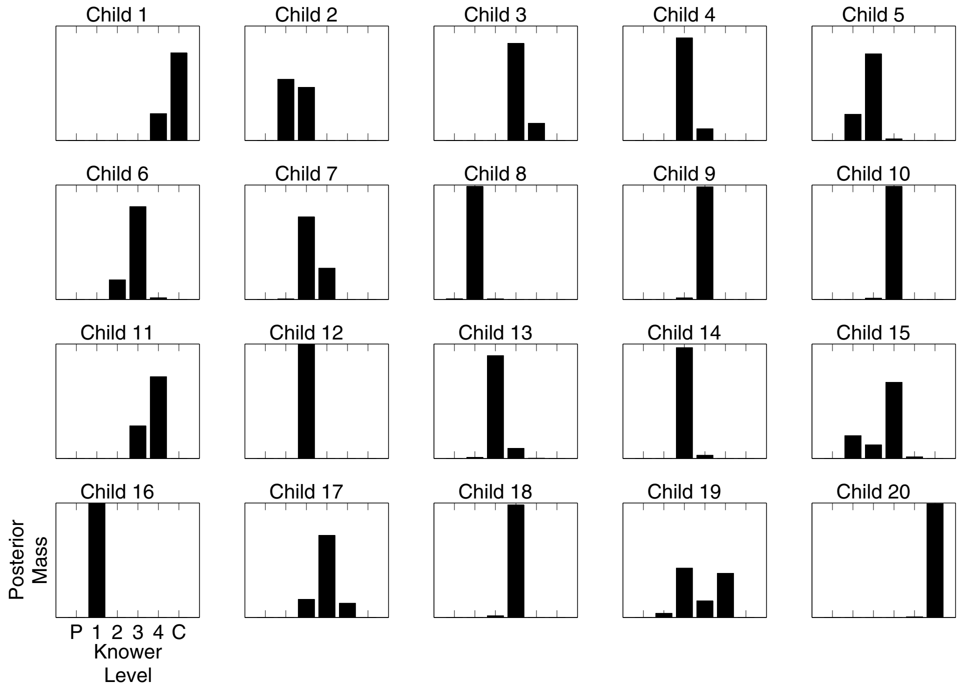


Fig. 19.8 Posterior over knower levels for 20 children, based on the Fast-Cards data. P="Pre-Number Knower" and C="Cardinality-Principle Knower."

Exercises

Exercise 19.2.1 Report the posterior for the evidence parameter v, and compare it to the value found in the Give-N analysis.

Exercise 19.2.2 Compare the posterior distributions over knower levels shown in Figure 19.8 with those inferred using the Give-N data in Figure 19.3.

