*Memory & Cognition*
*2009, 37 (4), 394-413*
*doi:10.3758/MC.37.4.394*

# Two pathways to stimulus encoding in category learning?

Tyler Davis, Bradley C. Love, and W. Todd Maddox
*University of Texas, Austin, Texas*

Category learning theorists tacitly assume that stimuli are encoded by a single pathway. Motivated by theories of object recognition, we evaluated a dual-pathway account of stimulus encoding. The part-based pathway establishes mappings between sensory input and symbols that encode discrete stimulus features, whereas the image-based pathway applies holistic templates to sensory input. Our experiments used rule-plus-exception structures, in which one exception item in each category violates a salient regularity and must be distinguished from other items. In Experiment 1, we found discrete representations to be crucial for recognition of exceptions following brief training. Experiments 2 and 3 involved multisession training regimens designed to encourage either part- or image-based encoding. We found that both pathways are able to support exception encoding, but have unique characteristics. We speculate that one advantage of the part-based pathway is the ability to generalize across domains, whereas the image-based pathway provides faster and more effortless recognition.

Theories of categorization treat object representations as input and assume that prior processes have extracted the relevant properties from the stimulus. However, there have been few systematic investigations of this extraction process. The purpose of the present article is to investigate whether there may be different ways to *encode* objects in categorization, and how they may differ from one another. Specifically, we evaluated whether dual-pathway accounts of stimulus encoding, one part-based and the other image-based, could lead to novel insights into how stimulus encoding is accomplished in category learning.

Whereas encoding is a central aspect of all categorization decisions, certain occasions place greater demands on encoding—for example, cases in which people must use subtle cues to identify individuals and subcategories that require special (i.e., exceptional) responses (cf. Logothetis & Sheinberg, 1996). For example, pet owners must be able to recognize their dogs among others of the same breed, and expert bird-watchers may spend hours searching for a rare species of finch, say, among more common birds with similar appearances.

Category learning paradigms have been used to study these types of learning problems in the laboratory. Rule-plus-exception designs are particularly well suited to studying the processes involved in the encoding, storage, and use of exception knowledge (e.g., Palmeri & Nosofsky, 1995; Sakamoto & Love, 2004). In these category structures, most items can be categorized according to an imperfect rule; for example, small items may tend to belong in one category, large items in another. The categories used in these designs also contain exception items that belong to the opposing category. Like the examples above, the exception items must be stored separately, and require different responses from other items that are otherwise very similar (e.g., the rule-following stimuli).

Models that can account for human performance in these tasks, such as SUSTAIN (Love, Medin, & Gureckis, 2004) and RULEX (Nosofsky, Palmeri, & McKinley, 1994), correctly predict enhanced recognition memory for the exception items (Palmeri & Nosofsky, 1995; Sakamoto & Love, 2004) by storing them separately from the other items in their respective categories. These models tacitly suggest that there is one way in which exception items are encoded, although the models differ in their underlying assumptions regarding how this is accomplished. For example, SUSTAIN always encodes an exception item in its own cluster at a particular location in multidimensional space, whereas RULEX always encodes an exception by storing a subset of the item's features.

Another possibility is that different pathways can be used to encode items in categorization tasks, depending on the properties of the stimulus, knowledge of the observer, cognitive demands of the task, and so on. Although multiple-system models are common in the category learning literature (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Erickson & Kruschke, 1998), these models also assume a common stimulus encoding. Like SUSTAIN and RULEX, they do not determine whether there are multiple routes to encoding exception items, probably because models of category learning have largely sidestepped the question of how object representations are constructed. Whereas models may differ in how they encode a stimulus, encoding assumptions have rarely been provided as grounds for choosing between models.

**T. Davis, thdavis@mail.utexas.edu**

Because of this lack of research into stimulus encoding, researchers often use different types of stimuli interchangeably. One distinction that tends to be overlooked is the difference between discrete and continuous-valued stimuli. Most category learning models[1] make no distinction between these types of stimuli, and research using one type is assumed to generalize to the other. By contrast, the data presented below imply a variety of important differences between these two types of stimuli, stemming from the ability of observers to encode discrete stimuli in a manner that allows them to stand in a fixed relationship to stored stimulus representations. Put simply, discrete stimuli can be mapped onto nominal values, and this mapping is likely to be stable from one viewing to the next. On the other hand, a vast literature on absolute judgment suggests that the perception of continuous-valued stimuli is heavily influenced by context (for a review, see Stewart, Brown, & Chater, 2005).

Another encoding issue that the category learning literature has largely neglected is the role that a learner's experience plays in shaping stimulus encoding (but see Schyns, Goldstone, & Thibaut, 1998). Indeed, whether or not a stimulus is viewed as discrete or continuous, or whether it can be individuated from other perceptually similar items, likely depends on specific aspects of the observer's history with the class of stimuli. For example, the ability of experienced carpenters to encode the different sizes of nails may be attributable to expert carpenters' ability to map these sizes onto discrete values (e.g., 2¼ in.). In contrast, novices are limited to using continuous encoding strategies and relative comparisons (e.g., "larger than").

In contrast to the category learning literature, work in object recognition focuses to a larger extent on how objects are encoded. Proposals for object representation have been diverse, including image-based (e.g., Poggio & Edelman, 1990) and symbolic, part-based, structural-description models (e.g., Biederman, 1987). More recent hybrid proposals preserve important aspects of both symbol and template approaches, such as the compositionality afforded by symbols and the speed and automaticity inherent in template models (Hummel & Stankiewicz, 1996; Ullman, Vidal-Naquet, & Sali, 2002; Zhang & Cottrell, 2005; see Palmeri & Tarr, 2008, for a review).

Given the overlap in topics and methodology between work in category learning and object recognition (Palmeri & Gauthier, 2004), substantial opportunity exists for cross-fertilization. In the following, we import concepts from the object recognition literature and evaluate how they may be used to make predictions for category learning tasks. The two-pathway framework for understanding how items are encoded in category learning tasks that we develop here is motivated by these hybrid approaches to encoding from the object recognition literature.

We use the term *pathway* here to emphasize that we are interested in different means (i.e., representations/encoding strategies) that can be used to achieve the same ends (i.e., categorization or recognition of an object). After a brief discussion of these two pathways, we will use this framework to predict how exception items are encoded in a series of category learning experiments that use rule-plus-exception designs.

### The Image-Based Pathway

One pathway for encoding objects in category learning is through the development of *image-based* representations, which are often characterized by the development of fine-grained templates that support automatic and rapid recognition of objects (see Gauthier & Tarr, 2002). These templates are often holistic, in that how well an object fits can only be assessed for the object as a whole; individual part matches and mismatches cannot be evaluated (cf. Farah, Wilson, Drain, & Tanaka, 1998). Template accounts of expertise correctly predict advantages for whole objects and canonical orientations (Diamond & Carey, 1986; Gauthier, Williams, Tarr, & Tanaka, 1998). However, because these templates are stimulus/domain specific and holistic, they are unlikely to be useful for encoding different classes of objects. Image-based representations are acquired following extensive experience, and are therefore unlikely to be operable during brief, single-session category learning experiments.

### The Part-Based Pathway

Another pathway for encoding objects in category learning is the *part-based* pathway. Part-based encoding involves the mapping of a stimulus's features onto a compositional and symbolic vocabulary. In contrast to the image-based pathway, this pathway creates representations consisting of semantically evaluable parts that can be used in different domains (Biederman, 1987; Fodor, 1975).

For part-based encoding to be successful, its users must be able to decompose and represent the aspects of an object that differentiate it from others; that is, it requires a vocabulary sufficiently rich for the context, as well as attentional resources to bind an object's features with symbols. Since these processes are likely analytic (Foard & Kemler Nelson, 1984; Regehr & Brooks, 1993), and possibly top-down (e.g., Ahissar & Hochstein, 2002), part-based encoding should be more time consuming and effortful than image-based encoding (see also Hummel & Stankiewicz, 1996).

### Predictions for Category Learning

Table 1 reviews the characteristics of the part- and image-based pathways for object encoding in category learning tasks. Considering the characteristics of these two path-

**Table 1**
**Pathways to Object Encoding**

| | Image-Based | Part-Based |
|---|---|---|
| Representation | Holistic | Featural/symbolic |
| Acquisition | Extensive direct experience | Culture, instruction, feature learning |
| Functional role | Permanent and rapid object encoding capacity | Bootstrapping performance in the absence of expertise, domain-general object encoding, etc. |
| Characteristics | Automatic and domain specific | Algorithmic and compositional. Requires attentional resources and a sufficient featural vocabulary. |

ways, we derive several predictions for category learning performance in tasks using rule-plus-exception designs.

First, we predict that subjects will need to successfully engage one of the two proposed pathways to master exception items in rule-plus-exception learning tasks. In these tasks, exceptions must be individuated from the category to which they belong. To successfully encode the exception items, subjects will require a vocabulary on which to map them that is rich enough to represent their distinguishing features (i.e., part-based encoding), or they will require exception-specific templates (i.e., image-based encoding). In contrast, individual rule-following items are not as demanding as exceptions, since they can likely be massed together at a lower resolution without harming performance. Extensive training involving rule-following items (encouraging image-based encoding) or having the ability to encode items' individual features via the part-based pathway should have little to no impact on rule-item performance.

