A Bayesian Latent Mixture Approach to Modeling Individual Differences in Categorization Using General Recognition Theory

Irina Danileiko (idanilei@uci.edu)
Michael D. Lee (mdlee@uci.edu)
Department of Cognitive Sciences, University of California, Irvine
Irvine, CA 92617 USA

Michael L. Kalish (mlkalish@syr.edu)
Psychology Department, Syracuse University
Syracuse, NY 13244 USA

Abstract

Decision-bound models of categorization like General Recognition Theory (GRT: Ashby & Townsend, 1986) assume that people divide a stimulus space into different response regions, associated with different categorization decisions. These models have traditionally been applied to empirical data using standard model-fitting methods like maximum likelihood estimation. We implement the GRT as a Bayesian latent mixture model to infer both qualitative individual differences in the types of decision bounds people use, and quantitative differences in where they place the bounds. We apply this approach to a previous data set with two category structures tested under different cognitive loads. Our results show that different participants categorize by applying diagonal, vertical, or horizontal decision bounds. Various types of contaminant behavior are also found, depending on the category structures and presence or absence of load. We argue that our Bayesian latent mixture framework offers a powerful approach to studying individual differences in categorization.

Keywords: category learning; decision bound models; General Recognition Theory; Bayesian inference; latent mixture model

Introduction

Categorization is a fundamental cognitive capability, forming a basis for structuring mental representations to capture meaning and enable prediction. Understanding and modeling how people make categorization decisions is a key challenge for cognitive science. One prominent and successful class of models, known as General Recognition Theory (GRT: Ashby & Townsend, 1986), assumes that categorization decisions are made based on decision bounds. For example, in a categorization task in which a person places a stimulus into one of two categories on each trial, GRT assumes decisions are based on a boundary that splits the stimulus space into two response regions. The decision-bound modeling approach is naturally contrasted with exemplar models of categorization, which assume that people remember all instances of a category and keep them in memory for comparison to novel stimuli to make categorization decisions (e.g., Nosofsky, 1986).

An important issue for any model of categorization relates to the possibility of individual differences. Different people may categorize differently, perhaps as a result of different starting knowledge, different training or learning experiences, different learning strategies, or different decision strategies. Many applications of category learning models ignore individual differences, and deal with behavioral data that are aggregated or averaged over people. Other applications apply models at the level of individual participants (e.g., Nosofsky, 1986). Most recently, there have been some attempts to extend categorization models to include models of individual differences (e.g., Bartlema, Lee, Wetzels, & Vanpaemel, 2014), using Bayesian methods, but these are restricted to exemplar and prototype models.

For decision-bound models, one important potential source of individual differences relates to the use of unidimensional versus multidimensional boundaries. A working hypothesis in the decision bound literature is that simple category structures that separate stimuli based on a single dimension are amenable to simple explicit rules that can be verbalized, whereas more complicated category structures that require integration across the dimensions need associatively learned boundaries that are more implicit. As a result, one focus of modeling individual differences using GRT is to infer whether people use a simple horizontal or vertical bound that partitions stimuli along one stimulus dimension, or a more general linear (diagonal) decision bound that is sensitive to both dimensions (e.g., Ell & Ashby, 2012). This modeling often also considers the possibility of some form of random responding, to identify contaminant participants.

Methodologically, GRT models that incorporate the possibility of individual differences (e.g., Ell & Ashby, 2012; Soto, Vucovich, Musgrave, & Ashby, in press) rely on maximum likelihood methods for parameter estimation, and model selection criteria like the Bayesian Information Criterion. While useful, these methods are limited. Maximum likelihood estimation does not allow for the uncertainty in where a person places a decision bound to be inferred, even though there will always be uncertainty remaining after observing their performance on a limited number of trials. Information criteria attempt to correct for the complexity of different possible decisions strategies, but do so in an approximate way that equates model complexity with counts of parameters. Using Bayesian methods automatically overcomes both of these limitations.