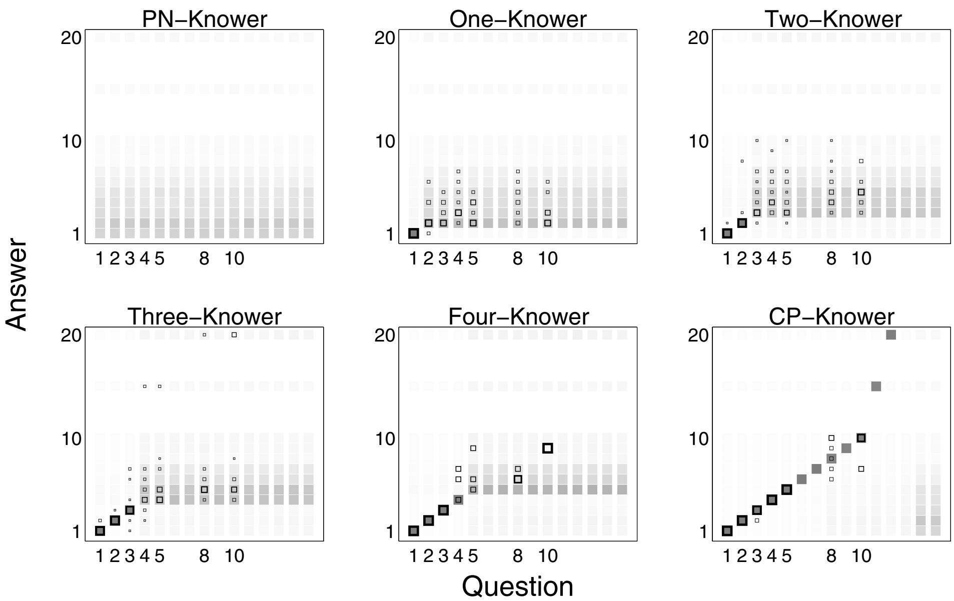


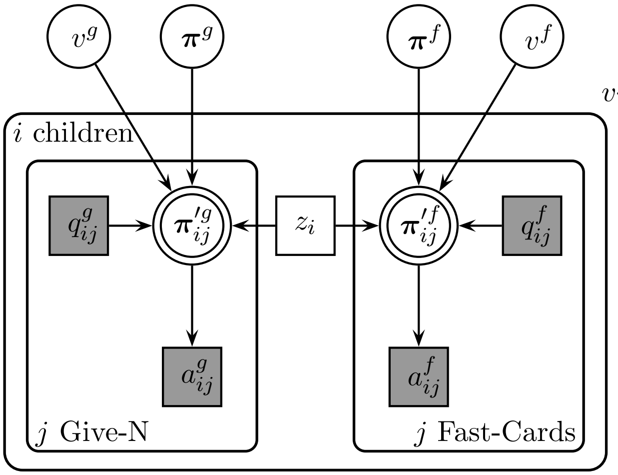
Fig. 19.9 Posterior prediction for the six knower levels on the Fast-Cards data. The posterior predictive mass of each question-and-answer combination is shown by shading. The distribution of the observed data aggregated over every child classified as belonging to that knower level is shown by squares.

Exercise 19.2.3 The uncertainty in the posterior distribution in Figure 19.8 always involves adjacent knower levels (e.g., it is uncertain whether child 2 is a one-knower or two-knower). Does this follow necessarily from the statistical definition of the z knower-level parameter? If so, how? If not, how do the patterns of uncertainty in Figure 19.8 arise?

19.3 Knower-level model for Give-N and Fast-Cards

Because of the within-subjects design of the data, in which each child did both the Give-N and Fast-Cards tasks, it is natural to think about combining both sources of behavioral data to infer knower levels. Conceptually, this is straightforward, with one underlying knower level generating both sets of behavioral data for any given child, according to the specific characteristics of each task.

The graphical model that integrates the two tasks in this way is shown in Figure 19.10. It is visually clear how it combines the graphical models in Figures 19.1 and 19.6, linking them through z_i, the common knower-level parameter. Notice that the base-rates and evidence-value parameters are allowed to be different for the two tasks. Thus, the way the data from the two tasks are modeled in Figure 19.10 corre-



$$
\begin{aligned}
\boldsymbol{\pi}^g &\sim \text{Dirichlet}(\overbrace{1,\ldots,1}^{15})\\
\boldsymbol{\pi}^f &\sim \text{Dirichlet}(\overbrace{1,\ldots,1}^{50})\\
\upsilon^g, \upsilon^g &\sim \text{Uniform}(1, 1000)\\
z_i &\sim \text{Categorical}(\tfrac{1}{6},\ldots,\tfrac{1}{6})\\
\pi'^g_{ijk} &\propto \begin{cases} \pi^g_k & \text{if } k > z_i\\ \upsilon^g \times \pi^g_{ijk} & \text{if } k \le z_i,\ k = q^g_{ij}\\ \frac{1}{\upsilon^g} \times \pi^g_{ijk} & \text{if } k \le z_i,\ k \ne q^g_{ij} \end{cases}\\
\pi'^f_{ijk} &\propto \begin{cases} \pi^f_k & \text{if } k > z_i\\ \upsilon^f \times \pi^f_{ijk} & \text{if } k \le z_i,\ k = q^f_{ij}\\ \frac{1}{\upsilon^f} \times \pi^f_{ijk} & \text{if } k \le z_i,\ k \ne q^f_{ij} \end{cases}\\
a^g_{ij} &\sim \text{Categorical}(\boldsymbol{\pi}'^g_{ij})\\
a^f_{ij} &\sim \text{Categorical}(\boldsymbol{\pi}'^f_{ij})
\end{aligned}
$$

Fig. 19.10 Graphical model for behavior on the Give-N task and the Fast-Cards task, coming from each child's underlying knower level.

sponds to assuming that the latent psychological state of number knowledge is the same in both tasks, but that base-rate responding, and the evidence levels different stimuli provide, can vary in task-specific ways.