Second, in brief category learning tasks, the ability to encode exceptions accurately should depend on the availability of the part-based pathway. This prediction follows from the framework *and* the first prediction; image-based representations should not be available for an *individual* item early in learning, because in order to develop them, extensive experience is required. Instead, accurate encoding of an exception should depend on the ability of a subject to map its parts onto a discrete vocabulary. Since one difference between continuous and discrete stimuli is how well they facilitate mapping onto nominal values, the framework suggests that, early in learning, whether subjects can successfully encode exceptions depends on whether stimuli are constructed from continuous or discrete-valued features. (Note that this difference, highlighted by the present framework, is not anticipated by the categorization literature, which largely treats continuous and discrete stimuli as interchangeable.)

Finally, given appropriate training, people should be able to develop competence along either pathway, and should exhibit behavior consistent with the use of the pathway for which they were trained. When a task encourages part-based encoding, people should be able to acquire a feature set that allows them to encode task-relevant differences between the stimuli (Schyns et al., 1998). Likewise, when a task encourages holistic processing, people should be able to develop image-based representations for encoding the stimuli.

Important behavioral differences should exist between subjects trained to use different pathways. One potential difference is that image-based encoding should be relatively rapid and automatic once it has fully developed, whereas part-based encoding should be more effortful and time consuming. However, following from the compositional nature of symbolic processing, part-based encoding should be applicable to a wider range of stimuli when a sufficiently rich vocabulary is acquired or provided.

**Overview of Experiments**

We report the data from three experiments that rigorously tested a number of predictions from the two-pathway framework. Each of these experiments involved training subjects on a rule-plus-exception task, and comparing whether they were able to recognize exception items at a higher rate than they could recognize rule-following items matched in terms of frequency of presentation ("frequency-matched rule-following items"). Enhanced exception recognition in these tasks suggests that subjects were able to individuate the exception items; it is, therefore, evidence that subjects were able to successfully encode these items' distinguishing features. The effect on learning and recognition performance of between-subjects manipulations, such as the availability of easily encodable item parts (Experiment 1) and training types (Experiments 2 and 3), allows us to infer differences in item representations when quantitatively examined via null hypothesis significance testing and formal modeling (discussed below).

Experiment 1 was designed to test the hypothesis that the availability of part-based encoding strategies is crucial, early in learning, to learning and recognition performance. We compared learning and recognition performance across conditions in which subjects were briefly trained in a task using continuous stimuli with or without providing an additional set of verbalizable labels. Consistent with the advantages afforded by part-based encoding, we found that subjects receiving these added discrete features were able to recognize the exception items significantly better than those who did not receive them. As predicted, no difference across conditions was found for rule-following items.

Whereas Experiment 1 manipulated the nature of the stimuli (discrete vs. continuous), Experiments 2 and 3 held the stimulus set constant (all stimuli were continuous) across conditions and instead examined how training subjects on different encoding strategies affects performance. These experiments were an extension of Experiment 1, in that they tested whether either encoding pathway can be made available under suitable training conditions. Manipulating the relative task compatibility of the two pathways between subjects further allowed us to test some of the hypothesized constraints on the two pathways, such as differences in encoding speed and ease of transfer to novel stimuli. Experiment 2A focused on extended training with continuous stimulus sets under conditions emphasizing direct task experience (encouraging image-based encoding), or learning a vocabulary of discrete values on which to map the continuous stimuli. In both cases, subjects were able to learn to recognize and categorize the exception items. Further, we provided evidence that subjects using a learned vocabulary for part-based encoding require more time and resources to be successful (Experiment 2B). In Experiment 3, we considered how subjects who received category learning training encouraging the use of image- or part-based encoding performed in a follow-up category learning task involving new categories constructed from the same stimulus set. Figure 1 provides a comparison of the training and test regimens for Experiments 1–3.

After we report the behavioral results for each set of experiments, we use a nested modeling procedure to extract additional information from subjects' recognition results (discussed formally in the next section). This modeling


**Figure 1. The various phases of each experiment are shown. In Experiment 1, subjects completed a category learning task, followed by recognition testing, with either discrete or continuous-valued stimuli. In Experiments 2 and 3, all stimuli were continuous-valued. In Experiments 2 and 3, subjects were given training prior to the final round of category learning training and recognition testing. In particular, subjects were either taught to assign labels to the features of the stimuli (i.e., identification learning [ID]), or completed additional category learning training. In the follow-up to Experiment 2, subjects from the ID condition were trained on a different category rule, and were given additional time on each trial to process category learning feedback. Experiment 3 combines the elements of Experiment 2 and Experiment 2's follow-up (i.e., additional feedback time and evaluation on two different category rules).**

approach helps to extend the basic statistical analysis by providing more rigorous comparisons of the encoding strategies that subjects use. Although our predictions and results are largely interpreted within a two-pathway framework, we will evaluate current single-pathway models in the General Discussion and consider how these proposals can be extended to account for our findings.

## Development of a Formal Model

To evaluate the results from our experiments, we fit a hierarchy of nested models inspired by Nosofsky and Zaki's (2003) hybrid similarity (HS-GCM) extension of the generalized context model (GCM; Nosofsky, 1986). The best-fitting parameter values from these models, along with the goodness of the overall fit, help us to evaluate whether one of the two proposed encoding pathways was engaged by subjects in a particular experimental condition. The HS-GCM embodies all of the assumptions of the GCM, but also incorporates a feature matching process inspired by Tversky's (1977) feature contrast model. This feature matching process added to the HS-GCM enables it to account for exception advantages in recognition memory that the GCM cannot address (e.g., Palmeri & Nosofsky, 1995; Sakamoto & Love, 2004). The feature matching processes embodied in the HS-GCM can be used to determine whether subjects are making use of one of the two encoding pathways. Our key behavioral measure is recognition memory following category learning; thus, we focus on how the GCM and HS-GCM account for the recognition data.

Recognition memory in the GCM is based on an item's familiarity, $F$, defined as the item's summed similarity to all exemplars stored in memory. In the case of a two-alternative forced choice (2AFC) recognition memory task in which item $i$ is old, and item $k$ is a new foil, the probability of correctly choosing item $i$ is given by

$$P(\text{correct} \mid ik) = F_i/(F_i + F_k), \tag{1}$$

and the similarity of an item $i$ to stored exemplar $j$ is given by

$$s_{ij} = \exp(-k * d_{ij}), \tag{2}$$

where $k$ is a scaling constant and $d_{ij}$ is the distance between probe item $i$ and stored exemplar $j$. For all applications described herein, distance will be given by the weighted city-block metric:

$$d_{ij} = \sum w_m * |x_{im} - x_{jm}|, \tag{3}$$

where $x_{im}$ is the probe item's value on dimension $m$, $x_{jm}$ is the value of exemplar $j$ on dimension $m$, and $w_m$ gives the attention weight on dimension $m$. These attention weights are constrained to sum to 1, and thus in the present applications, in which all of the stimuli have two dimensions, there will be a single attention weight parameter, Wr, which represents the attention to whichever dimension is rule-relevant.

Nosofsky and Zaki (2003) developed the HS-GCM because the GCM could not account for the fact that items possessing discrete distinctive features are recognized by people at a higher rate than more similar items are. The GCM predicts better recognition for items having a high rather than a low familiarity. The HS-GCM solves this problem by allowing items that possess distinctive features to become more self-similar and thus recognized at a higher rate. This is accomplished by calculating familiarity as the sum of hybrid similarities given by

$$sh_{ij} = C * D * s_{ij}, \tag{4}$$

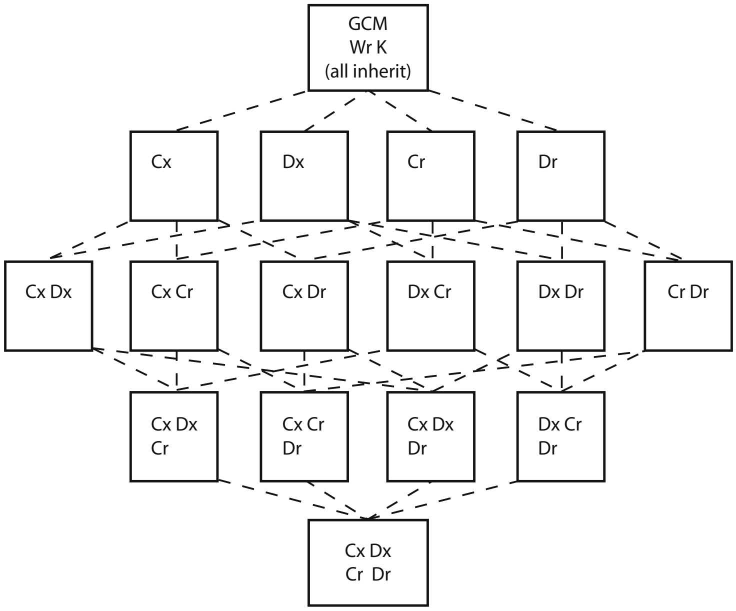


**Figure 2. Hierarchy of nested models used for the model-based analysis of the experiments. At the top of the hierarchy is the GCM, which includes two free parameters, an attention weight, Wr, and the scaling parameter K. Each level in the hierarchy adds a single hybrid parameter (see text for parameter descriptions). Analysis involves fitting each of these models and using nested model testing to select the best model while penalizing for complexity.**

where $C$ ($C > 1$) is the increase in similarity a probe item gets for matching the distinctive features of an exemplar in memory, $D$ ($0 < D < 1$) is the reduction in similarity from mismatching these distinctive features, and $s_{ij}$ is the spatial similarity metric used in the GCM (Equation 2, above).