Accordingly, our goal in this paper is to demonstrate a Bayesian latent mixture approach to modeling individual differences within the GRT framework. The structure of this paper is as follows. We first describe the experimental data

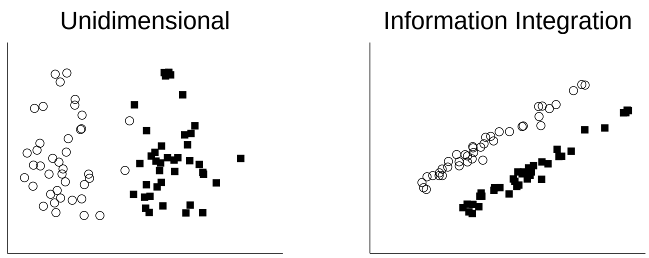


Figure 1: The unidimensional and information integration category structures.

set that we re-analyze. We then describe our formulation of a model, with six possible categorization strategies, in latent mixture terms to allow for individual differences, and its implementation as a graphical model to allow for fully Bayesian inference. We examine the inferences this model makes about individual differences in the decision strategies and decision bounds for the four experimental conditions in the data set. Finally, we discuss the benefits, as well as possible refinements and extensions of our modeling approach.

Zeithamova and Maddox (2006) Experiment

Zeithamova and Maddox (2006) conducted a category learning experiment with four conditions, involving a total of 170 participants in a between-participants design. Each condition consisted of five 80-trial blocks, during which each stimulus was presented once with corrective feedback. The stimuli consisted of Gabor patches, varying in the dimensions of spatial frequency and spatial orientation. The two category structures used are shown in Figure 1, with stimuli in Category A shown in black, and stimuli in Category B shown in white. The unidimensional structure on the left involves a single dimension that separates the categories, while the information integration category structure on the right involves both dimensions. Both of the category structures were presented with and without an additional memory load task, to give the total of four conditions. These category structures and load conditions provide important tests of theories contrasting verbal and implicit category learning systems, and have been replicated and re-analyzed by Newell, Dunn, and Kalish (2010).

Individual Differences GRT Model

Latent mixture models assume that observed data arise from a number of different sources, which combine or mix to produce the overall data. In the case of individual differences in categorization, the different sources correspond to the different categorization strategies used by different people. The latent nature of the mixture means which strategy each individual uses is not known, but rather there are latent parameters assigning people to strategies that need to be inferred.

Table 1: Number and percentage of participants inferred to use vertical, diagonal, or other strategies in each condition.

Condition	Vertical	Diagonal	Other
Unidimensional no load	34 (68%)	6 (12%)	10 (20%)
Unidimensional load	23 (46%)	7 (14%)	20 (40%)
Information integration no load	15 (30%)	25 (50%)	11 (20%)
Information integration load	17 (34%)	17 (34%)	16 (32%)

Modeling Assumptions

The model we develop is tailored to the Zeithamova and Maddox (2006) experiment. It includes six categorization strategies that could be applied to the categorization structures in the experiment. The latent mixture modeling methods we use, however, are general, and could naturally be extended or modified to incorporate different assumptions about individual differences in categorization strategies or types of stimuli.

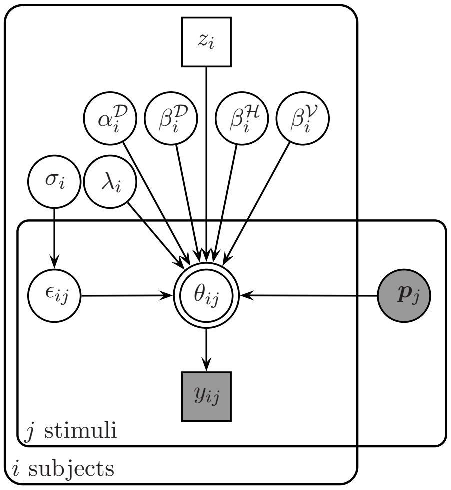
The most obvious categorization strategies to include, in the context of GRT, are vertical boundaries, which are applicable to the unidimensional category structure, and general linear (diagonal) decision boundaries, which are applicable to the information integration structure. We also decided to include a horizontal boundary strategy for completeness.