The script `NumberConcept_3.txt` implements the graphical model. The code `NumberConcept_3.m` or `NumberConcept_3.R` applies the model to the Given-N and Fast-Cards data:

```
# Knower Level Model Applied to Give-N and Fast-Cards Data
model{
  # Give-N Part
  # Data
  for (i in 1:ns){
    for (j in 1:gnq[i]){
      ga[i,j] ~ dcat(npiprime[z[i],gq[i,j],1:gn])
    }
    # Posterior Predictive
    for (j in 1:gn){
      predga[i,j] ~ dcat(npiprime[z[i],j,1:gn])
    }
  }
  # Model
  for (i in 1:nz){
    for (j in 1:gn){
      for (k in 1:gn){
        piprimetmp[i,j,k,1] <- pi[k]
        piprimetmp[i,j,k,2] <- 1/gv*pi[k]
        piprimetmp[i,j,k,3] <- gv*pi[k]
        ind1[i,j,k] <- step((i-1)-k)
        ind2[i,j,k] <- equals(k,j)
```

```
      ind3[i,j,k] <- equals(i,1)
      ind4[i,j,k] <- equals(i,nz)
      ind5[i,j,k] <- ind3[i,j,k]+ind4[i,j,k]*(2+ind2[i,j,k])
                     + (1-ind4[i,j,k])*(1-ind3[i,j,k])
                     * (ind1[i,j,k]+ind1[i,j,k]*ind2[i,j,k]+1)
      piprime[i,j,k] <- piprimetmp[i,j,k,ind5[i,j,k]]
      npiprime[i,j,k] <- piprime[i,j,k]/sum(piprime[i,j,1:gn])
    }
  }
}
# Fast-Cards Part
# Data
for (i in 1:ns){
  for (j in 1:fnq[i]){
    fa[i,j] ~ dcat(fnpiprime[z[i],fq[i,j],1:fn])
  }
  # Posterior Predictive
  for (j in 1:gn){
    predfa[i,j] ~ dcat(fnpiprime[z[i],j,1:fn])
  }
}
# Model
for (i in 1:nz){
  for (j in 1:gn){
    for (k in 1:fn){
      fpiprimetmp[i,j,k,1] <- fpi[k]
      fpiprimetmp[i,j,k,2] <- 1/fv*fpi[k]
      fpiprimetmp[i,j,k,3] <- fv*fpi[k]
      find1[i,j,k] <- step((i-1)-k)
      find2[i,j,k] <- equals(k,j)
      find3[i,j,k] <- equals(i,1)
      find4[i,j,k] <- equals(i,nz)
      find5[i,j,k] <- find3[i,j,k]+find4[i,j,k]*(2+find2[i,j,k])
                      + (1-find4[i,j,k])*(1-find3[i,j,k])
                      * (find1[i,j,k]+find1[i,j,k]*find2[i,j,k]+1)
      fpiprime[i,j,k]  <- fpiprimetmp[i,j,k,find5[i,j,k]]
      fnpiprime[i,j,k] <- fpiprime[i,j,k]/sum(fpiprime[i,j,1:fn])
    }
  }
}
# Posterior Prediction For Knower Levels
for (i in 1:nz){
  for (j in 1:gn){
    predgaz[i,j] ~ dcat(npiprime[i,j,1:gn])
    predfaz[i,j] ~ dcat(fnpiprime[i,j,1:fn])
  }
}
# Base rates
for (i in 1:fn){
  fpitmp[i] ~ dunif(0,1)
  fpi[i] <- fpitmp[i]/sum(fpitmp[1:fn])
}
for (i in 1:gn){
  pitmp[i] ~ dunif(0,1)
  pi[i] <- pitmp[i]/sum(pitmp[1:gn])
}
```