We use a hierarchy of models inspired by the HS-GCM as a measurement tool for assessing whether subjects were successful at encoding items, and to determine how much weight subjects placed on matching the distinctive features of certain types of items (exceptions vs. rule-following items) in their recognition memory responses. Our approach is to fit a hierarchy of nested models (see Figure 2), starting with a two-parameter version of the GCM, and building to a full model that includes six parameters. By definition, models with more parameters will fit at least as well as simpler, nested models. Using model selection statistics (general linear test), we evaluate members of the model family in a manner than both rewards improved fit and penalizes increased complexity in order to arrive at the most psychologically viable model.

In all experiments, there will be two primary item types: exceptions and frequency-matched rule-following (control) stimuli. Because the framework makes predictions regarding when items will be encoded and which items will be encoded, the model needs to be expanded to include C and D parameters specific to matching the distinctive features of either the frequency-matched rule-following or exception items. For example, when a probe item matches the distinctive features of an exception stored in memory, it receives a boost in similarity of Cx, and when a probe item mismatches the distinctive features of an exception in memory, its similarity is reduced by Dx. Likewise, Cr and Dr parameters mark whether a probe item matches or mismatches the features of a frequency-matched rule-following item stored in memory. These additional parameters allow the model to account for exception advantages in recognition memory when they are present in the data. We emphasize that even though these parameters are specific to item types, we follow Nosofsky and Zaki (2003) in interpreting them as weighting the presence/absence of feature types. Whereas in the present study items only share distinctive features with themselves, it would be possible to test this interpretation of the parameters, as Nosofsky and Zaki did, by including different foil types.

Our nested model hierarchy allows us to quantitatively test a variety of the predictions outlined above. First, since only exception items should require individuated encoding via one of the two pathways, the only parameters that should prove to be useful in accounting for recognition data in the following experiments are the ones specific to the exceptions; that is, since the rule-following items do not require individuated encoding, including Cr and Dr parameters should not produce significantly better fits than those from nested models that assume Cr and Dr are fixed at 1. Further, of the two exception parameters (Cx and Dx), the Cx parameter should contribute most to the fits for conditions in which subjects successfully individuate the exception items, because Cx signifies that subjects are using a direct match between the features of a probe and a stored exception item in their recognition memory decisions.

Second, early in learning, subjects should only be able to use the distinctive features of the exception items if they have a discrete vocabulary with which to translate them; we addressed this question in Experiment 1. Fits of the nested model hierarchy to Experiment 1 should yield higher values of Cx for groups of subjects who receive stimuli with additional discrete labels in comparison to groups of subjects who receive the continuous stimuli without these aids. This follows from the framework, because neither pathway should be able to represent continuous-valued stimuli when they are completely novel. Further, in the conditions that do not receive the additional discrete vocabulary to bolster part-based encoding, the added Cx parameter should not produce significantly better fits in comparison with the GCM. In contrast, when these discrete labels are provided, the model with the added Cx parameter should fit significantly better than the GCM.

Finally, in Experiments 2 and 3, we examined the prediction that subjects should be able to learn to encode the distinctive features of the exception items through either of the two pathways, if trained appropriately and if the conditions for the proper functioning of the pathway are met. When a person uses either pathway to encode the exception items, the Cx parameter should contribute significantly to model fits.

## EXPERIMENT 1

Experiment 1 examined whether the ability to use part-based encoding is important for successful recognition of exception items following brief rule-plus-exception category training. Subjects in the continuous condition received rule-plus-exception training with the continuous-valued stimuli in Figure 3, followed by a recognition test. Subjects in the discrete condition followed the same procedure with identical stimuli, except that the stimuli also contained redundant verbal cues, as displayed in Figure 3, that were intended to aid part-based encoding.

The verbal cues present in the discrete condition were completely redundant, in that all stimulus information could, in principle, be encoded in the absence of these labels. However, if being able to map the stimuli onto a symbolic vocabulary is important for recognizing and storing the exceptions individually, subjects in the discrete condition should perform better at categorizing and recognizing exception items. Because rule-following items do not require separate encoding, performance for these items should be comparable across conditions.

Following presentation of the statistical analysis of the results from Experiment 1, we present results from the model hierarchy developed above. These fits should support the statistical analyses of the experimental results. In particular, in conditions in which an exception advantage is observed, model fits should indicate that subjects successfully encoded the distinctive features of the exception items. In terms of the model hierarchy (see Figure 2), models including the Cx parameter should fare better than the GCM model. Also, because we predict that subjects are not individuating rule-following items, we predict that including the Cr and Dr parameters should not significantly improve model fit.

### Method

**Subjects**. One hundred three students from the University of Texas at Austin participated for course credit. Each subject was randomly assigned to either the continuous or discrete condition.

**Materials**. Example stimuli for the discrete and continuous conditions are shown in Figure 3. Stimuli were rectangles in which there were variations in height and in the position of a vertical line segment along the lower base of the figure. Each dimension displayed six unique values, yielding 36 total stimuli. Neighboring values

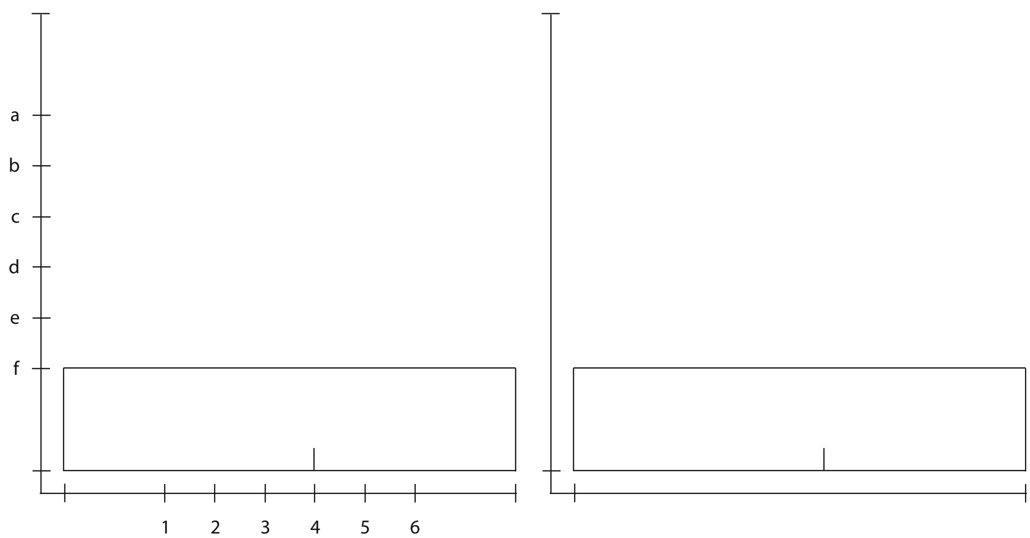


**Figure 3. An example of a discrete (left) and a continuous (right) stimulus from Experiment 1. The height of each stimulus varies, as does the position of the line segment along the lower boundary of each stimulus.**

along a dimension were separated by 14 mm (approximately 1.1º of visual arc).

Each stimulus dimension was bounded by a fixed-length axis. In the discrete condition, these axes contained labels and tick marks at each of the six dimension values. This labeling scheme is illustrated in Figure 3. Axes in the continuous condition did not have these markings.

The primary figure was blue, the axes were purple (including tick marks in the discrete condition), all text was black (including labels in the discrete condition), and the screen background was white.

**Design**. Subjects were trained in a category learning task using a rule-plus-exception structure, in which they had to place items into one of two contrasting categories. This was followed by a 2AFC recognition phase.