The other three categorization strategies we consider correspond to contaminant models. In the latent mixture approach, with its focus on generative modeling of observed behavior, contaminants are not "removed" by processing the data on the basis of accuracy or other summary criteria, but by modeling the contaminant behavior itself (Zeigenfuse & Lee, 2010). We allow for three different types of contaminant behavior. One corresponds to guessing, in which the participant is equally likely to categorize any stimulus as belonging to Category A as Category B, and the other two assume that all, or almost all, of the stimuli are repeatedly placed in either Category A or Category B.

Graphical Model Implementation

We formalize the latent mixture of these six strategies as a graphical model, which is a common way of formalizing probabilistic cognitive models (Lee & Wagenmakers, 2013). Graphical models have the conceptual advantage of providing an intuitive visualization of the generative process assumed to produce data, and its dependence on psychological variables represented by parameters. They have a practical advantage of making it straightforward to do fully Bayesian inference using computational sampling methods.

In a graphical model, nodes represent the data, parameters, and other variables of interest, and their dependencies are indicated by the graph structure. Latent parameters are shown as unshaded nodes while the observed parameters and the data are shown as shaded nodes. Discrete variables are represented with square nodes while continuous variables are represented with circular nodes. Stochastic variables are shown as single-bordered nodes while deterministic ones are shown



$$\begin{aligned}
\alpha_i^{\mathcal{D}} &\sim \text{Uniform}(-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \\
\beta_i^{\mathcal{H}}, \beta_i^{\mathcal{V}}, \beta_i^{\mathcal{D}} &\sim \text{Uniform}(-\tfrac{1}{2}, \tfrac{1}{2}) \\
\sigma_i &\sim \text{Uniform}(0, 1) \\
\lambda_i &\sim \text{Uniform}(0, 10) \\
\epsilon_{ij} &\sim \text{Gaussian}(0, \tfrac{1}{\sigma_i^2}) \\
z_i &\sim \text{Categorical}(\tfrac{1}{6}, \ldots, \tfrac{1}{6}) \\
\theta_{ij} &= \begin{cases} \frac{1}{2} & \text{if } z_i = 1 \\ 0.01 & \text{if } z_i = 2 \\ 0.99 & \text{if } z_i = 3 \\ \Phi([p_{j1} - \beta_i^{\mathcal{V}} + \epsilon_{ij}]/\lambda_i) & \text{if } z_i = 4 \\ \Phi([p_{j2} - \beta_i^{\mathcal{H}} + \epsilon_{ij}]/\lambda_i) & \text{if } z_i = 5 \\ \Phi([\frac{p_{j2} - \tan\alpha_i^{\mathcal{D}} p_{j1} - \beta_i^{\mathcal{D}}}{\tan^2\alpha_i^{\mathcal{D}} + 1} + \epsilon_{ij}]/\lambda_i) & \text{if } z_i = 6 \end{cases} \\
y_{ij} &\sim \text{Bernoulli}(\theta_{ij})
\end{aligned}$$

Figure 2: Graphical model representation of our model for inferring individual differences in categorization using GRT.

as double-bordered nodes. The rectangular plates show independent replications of the graph structure within the model.

Figure 2 shows the graphical model representation of our latent mixture model. The discrete parameter z_i acts as strategy selection variables, indicating which of the six possible decision strategies is used by the ith participant. This determines how θ_{ij}, the probability that the ith participant categorizes the jth stimulus into Category A, is calculated. If $z_i = 1$, indicating the guessing strategy, then $\theta_{ij} = \frac{1}{2}$. For the repeated-choice contaminant models, with $z_i = 2$ and $z_i = 3$, θ_{ij} is assumed to be 0.01 and 0.99, respectively.