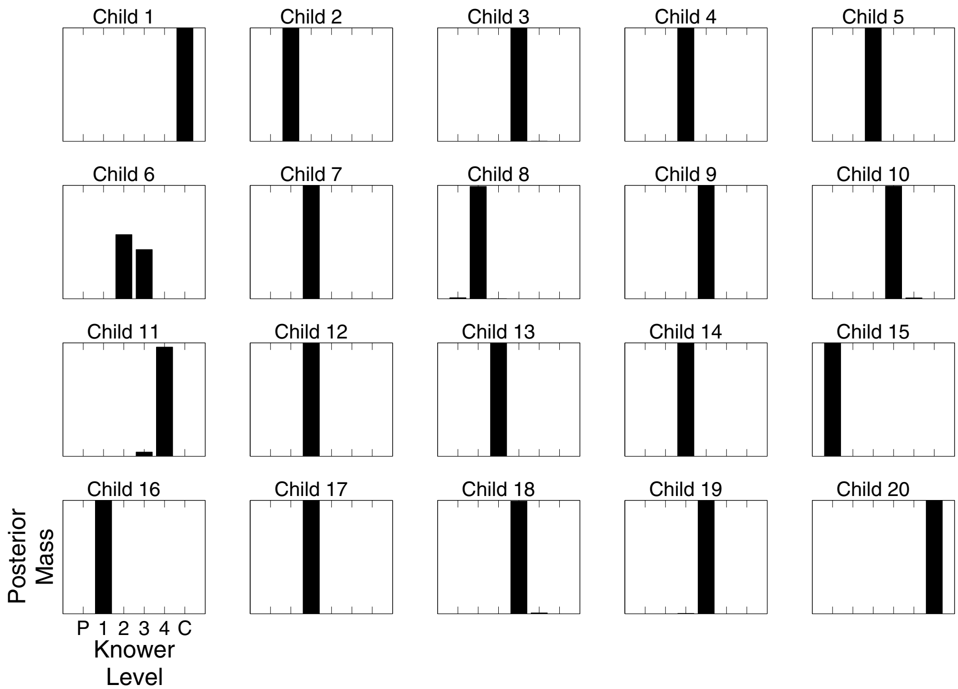


Fig. 19.11 Posterior over knower levels for 20 children, based on both the Give-N and Fast-Cards data. P="Pre-Number Knower" and C="Cardinality-Principle Knower."

```
  predpi  ~ dcat(pi[1:gn])
  predfpi ~ dcat(fpi[1:fn])
  # Priors
  gv ~ dunif(1,1000)
  fv ~ dunif(1,1000)
  for (i in 1:ns) {
    z[i] ~ dcat(priorz[])
  }
  for (i in 1:nz){
    priorz[i] <- 1/6
  }
}
```

Figure 19.11 shows the posterior distribution over the six knower levels for each child. For almost all of the children—with child 6 being the one exception—the posterior distributions have no uncertainty, and children are confidently inferred as belonging to a single knower level. The key point is that, by using both sources of empirical evidence simultaneously, clearer inferences are able to be made than were possible from either alone.

Exercise

Exercise 19.3.1 The behavioral data for Child 18 are detailed in Table 19.1, showing their answers to every question in both tasks. Explain why the inference based on only the Give-N data in Figure 19.3 has most posterior mass on four-knower, but there is some uncertainty, with three-knower also a possibility. Explain why the inference based on only the Fast-Cards data in Figure 19.8 has posterior mass almost entirely on the three-knower possibility. Explain why the combined inference in Figure 19.11 favors three-knower, and why the possible four-knower inference in the Give-N analysis could be viewed as arising from the nature of the task itself, rather than from the actual number knowledge of the child that is of primary interest developmentally.

Table 19.1 Behavior of Child 18 on Give-N and Fast-Cards tasks.

Question	Give-N Answers	Fast-Cards answers
1	1, 1, 1	1, 1, 2
2	2, 2, 2	2, 2, 2
3	3, 3, 3	3, 3, 3
4	4, 5, 4	50, 15, 3
5	9, 7, 6	15, 4, 4
8	7, 15, 5	20, 50, 5
10	10, 6, 5	20, 20, 20