The category structure (see Figure 4) was designed so that subjects could successfully categorize most stimuli using an imperfect rule on one of the dimensions of variation. The dimension that the imperfect rule corresponded to was counterbalanced between subjects. The structure used when height was the rule-relevant dimension was equivalent to a 90º rotation of that used when the line-segment dimension was rule-relevant (shown in Figure 4). Each category also included one exception item that did not follow this rule. An example category structure is shown in Figure 4. In this figure, the rule-relevant dimension is the line-segment position. Each letter in the 6 × 6 grid stands for the location of one stimulus. All of the lowercase letters stand for regular rule-following items, the cursive letters stand for exceptions, and the capital letters stand for rule-following items that matched the exceptions in terms of frequency of presentation. The ?s were spaces reserved for foils to be used in the recognition phase, and were not shown in the category learning phase of the experiment. The placement of these special items (exceptions, foils, and frequency-matched rule-following items) was randomized between subjects, such that the exception items could appear in one of the four locations given by the circles in Figure 4. Once the exception placement was determined, the placement of the other items was determined by the following rule: frequency-matched rule-following items were always in the position diagonal to the exception, and the remaining two positions were reserved for the foils.

The category learning phase consisted of three training blocks. In every training block, each regular item (lowercase in Figure 4) appeared once, and each exception and frequency-matched rule-following item appeared three times. Foil items (the ?s in Figure 4) did not appear during the category learning phase. Thus, each subject completed 120 category learning trials. The trial order was randomized for each block, for each subject.

Following training, subjects completed the recognition phase. On each recognition trial, subjects judged which of two presented items appeared during the category learning phase. One item of each presented pair was one of the four foils not shown during category learning. The other item was either one of the exceptions or a frequency-matched rule-following item. Subjects completed two blocks of recognition. Each block consisted of all possible pairings described above. Thus, each subject completed 32 recognition trials. The trial order was randomized for each block for each subject.

**Procedure**. Directions were displayed on the screen prior to both the category learning and recognition phases. Subjects wore headphones to deliver auditory feedback and to dampen background noise.

On each category learning trial, the stimulus was presented in the central area of the screen, along with the prompt "Category A or B." Subjects responded by pressing the A or B key. After responding, the stimulus remained on the display. A low- or high-pitched tone sounded, depending on whether the subject's response was correct or incorrect. Additionally, the text "Correct" or "Incorrect" and "The answer is Category A" or "The answer is Category B" was displayed for 2,000 msec after the response. After this feedback period, the screen was blanked for 500 msec prior to the start of the next trial.

On each recognition trial, two stimuli were presented side by side in the central area of the screen. The locations of the stimuli were offset from one another by 35–40 mm on the horizontal dimension and 10–15 mm on the vertical dimension. The offset for each trial was determined randomly by sampling from a uniform distribution. Along with these two stimuli, the text prompt "Old: Left (q) or Right (p)?" was displayed at the center of the screen, and subjects responded by pressing the respective key. After responding, the screen was blanked and the text "Thank you" was displayed for 2,000 msec. This was followed by a blank screen for 500 msec prior to the start of the next trial.

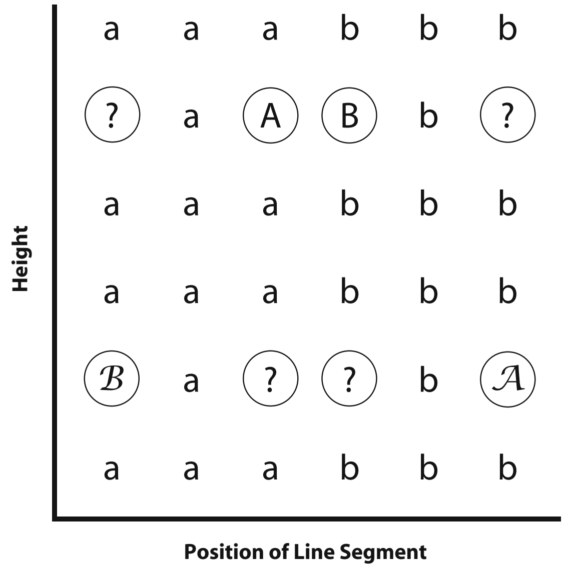


**Figure 4. A sample category structure, with exceptions in cursive and frequency-matched rule items in capital letters. The ?s represent foils used in the recognition phase. The center of the figure, which separates the two categories, represents the decision bound between them.**

## Results

Data were analyzed from the final block of category learning training and both blocks of the recognition phase. The first two category learning blocks are not presented because all subjects make a substantial number of errors early in training and our focus is on their ability to master the category structures under various conditions.

**Category learning**. Figure 5 displays categorization accuracy for the final block of training. A factorial condition (discrete or continuous) × item type (exception or frequency-matched rule item) ANOVA was conducted on category learning accuracy. A significant condition × item type interaction was observed [$F(1,101) = 22.07$, $MS_e = .055$, $p < .001$, $\eta_p^2 = .179$]. Exception item accuracy was significantly higher in the discrete condition (.67) than in the continuous condition (.37) [$t(101) = 5.55$, $p < .001$], whereas accuracy for the frequency-matched rule-following items did not differ (discrete = .88; continuous = .88) across conditions ($t < 1$).

**Recognition**. Figure 6 displays recognition accuracy. A factorial condition (discrete or continuous) × item type (exception or frequency-matched rule item) ANOVA was conducted on recognition accuracy. The interaction was

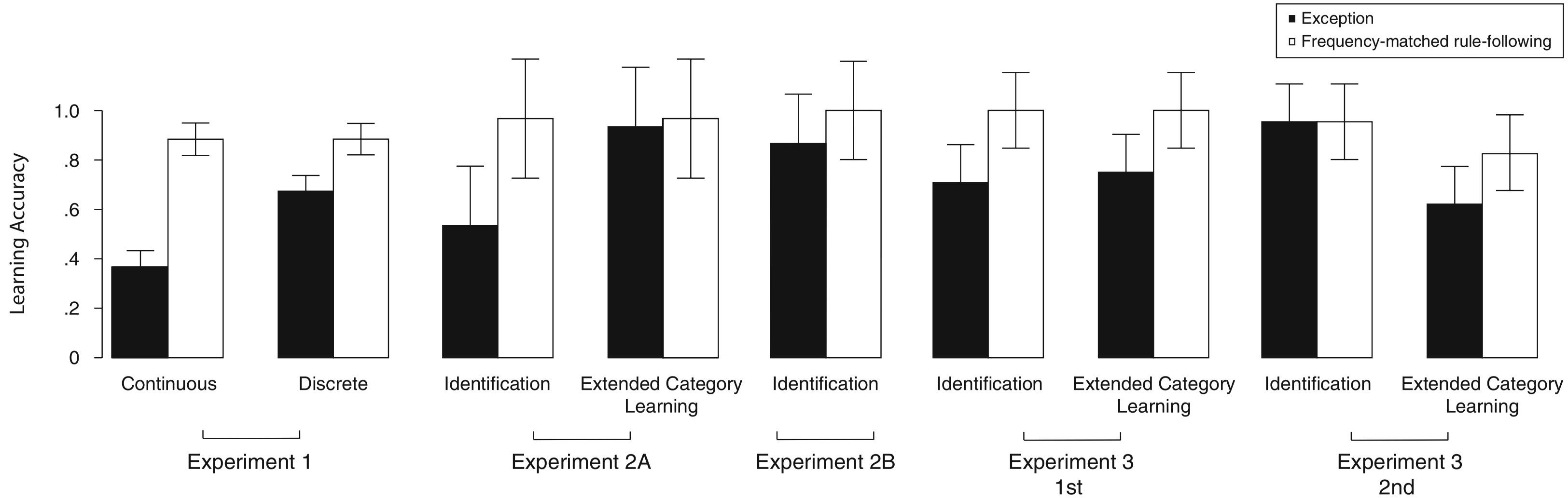


**Figure 5. Category learning performance for all experiments. Error bars are 95% within-subjects confidence intervals computed within each individual experiment (Loftus & Masson, 1994).**

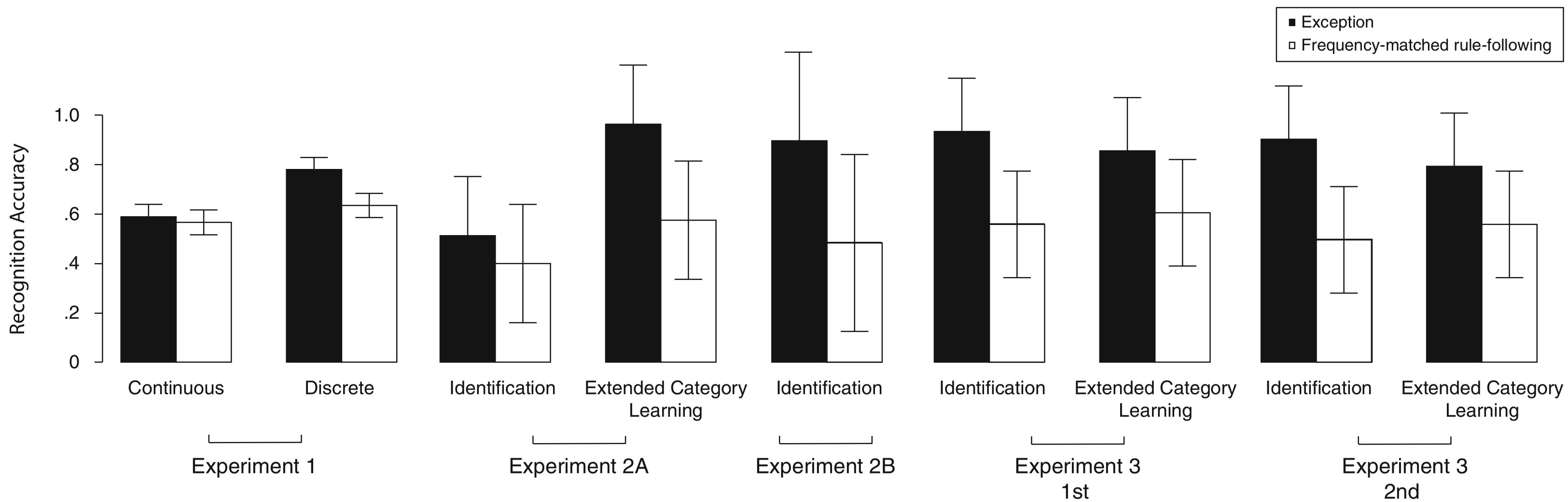


**Figure 6. Recognition performance for all experiments. Error bars are 95% within-subjects confidence intervals computed within each individual experiment (Loftus & Masson, 1994).**