The other three strategies involve decision bounds, and are parameterized. If $z_i = 4$, then a vertical decision bound is used with x-axis value $\beta_i^{\mathcal{V}}$. The jth stimulus, with coordinate location $\boldsymbol{p}_j = (p_{j1}, p_{j2})$ is more likely to be in Category A if it lies to the right of this boundary. Following GRT, we assume the difference between the boundary and the location of the stimulus on the relevant dimension is corrupted by additive Gaussian noise, and use a probit function to map noisy distances to response probabilities. Formally, this gives

$$\theta_{ij} = \Phi\big([p_{j1} - \beta_i^{\mathcal{V}} + \varepsilon_{ij}]/\lambda_i\big)$$

for the vertical boundary strategy, where $\varepsilon_{ij} \sim \text{Gaussian}\big(0, \frac{1}{\sigma_i^2}\big)$ is the noise term, parameterized by the standard deviation of the noise σ_i for the ith participant, and λ_i scales the probit transfer function, controlling how categorization probabilities for the ith participant vary with the distance of a stimulus from the decision bound. By applying the noise term directly to the distance, we conceive of it combining both the variability in the perceptual information provided by the stimulus, and variability in memory for the decision bound (Maddox & Ashby, 1993). The horizontal strategy when $z_i = 5$ is formalized analogously, in terms of a y-axis value $\beta_i^{\mathcal{H}}$ as

$$\theta_{ij} = \Phi\big([p_{j2} - \beta_i^{\mathcal{H}} + \varepsilon_{ij}]/\lambda_i\big).$$

The general linear (diagonal) decision bound when $z_i = 6$ is parameterized by the angle of the slope $\alpha_i^{\mathcal{D}}$ and intercept $\beta_i^{\mathcal{D}}$. Using standard geometric results giving the distance from a point to a line parameterized this way gives

$$\theta_{ij} = \phi\big([\frac{p_{j2} - \tan\alpha_i^{\mathcal{D}} p_{j1} - \beta_i^{\mathcal{D}}}{\tan^2\alpha_i^{\mathcal{D}} + 1} + \varepsilon_{ij}]/\lambda_i\big).$$

Since $\alpha_i^{\mathcal{D}}$ is an angle, it is natural to make its prior uniform over all possibilities, so that $\alpha_i^{\mathcal{D}} \sim \text{Uniform}\big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$. To simplify the setting of priors, we normalized the coordinate locations of the stimuli to lie in a square of unit length, centered on the origin. This makes $\beta_i^{\mathcal{V}}, \beta_i^{\mathcal{H}}, \beta_i^{\mathcal{H}} \sim \text{Uniform}\big(-\frac{1}{2}, \frac{1}{2}\big)$ reasonable vague assumptions. We also use vague uniform priors on the noise standard deviation of probit scaling parameters, with $\sigma_i \sim \text{Uniform}\big(0, 1\big)$, $\lambda_i \sim \text{Uniform}\big(0, 10\big)$, and give each of the possible categorization strategies equal prior probability $z_i \sim \text{Categorical}\big(\frac{1}{6}, \ldots, \frac{1}{6}\big)$

Modeling Results

We implemented the graphical model in Figure 2 in JAGS (Plummer, 2003), and applied to to the data from the final block of trials for every participant in all four conditions of Zeithamova and Maddox (2006). Our results are based on 6 independent chains with 10,000 samples each, collected after discarding the first 50,000 burn-in samples from each chain, and thinning by collecting only every 5th sample. The chains were inspected visually for convergence, and using the standard $\hat{R}$ statistic. In a few cases, individual chains that had clearly failed to mix were discarded.

Table 1 summarizes the overall results, listing how many participants are inferred as using the vertical, diagonal, or other categorization strategy—grouping the contaminant and horizontal strategies as "other" strategies, since they do not allow for accurate categorization behavior—for all four experimental conditions. Individual-level results are shown in

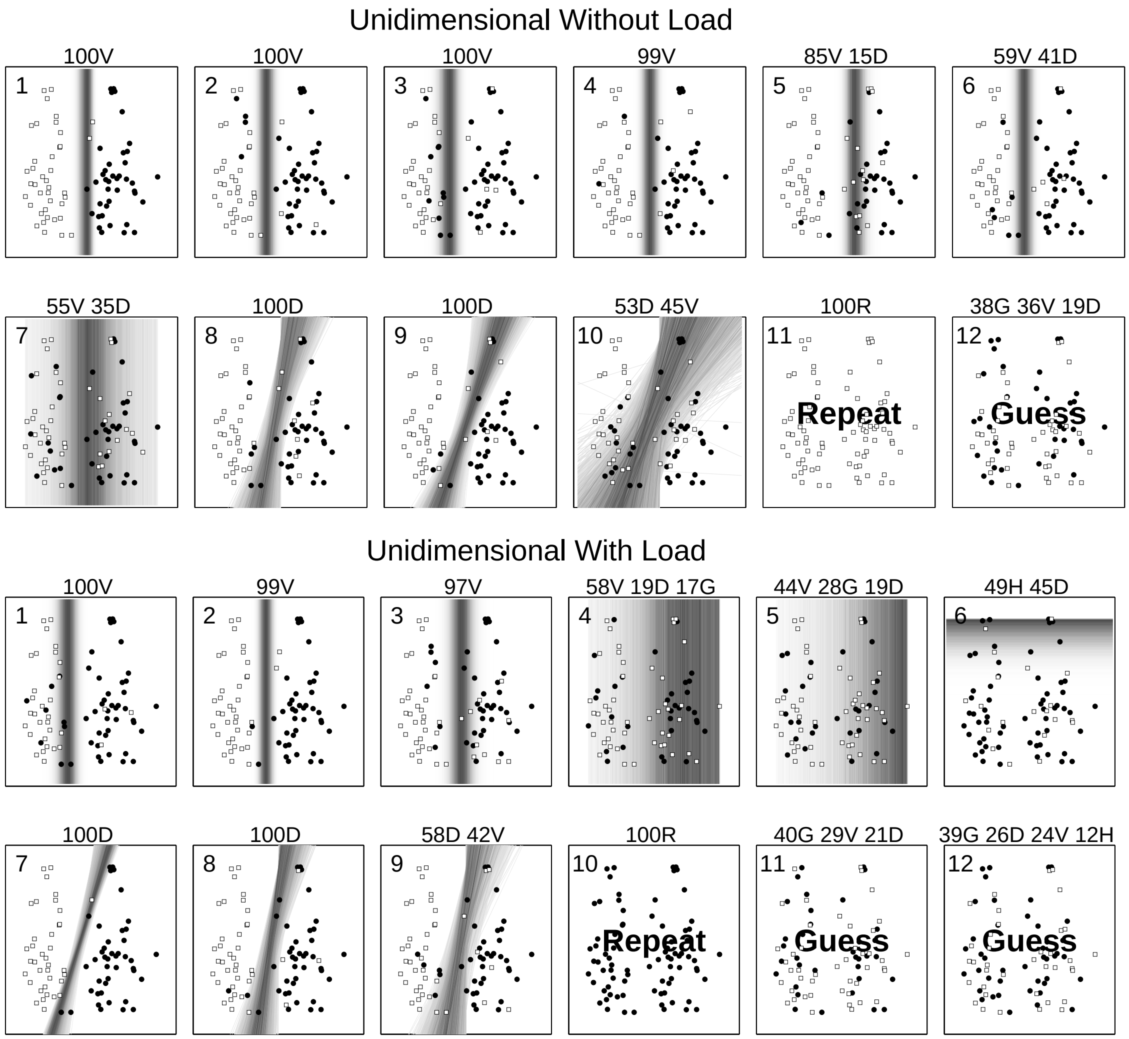


Figure 3: Inferences about catergorization strategies and decision bounds for 12 representative participants in the unidimensional with no load (top two rows) and unidimensional with load (bottom two rows) conditions.