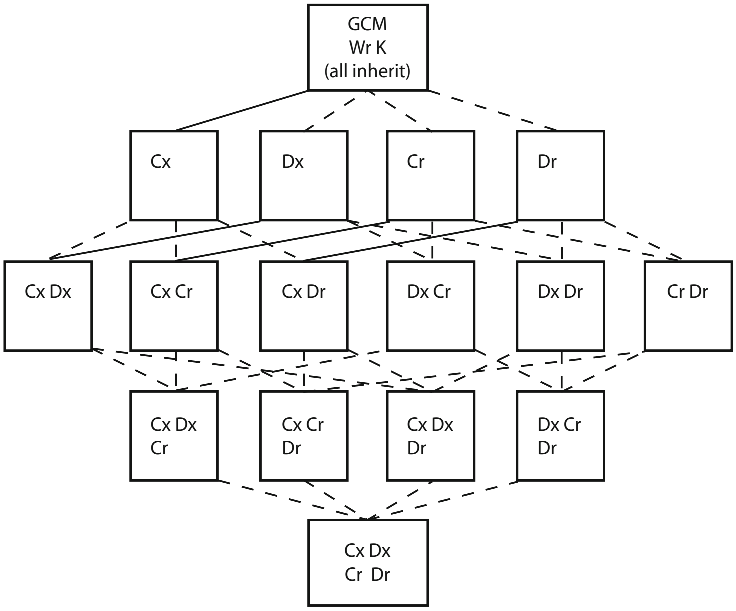


**Figure 7. Model hierarchy for Experiment 1 discrete condition. Each level of the hierarchy adds an additional free parameter from the possible free parameters discussed in the introduction. Solid connections show significant increases in variance accounted for between levels (general linear test).**

significant [$F(1,101) = 6.043$, $MS_e = .032$, $p = .016$, $\eta_p^2 = .056$]. Exception items (.78) were recognized at a significantly higher rate than were the frequency-matched rule-following items (.63) in the discrete condition [$t(52) = 4.09$, $p < .001$], but not in the continuous condition (.59 and .57, respectively) ($t < 1$). The difference in recognition performance between conditions was significant for the exception items [$t(101) = 4.51$, $p < .05$], and marginal for the frequency-matched rule-following items [$t(101) = 1.66$, $p = .10$].

## Model-Based Analysis

In this section we examine the results of Experiment 1, using the model hierarchy described above. By allowing us to rigorously test the behavioral profiles of subjects in each of the conditions in relation to the predictions of the two-pathway framework, this analysis goes beyond the results of the null hypothesis significance tests and point estimates. The primary goal of the model-based analysis is to determine how subjects encoded the different types of items, exceptions and rule following. Testing whether the hybrid parameters Cx and Dx (exceptions) and Cr and Dr (rule following) contributed to the model fits allows us to assess both whether the subjects encoded the respective items as individuals and the weight that they placed on these distinguishing features in recognition decisions.

**Fitting procedure**. In modeling the present data, we fit a hierarchy of nested models involving all possible combinations of hybrid parameters (Cx, Dx, Cr, Dr; see Figure 2) to the data in the discrete and continuous conditions. We used nested model testing procedures (general linear test) to select the best model (trading off complexity and goodness of fit). Each model was fit to eight data points representing the average probability that subjects within a condition[2] would correctly recognize an old item (frequency-matched rule-following or exception) when matched with a particular type of foil (see Figure 4). We used only pairings with unique similarity relations. These can be found by factorially crossing target type (frequency-matched rule vs. exception), foil position (matching exception on relevant vs. irrelevant dimension), and exception placement (whether the placement of the exception item was on the outside or inside of the category distribution; see Experiment 1 Method).[3] For example, four data points would be calculated from conditions with exception items placed on the inside of the category distribution (as in Figure 4). These would be exceptions matched with foils along the rule-relevant dimension, exceptions matched with foils along the irrelevant dimensions, and likewise for the frequency-matched rule-following items. The other four data points would be calculated using these same target–foil pairings from conditions in which the exception items were on the outside of the category distribution. We fit the models for all of the tasks by minimizing the sum of squared deviations between predicted and observed probabilities, and we tested the significance of added parameters using the general linear test.

**Modeling results**. The modeling results for the discrete condition are displayed in Figure 7. Table 2 shows the goodness of fit (SSD) for each model. Significant changes ($p < .05$) in variance accounted for are given by

**Table 2**
**Fits (SSD) of Each of the HS-GCM Models in the Hierarchy to the Recognition Data**

| | Experiment | | |
|---|---|---|---|
| | 1 (Discrete) | 1 (Continuous) | 2A, 2B, and 3 (Aggregate) |
| Cx Dx Cr Dr Wr K | .01455 | .08505 | .02063 |
| Cx Cr Dr Wr K | .01455 | .08772 | .02139 |
| Cx Dx Cr Wr K | .01504 | .08505 | .02067 |
| Cx Dx Dr Wr K | .01455 | .08505 | .02063 |
| Dx Cr Dr Wr K | .0547 | .08570 | .24827 |
| Cx Cr Wr K | .01504 | .08772 | .03317 |
| Cx Dx Wr K | .01504 | .08505 | .02067 |
| Cr Dr Wr K | .0547 | .08837 | .24827 |
| Cx Dr Wr K | .01455 | .08772 | .02442 |
| Cr Dx Wr K | .0555 | .08570 | .26450 |
| Dx Dr Wr K | .0547 | .08510 | .24827 |
| Cx Wr K | .01504 | .08772 | .03317 |
| Dx Wr K | .0555 | .08570 | .26450 |
| Cr Wr K | .0555 | .08837 | .27615 |
| Dr Wr K | .0547 | .08837 | .24827 |
| Wr K | .0555 | .08837 | .27615 |

Note—See Figure 2.

solid connections in the hierarchy, whereas nonsignificant changes are dashed. As is clear from Figure 7, in the discrete condition, the Cx parameter is the only parameter that contributes significantly to the model fits beyond the base GCM parameters (Wr and K). In the continuous condition, no model fit significantly better than the GCM. As predicted from the framework, being able to encode the exception item's distinctive features (reflected in the Cx parameter) via the provided tick marks and labels was critical for recognition (see Tables 3 and 4 for the predicted/observed probabilities from both sets of experiments, and the best-fitting parameter values, respectively). Finally, in addition to the results of the nested model testing, Cx had a smaller obtained value (1.15) in the continuous condition than did the Cx observed in the discrete condition (2.67). This higher value for the Cx parameter in the discrete condition indicates greater weight on the exception items' distinctive features in the discrete condition than in the continuous condition. Together, these two measures provide strong evidence that these conditions did differ in the extent to which subjects encoded and used information about the exception items.

**Table 3**
**Observed and Predicted (HS-GCM With Only Cx Parameter + GCM) Probabilities Correct by Item Type**

| | Observed | Predicted (HS-GCM) |
|---|---|---|
| Experiment 1 (Discrete) | | |
| Exception | .78 | .78 |
| Rule | .64 | .64 |
| Experiment 1 (Continuous) | | |
| Exception | .59 | .59 |
| Rule | .57 | .57 |
| Experiments 2 and 3 (Aggregate) | | |
| Exception | .90 | .90 |
| Rule | .55 | .55 |

**Table 4**
**Parameter and $R^2$ Values for the Best-Fitting Models**

| | Cx | K | Wr | $R^2$ |
|---|---|---|---|---|
| Experiment 1 (Discrete) | 2.67 | 17.00 | .98 | .74 |
| Experiment 1 (Continuous) | 1.15 | 16.11 | .99 | .65 |
| Experiments 2 and 3 (Aggregate) | 13.38 | 16.87 | .99 | .88 |

## Discussion

The predictions derived from the two-pathway framework were supported by the results of Experiment 1. Following brief training in the category learning task, only subjects in the discrete condition displayed robust exception recognition. For subjects in the discrete condition, the redundant verbal cues that enable part-based encoding were sufficient to bolster exception recognition. Categorization of the rule-following items did not require them to be encoded separately from the rest of the category, so performance was equivalent for these items in the discrete and continuous conditions. This interpretation is strengthened by the modeling results, which illustrate how subjects weighted the item features in their recognition decisions. In the discrete condition, subjects were able to encode the distinguishing features of the exception items, as evidenced by the model fits and the higher value of the Cx parameter. This was not the case in the continuous condition, where no model yielded a better fit than the GCM.