Figures 3 and 4 for the unidimensional and information integration category structures. Each panel corresponds to a participant, and a total of 12 representative participants are shown for each condition. Results for all participants in all conditions are available at `https://osf.io/dmjs7`.

Each panel shows the categorization decisions made by the participant, with stimuli placed in Category A shown as black circles, and stimuli placed in Category B shown as white squares. The label above the panel summarizes the posterior of the z_i indicator variable, using the abbreviations D=Diagonal, V=Vertical, H=Horizontal, G=Guess, R=Repeat. Any strategy with more than 10% posterior mass is included so that, for example, "38G 36V 19D" means the posterior probabilities were 0.38 for the guessing strategy, 0.36 for the vertical bound strategy, and 0.19 for the general linear (diagonal) strategy, with the small remaining posterior mass distributed among the other strategies.

Each panel also shows the posterior distribution for the inferred decision boundary for the participant. This is based on the maximum a posteriori (MAP) strategy—that is, the one with the greatest posterior probability—and shows the posterior of each boundary by shading, with darker shades indicating more likely boundaries. For those participants inferred to be contaminants, the label "Repeat" or "Guess" is shown.

The first 7 participants shown for unidimensional without load condition in Figure 3 are inferred to be using a vertical decision boundary, the next 3 are inferred to use diagonal boundaries, and the last two participants are inferred to be contaminants. The inferred locations of these boundaries vary across the participants with, for example, the vertical boundary for the 3rd participant much further to the left than the vertical boundary for the 5th participant. These inferences are consistent with the different categorization decisions made by the participants since, for example, the 3rd participant makes Category A decisions for stimuli much further to the left.

There are different levels of uncertainty in the inferences

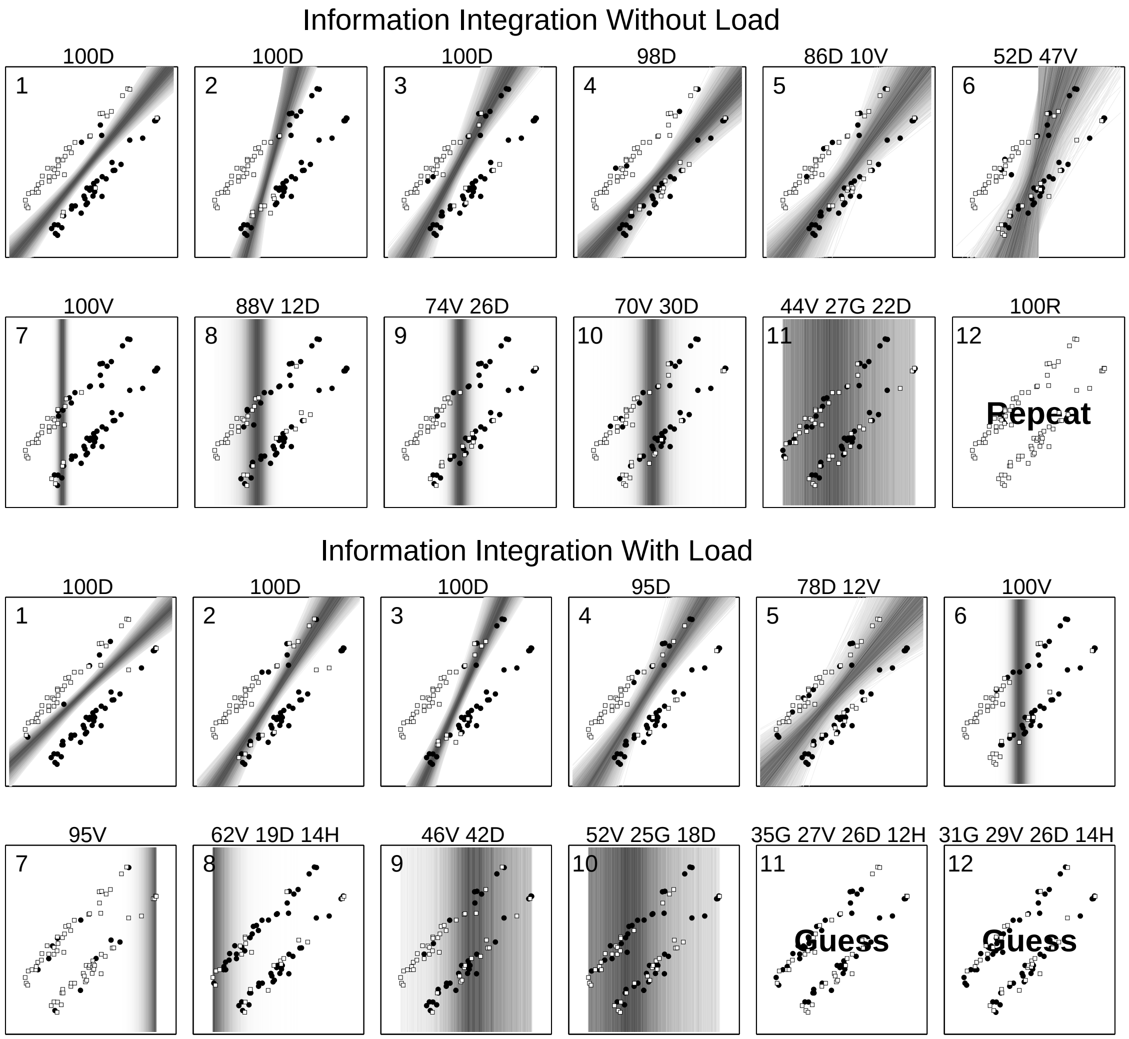


Figure 4: Inferences about categorization strategies and decision bounds for 12 representative participants in the information integration with no load (top two rows) and information integration with load (bottom two rows) conditions.

about the boundaries participants use. The vertical boundary for the seventh participant, for example, is much more uncertain than the vertical boundaries inferred for the first 6 participants. It is also relatively uncertain whether this participant used a vertical or diagonal decision bound, with their posterior probabilities of 0.55 and 0.35 respectively. The complete representation of uncertainty about inferences, at both the level of which categorization strategy a participant used, and where they placed their decision boundaries for these strategies, is an important advantage of the Bayesian approach.

The representative participants chosen for the unidimensional with load condition in Figure 3 show the greater variability in the categorization strategies inferred to be used in this condition. There are more contaminant participants, and less consistent use of vertical boundaries. More participants are inferred to use diagonal and even horizontal boundaries, and the uncertainty in the location of these boundaries is greater. These differences are naturally attributed to the effects of cognitive load.

A similar pattern of modeling results is found for the information integration conditions in Figure 4. The first 6 representative participants in the no load condition are inferred to use diagonal decision boundaries, but there are significant individual differences in where these boundaries are placed, and the certainty of the inferences about these locations. The next 5 representative participants are inferred to use a simpler vertical decision bound categorization strategy, while the 12th participant is inferred to be a repeat contaminant.

The 5th and 6th participants in this condition—with significant posterior uncertainty between the diagonal and vertical categorization strategies—highlight a powerful property of the latent mixture approach. One way to conceive of the latent mixture in cases like this is not as an inference between two incommensurable possibilities, but as a single general model with a theoretically rich prior. Since a vertical boundary is a special case of a diagonal boundary, the 5th participant's posterior uncertainty of 0.86 for diagonal and 0.10 for vertical boundaries could be interpreted as a diagonal boundary with 86% of the overall posterior coming from the slope and intercept, and an additional 10% coming from the posterior for the vertical boundary. This corresponds to inferring only a diagonal boundary, but with a prior that places significant prior density on boundaries with infinite slope.