One key question is whether subjects in the two conditions represented the stimuli in fundamentally different ways. One signature of continuous representation that has been studied extensively in the absolute identification literature is *bow effects* (for a review, see Stewart et al., 2005). Bow effects occur when there is increased accuracy for identifying items at the extremes on a dimension of variation. Additional analysis of our data revealed bow effects for the continuous, not the discrete, condition. In particular, subjects in the continuous (but not discrete) condition showed better recognition for whichever item (frequency-matched rule-following or exception) was farther from the decision bound (i.e., on the extreme edge of the rule-relevant dimension; see Figure 4).[4]

That a bow effect did not occur in the discrete condition is evidence that encoding the exception via the provided verbal cues involved a change in representation from a continuous perceptual space to a symbolic one. The presence of bow effects in the continuous condition and their absence in the discrete condition bolsters the two-pathway framework.

A final important aspect of Experiment 1 is its relationship to Nosofsky and Johansen's (2000) critique of a similar experiment by Erickson and Kruschke (1998): When tick marks and labels are added to stimuli, subjects might completely ignore the physical stimulus parts and encode only the alphanumeric labels. Whereas Nosofsky and Johansen did not explicitly offer a dual-pathway model, their critique anticipates the need for a theory, such as the one we propose here, that incorporates both part- and image-based encoding. The possibility that some subjects were using this strategy is, therefore, consistent with the present approach, and would be indicative of a feature matching process that relies on the mapping of stimulus features onto a discrete vocabulary.

Whether part-based encoding is completely reliant on the presence of the added features used in Experiment 1 is addressed in the following experiments, in which these features are removed from the stimuli. To foreshadow, comparing both the behavioral and modeling results indicates even stronger weighting (higher Cx parameter values) on matching the exception items' distinctive features when the alphanumeric labels have been removed (Experiments 2 and 3) than when they are present (Experiment 1). This suggests that subjects were not relying only on these labels to make their decisions.

**Overview of Extended-Training Experiments**

In Experiment 1, the nature of the stimuli (continuous vs. discrete) was manipulated across conditions as tick marks and verbal labels were introduced in the discrete condition. We hypothesize that these additions facilitated the encoding of the exception items via the part-based pathway, as evidenced by these subjects' enhanced abilities to categorize and recognize exceptions. In Experiments 2 and 3, we examined encoding along both pathways with stimuli that lack these additions. One key question was whether subjects could master exception items in the absence of this added discrete information. Subjects in the continuous condition of Experiment 1 could not master the exceptions, but subjects in Experiments 2 and 3 might have, with training that encouraged successful encoding along either the image-based or part-based pathway.

Most objects in real-world categorization tasks lack the external labels present in Experiment 1's discrete condition, yet people can often individuate objects. One possibility is that people encode objects by either generating their own internal labels for part-based encoding or by developing the capacity to encode objects holistically and automatically through the image-based pathway.

In Experiments 2 and 3, we trained subjects on tasks that encourage the use of image- or part-based encoding prior to testing them on their ability to encode exception items. To anticipate the results: We found that subjects trained to use either type of encoding were able to encode the exception items, but that subjects showed critical differences, depending on their initial training conditions. In particular, when subjects were trained to use part-based encoding, they required additional time to encode exception items.

In terms of the model hierarchy, we predict that models including the Cx parameter (see Figure 2) should fit better than the GCM model in any of the conditions in which subjects are able to encode the exception items. We also predicted that including the Cr and Dr parameters should not significantly improve model fit, since individual encoding should not be necessary for learning the rule-following items. Finally, because these experiments involved extensive training on only a few subjects, the model is fit to aggregate data from Experiments 2 and 3 and will be presented after Experiment 3.

## EXPERIMENT 2A

Experiment 2A examined the ability of subjects to master exception stimuli (that do not include external cues) under conditions that encouraged either image- or part-based encoding. Prior to completing category learning training and recognition testing using the same methods as in Experiment 1's continuous condition, subjects completed an initial training regimen. In the extended category learning condition, the initial training regimen consisted of extended category learning training. This extensive on-task experience should encourage the development of fine-grained image-based representations for the exception items. In the identification condition, subjects also completed an extensive initial training regimen. Instead of category learning, these subjects engaged in an identification task that required them to learn to identify the stimuli by assigning labels to the individual stimulus parts (i.e., height or line segment position), a task that should encourage the formation of an internal vocabulary useful for encoding the stimulus parts.

The two-pathway account predicts that, following these initial training regimens, subjects in both conditions should be able to master exception items, albeit via different pathways. Such a result would contrast with the failure of subjects in Experiment 1's continuous condition to master the exception item in the absence of external labels.

### Method

**Subjects**. Ten students from the University of Texas at Austin participated in Experiment 2. They received monetary reimbursement of $7 per session for four 1-h sessions. In addition, they were offered a final day bonus of $23, depending on their performance; all subjects earned the bonus. Each subject was randomly assigned to either the extended category learning or identification condition.

**Materials**. Materials for the category learning and recognition phases were the same as those used in Experiment 1's continuous condition. The same stimuli were also used during the initial training regimen in the extended category learning condition. The initial training regimen for the identification condition also involved these same stimuli, but these stimuli were coupled with an additional grid, as shown in Figure 8. This grid was 75 × 75 mm, with 36 evenly spaced cells. As explained below, this grid was used during identification learning.

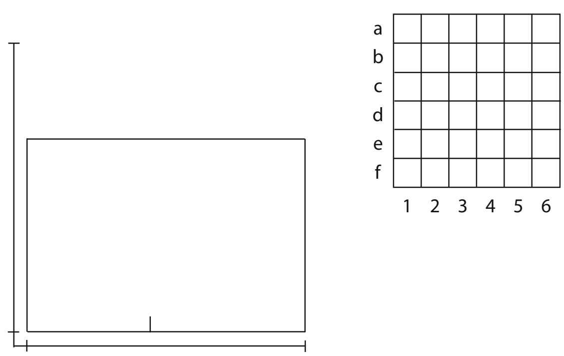


**Figure 8. An illustration of identification learning procedure. Subjects are queried first on the vertical dimension, at which point a black line goes through the grid at the location of their response. Next, they are queried on the horizontal dimension and a black dot appears with their final answer. If correct, the dot turns green during feedback, and if incorrect it turns red and the correct answer appears as a green dot.**

**Design**. Figure 1 provides an overview of Experiment 2's design. The final category learning and recognition phases in Experiment 2 were the same as those used in the continuous condition in Experiment 1. Where the conditions used in Experiment 2 differed from one another was in the initial training regimens. Subjects in the extended category learning condition received prolonged training in the original category learning task used in Experiment 1, whereas subjects in the identification condition received training of equal duration in an identification task. Both training regimens consisted of 14 blocks of training for each of the three training sessions. The definition of a block was the same as in Experiment 1 (i.e., one complete pass through the stimulus set in a random order).

In the fourth and final session, subjects in both conditions completed three additional blocks of their respective training regimens. This final training phase was followed by a final category learning and recognition phase identical for both conditions and identical to those used to assess performance in the continuous condition in Experiment 1.

**Procedure**. The procedures used for the category learning training and recognition phases were identical to those used in Experiment 1. The initial training regimen for the extended category learning condition also used Experiment 1's category learning procedure.

The procedure for the identification training required subjects to sequentially assign a value to both dimensions on each stimulus. On a given trial, a stimulus would appear on the screen along with the grid (see Figure 8). Subjects would be prompted to answer the question "What is the vertical position a–f?" and, immediately afterward, the question "What is the horizontal position 1–6?" They would answer each prompt by using the keys corresponding to those letters and numerals on the keypad. After the horizontal position was entered, a black dot would appear in the appropriate cell, and the prompt "This is your answer, hit 'n' to change, or hit any key to continue" appeared. If the subject chose to continue, the dot in the grid would change from black to red or green, for incorrect and correct answers, respectively. If the response was incorrect, a green dot would appear at the correct location. Auditory feedback was given using the same tones from Experiment 1, and subjects were informed, "Correct/incorrect. This is the correct answer. Hit 'enter' to continue." The subject's answer, the correct answer, and the stimulus remained on the screen during feedback. Everything was self-timed and each trial ended with a blank screen for 500 msec.

## Results

As in Experiment 1, data from the final block of category learning training and both blocks of the recognition phase were analyzed.[5] Additionally, data from the initial training regimen are briefly presented.