The 11th participant in the information integration without load condition shows significant uncertainty about the categorization strategy. Their decisions are somewhat consistent with a vertical boundary towards the middle-left of the stimuli, but also appear somewhat consistent with guessing. The posterior uncertainty in their z_i model indicator parameter shows that none of the categorization strategies assumed in modeling provide a good account of their behavior, and suggests the need for further model development. One possibility in cases like this is that the participant changed strategies during the trial block, and some sort of more complicated multi-strategy model is required.

The representative participants shown for the information integration with load condition are inferred to use both diagonal and vertical decision bounds, but there is more use of the simpler vertical strategy, and more contaminant behavior. Examples like the 7th participant, who is confidently inferred to use a vertical bound near the far right of the stimulus space, highlight the possibility of improving the current model using more informative priors. The current inference is that the participant places a near-degenerate decision bound to distinguish just one or two stimuli from the others. It is debatable whether this use is consistent with the theoretical motivations of decision bound models, which usually expect significant number of stimuli to be separated. One way to include this theoretical assumption, only possible in a Bayesian approach, is to change the prior on β_i^v to capture the expectation that the vertical boundary will be close to the middle of the stimulus space (Vanpaemel & Lee, 2012).

Conclusion

Latent mixture modeling is a general framework for modeling individual differences in human behavior. In this paper, we applied the approach to the challenge of understanding the different decision bounds people might use to categorize stimuli, consistent with previous theorizing and modeling applying General Recognition Theory. We developed a latent mixture with six possible categorization strategies, and applied it to previously modeled data reported by Zeithamova and Maddox (2006).

Our results highlight a number of powerful features of the Bayesian approach. It represents uncertainty about model use and the parameterization of those models, in contrast to traditional inference methods like maximum likelihood estimation or least-squares fitting. This means inferences are sensitive to all aspects of the complexity of the various possible categorization strategies. Importantly, it is possible to propose any sort of categorization strategy, including explicit models of contaminant behavior. We focused on linear decision bounds from GRT, but other more general decision bounds like bilinear or quadratic bounds, or different categorization models such as exemplar models, could be incorporated in the graphical modeling framework. It is also straightforward to extend our account of individual differences to include hierarchical structure, allowing both variability in classification strategies and parameterizations of those strategies to be modeled (Bartlema et al., 2014). These possibilities provide a natural direction for future research in understanding the individual differences in how people learn and use category structures.

Acknowledgments

This work was supported by NSF Award 1431635.

References

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–79.

Bartlema, A., Lee, M., Wetzels, R., & Vanpaemel, W. (2014). A Bayesian hierarchical mixture approach to individual differences: Case studies in selective attention and representation in category learning. *Journal of Mathematical Psychology*, *59*, 132–150.

Ell, S. W., & Ashby, F. G. (2012). The impact of category separation on unsupervised categorization. *Attention, Perception & Psychophysics*, *74*, 466–475.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception and Psychophysics*, *53*, 49–70.

Newell, B. R., Dunn, J. C., & Kalish, M. L. (2010). The dimensionality of perceptual category learning: A state-trace analysis. *Memory & Cognition*, *38*, 563–581.

Nosofsky, R. M. (1986). Attention, similarity and the idenitification-categorization relationship. *Journal of Experimental psychology: General*, *115*, 39–57.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, & A. Zeileis (Eds.), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing.* Vienna, Austria.

Soto, F. A., Vucovich, L., Musgrave, R., & Ashby, F. G. (in press). General recognition theory with individual differences: A new method for examining perceptual and decisional interactions with an application to face perception. *Psychonomic Bulletin & Review.*.

Vanpaemel, W., & Lee, M. D. (2012). Using priors to formalize theory: Optimal attention and the Generalized Context Model. *Psychonomic Bulletin & Review*, *19*, 1047–1056.

Zeigenfuse, M. D., & Lee, M. D. (2010). Finding the features that represent stimuli. *Acta Psychologica*, *133*, 283–295.

Zeithamova, D., & Maddox, W. T. (2006). Dual task interference in perceptual category learning. *Memory & Cognition*, *34*, 387–398.