**Initial training regimen**. The purpose of the initial training regimen was to encourage part-based or image-based encoding, and thus is not the main focus of analysis. Even so, it is important to verify that subjects in both conditions achieved a high level of performance. As shown in Figure 9, subjects in both conditions approached asymptote by the end of the initial training regimen.

**Category learning**. Figure 5 displays categorization accuracy for the final block of training. A factorial condition (identification vs. extended category learning) × item (exception vs. frequency-matched rule) ANOVA was conducted on category learning accuracy. A significant condition × item type interaction was observed [$F(1,8) = 5.236$, $MS_e = .038$, $p < .05$, $\eta_p^2 = .396$]. Exception item accuracy was higher (.93) in the extended category learning condition than in the identification condition (.53) [$t(8) = 2.203$, $p = .08$],[6] whereas accuracy for the frequency-matched rule-following items did not differ across conditions (.97 for both; $t < 1$).

**Recognition**. Figure 6 displays recognition accuracy. A factorial condition (discrete or continuous) × item type (exception or frequency-matched rule item) ANOVA was conducted on recognition accuracy. The interaction was not significant [$F(1,8) = 1.763$, $MS_e = .054$, $p = .22$, $\eta_p^2 = .181$]. However, the main effect for item [$F(1,8) = 5.829$, $MS_e = .054$, $p < .05$, $\eta_p^2 = .421$] and the main effect for condition [$F(1,8) = 35.461$, $MS_e = .014$, $p < .05$, $\eta_p^2 = .816$] were both significant. Despite the failure to reach a significant interaction, exception items were recognized at a significantly higher rate (.96) than the frequency-matched rule-following items (.58) in the extended category learning [$t(4) = 3.56$, $p < .05$], but not in the identification condition (.51 and .40, respectively; $t < 1$). The difference in recognition performance between conditions was significant for the exception items [$t(8) = 4.97$, $p < .05$], but not for the frequency-matched rule-following items [$t(8) = 1.28$, $p = .24$].

## Discussion

In contrast to subjects' performance in the continuous condition in Experiment 1, subjects who experienced the same stimulus set in Experiment 2A's extended category learning condition were able to master exception items after extensive on-task (i.e., category learning) experience, which presumably led to the development of the fine-grained holistic representations that underlie image-based encoding. Contrary to the predictions we derived from the two-pathway framework, subjects in Experiment 2A's identification condition did not fully master the exceptions, despite excelling in an initial training regimen designed to enable part-based encoding.

Nevertheless, this failure to master exception items is not necessarily incompatible with the two-pathway framework. One possible explanation is that subjects in the identification condition did not have sufficient time during the category learning trials to properly encode the exception stimuli. As discussed in the beginning of this article, part-based encoding is hypothesized to involve analytic processes that require considerable time to complete. If this effortful encoding process was initiated after receiving negative feedback on exception items (full encoding is not required for rule-following items), the 2,000-msec feedback period might have been insufficient to process the feedback and map the exception items onto the trained vocabulary. Although these same subjects were successful at the identification task in the initial training regimen, the identification task was self-paced and the median time subjects took to respond to both dimensions in the final session was 2,424 msec.

## EXPERIMENT 2B

Subjects in Experiment 2A's identification condition, unlike those in Experiment 1's discrete condition, had to generate their own labels. One possibility is that 2,000 msec is not sufficient to process category learning

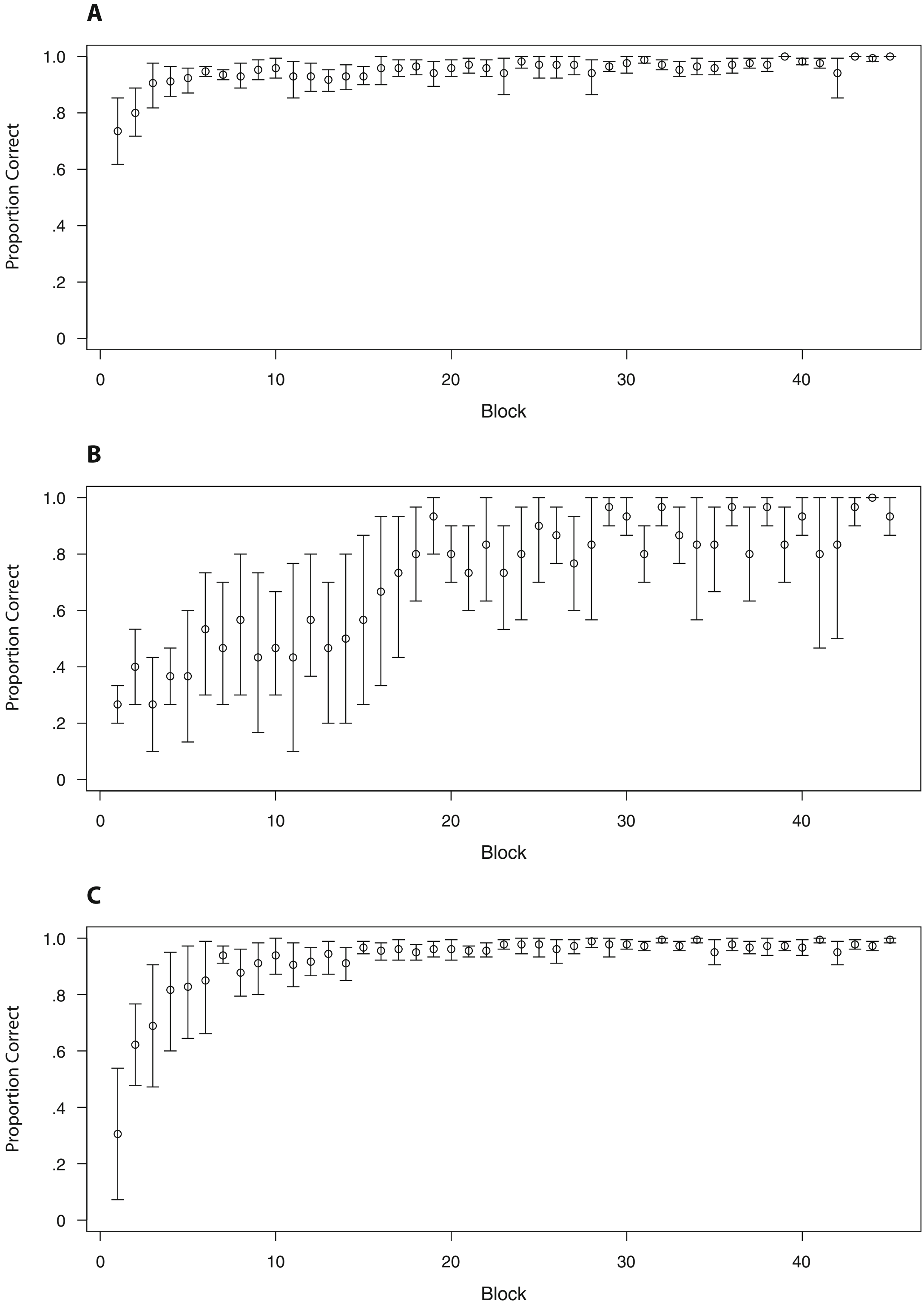


**Figure 9. Panels A and B show training performance by block for the extended category learning condition of Experiment 2A for rule-following items and exceptions, respectively. Panel C shows training performance by block for the identification condition. Error bars show bootstrapped 95% between-subjects confidence intervals for the mean performance at each block.**

feedback under these conditions. To test the hypothesis that subjects in the identification condition did not have enough time to perform part-based encoding, in Experiment 2B we had the subjects from the identification condition return for another testing session that was a slightly modified version of the fourth and final session of Experiment 2A. Specifically, these subjects completed category learning and recognition tasks identical to the fourth session, except that the rule-relevant dimension and exception placement were switched for each subject, and the feedback period of category learning trials was extended from 2,000 to 4,000 msec. If insufficient encoding time at feedback was the reason for the inability of subjects in the identification condition to master the exception items, extending this time should result in improved performance.

### Method

**Subjects**. The 5 subjects from the identification condition in Experiment 2A participated in Experiment 2B. They were given monetary reimbursement of $7 for a single 1-h session. In addition, they were offered a bonus of $8 depending on their performance; all subjects earned the bonus.

**Materials**. The materials were identical to those used in Experiment 2.

**Design**. (Consult Figure 1 for an overview of the experiment.) Subjects completed six blocks of identification training to serve as a refresher. This was followed by a testing phase consisting of a category learning and recognition task identical in format to those completed in Experiment 2A, except that the category structures were adjusted to minimize the impact of the previous testing phase. The category structures for this testing phase were constructed for each subject using the rule-relevant dimension and exception placement opposite from the one they had received in Experiment 2A. Since the category structure used when one dimension was rule relevant was a 90º rotation of that used when the other dimension was rule relevant (see Experiment 1 design), all of the items that were exceptions, foils, or frequency-matched rule-following items in Experiment 2A were ensured to be regular rule-following items in Experiment 2B.

**Procedure**. The procedure for all phases was identical to those in Experiment 2A, except that in the category learning test phase all stimuli remained on the screen with feedback for 4,000 msec instead of 2,000 msec. All other timing was identical to that in Experiment 2A.

### Results

As in Experiments 1 and 2A, data from the final block of both category learning phases and both blocks of the recognition phase were analyzed. Since the subjects for Experiment 2B were repeat subjects from Experiment 2A, most of this analysis involved comparing performance across sessions. However, any conclusions as to the ability of subjects in Experiment 2B to encode the exception items can be made without relying on these comparisons. Data from the identification warm-up blocks were omitted, since their only purpose was to serve as a refresher for the previous training.

**Category learning**. Figure 5 displays categorization accuracy for the final block of the category learning phase. A factorial item type (exception vs. frequency-matched rule) × session (Experiment 2A vs. 2B) ANOVA was conducted on the accuracy for the final block of the category learning phase. The condition × item type × session interaction was marginally significant [$F(1,4) = 4.3784$, $MS_e = .026$, $p = .105$, $\eta_p^2 = .52$]. Exception item accuracy was numerically higher in the second session (.87; Experiment 2B) relative to the first session (.53; Experiment 2A), but this was not significant [$t(4) = 1.9069$, $p = .129$]. Accuracy for the frequency-matched rule-following items did not differ across sessions (.97 and 1 for the first and second sessions, respectively; $t < 1$).

**Recognition**. Figure 6 displays accuracy for the recognition phase. A factorial condition item type (exception vs. frequency-matched rule item) × session (Experiment 2A vs. 2B) ANOVA was conducted on the accuracy for the recognition phase. The interaction was not significant [$F(1,4) = 1.360$, $MS_e = .0198$, $p = .308$, $\eta_p^2 = .25$]; however, there were significant main effects for session [$F(1,4) = 14.227$, $MS_e = .03057$, $p < .05$, $\eta_p^2 = .78$] and type [$F(1,4) = 11.272$, $MS_e = .03057$, $p < .05$, $\eta_p^2 = .74$]. Despite the failure to reach a significant interaction, exception items were recognized at a significantly higher rate (.90) than the frequency-matched rule-following items (.49) in the second session (Experiment 2B) [$t(4) = 3.56$, $p < .05$], but not in the first session (Experiment 2A) (.51 and .40, respectively; $t < 1$). The difference in recognition performance between sessions was significant for the exception items [$t(4) = 5.17$, $p < .05$], but not for the frequency-matched rule-following items ($t < 1$).

### Discussion

Whereas subjects in Experiment 2A's identification condition were unable to master the exception items when originally tested, these same subjects were able to do so with the feedback time increased from 2,000 to 4,000 msec in Experiment 2B. This result follows from the present framework, since part-based encoding is posited to involve more analytic and effortful processes, and subjects in the identification condition were trained in a manner designed to encourage this type of encoding.

## EXPERIMENT 3

In Experiment 2A, the extended category learning subjects successfully categorized and recognized the exception items. However, the identification subjects showed comparatively impaired performance with the same exception items. In Experiment 2B, the same subjects from the identification condition who were unable to master the exception items were able to learn to categorize and recognize these items when they were provided with additional time to process the feedback. This suggests that encoding a stimulus via part-based encoding is more effortful than doing so through image-based encoding.

As a more thorough test of the hypothesis that part-based encoding can lead to successful recognition and categorization of exception items if the feedback duration is sufficient, Experiment 3 was conducted with a new set of subjects. For Experiment 3, subjects in two conditions, extended category learning and identification, were trained using the same procedure as in Experiment 2A (see Figure 1 for an overview of the procedure). As in Experiment 2B, during the final session's category learning test, the feedback duration was increased from 2,000 to 4,000 msec. In addition, all subjects completed a second series of category learning and recognition tests using a category structure with the rule-relevant dimension and exception placement opposite those used in the previous category learning phase. These added category learning and recognition phases were included to shed light on the flexibility of both pathways in their ability to generalize to new contexts. On the basis of results from Experiment 2B, and the notion that the representations used by the part-based pathway are compositional, it was expected that subjects in the identification condition would master the exceptions in both tasks. It was less certain how subjects in the extended category learning condition would perform, since results from the expertise literature suggest that some of the new task should assimilate to their prior expertise (Tanaka, Curran, & Sheinberg, 2005; Tarr & Gauthier, 1998) but, at the same time, they would be expected to exhibit some interference from this prior learning.

### Method

**Subjects**. Eight students from the University of Texas at Austin participated in Experiment 3. They received monetary reimbursement of $7 per session for four 1-h sessions. In addition, they were offered a final day bonus of $23 depending on their performance; all subjects earned the bonus. Each subject was randomly assigned to either the extended category learning or identification condition.

**Materials**. The materials were identical to those used in Experiment 2.

**Design**. The experimental overview is presented in Figure 1. Both conditions were trained in a manner identical to those in Experiment 2 up until the final test on the fourth day. Immediately after the recognition phase following the first category learning test phase, subjects completed an additional category learning and recognition phase. The category structures for these second testing phases were constructed for each subject using the same procedure as outlined in Experiment 2B. This ensured that all of the items that were exceptions, foils, or frequency-matched rule-following items in the first phase were regular rule-following items in the second phase.

**Procedure**. The procedures for all phases were identical to those in Experiment 2A, except that in both category learning test phases on the final day, all stimuli remained on the screen with feedback for 4,000 msec instead of the 2,000 msec used in the previous two experiments. All other timing during the training and recognition phases was identical.

### Results

As in Experiments 1 and 2, data from the final block of both category learning phases and both blocks of each of the recognition phases were analyzed. Results from the initial training regimens were very similar to those from Experiment 2A.

**Category learning**. Figure 5 displays categorization accuracy for the final block of training for both category learning phases. A factorial condition (identification vs. extended category learning) × item (exception vs. frequency-matched rule) × phase (1st vs. 2nd) ANOVA was conducted on the accuracy for the final block of the category learning phase. The highest order interaction was not significant [$F(1,6) = 2.000$, $MS_e = .015625$, $p = .2070$, $\eta_p^2 = .25$]. Furthermore, there were no significant lower order interactions or main effects. In the first category learning phase, there was no difference between conditions in exception item accuracy (extended category learning = .75; identification = .71) ($t < 1$), nor did accuracy for the frequency-matched rule-following items differ (1 for both extended category learning and identification) across conditions ($t < 1$). In the second category learning phase, exception item accuracy was numerically lower in the extended category learning condition (.63) than in the identification condition (.96), but this was not significant [$t(6) = 1.372$, $p = .22$]. Accuracy for the frequency-matched rule-following items did not differ (.83 and .96 for extended category learning and identification, respectively) across conditions ($t = 1$).

**Recognition**. Figure 6 displays recognition accuracy for both recognition phases. A factorial condition (identification vs. extended category learning) × item (exception vs. frequency-matched rule) × phase (1st vs. 2nd) ANOVA was conducted on accuracy for the recognition phases. The highest order interaction was not significant [$F(1,6) = 0.0356$, $MS_e = .030884$, $p = .8566$, $\eta_p^2 = .0007$]. Only the main effect for type was significant [$F(1,6) = 16.5264$, $MS_e = .04846$, $p < .05$, $\eta_p^2 = .734$]. In the first recognition phase, exception items were recognized at a numerically higher rate than was the frequency-matched rule-following items in the extended category learning condition (.86 and .61 for exceptions and rule-following, respectively), but this was not significant [$t(3) = 1.41$, $p = .25$]. This effect was marginally significant in the identification condition (.94 and .56 for exceptions and rule-following, respectively) [$t(3) = 2.48$, $p = .09$]. In the first recognition phase, there were no differences between conditions in exception or frequency-matched rule-following performance ($t$s $< 1$). In the second recognition phase, exception items were recognized at a significantly higher rate (.80) than the frequency-matched rule-following items (.56) in the extended category learning condition [$t(3) = 3.64$, $p < .05$]. This effect was marginally significant in the identification condition (.91 and .50 for exceptions and rule-following, respectively) [$t(3) = 2.79$, $p = .07$]. As with the first session, there were no performance differences between conditions for either the exception or frequency-matched rule-following items ($t$s $< 1$).

### Discussion

In Experiment 3, subjects in both the identification and extended category learning conditions learned to categorize and recognize the exception items in each of the testing phases. This is in contrast to Experiment 2A, in which subjects in the identification condition were unable to master the exception items, but is supportive of the findings from Experiment 2B, which showed that the same subjects could successfully encode the exception items if given additional time during feedback.

Taken with results from previous experiments, these results converge with predictions derived from the present two-pathway framework. Both pathways, part- and image-based encoding, allow subjects to master the exception items but differ in how they make it possible for this to happen. Both are time consuming and effortful, but image-based encoding is more so. Furthermore, as predicted by the compositionality of part-based representations, this pathway supports performance in new contexts—for example, when the exception items and rule-relevant dimension are changed. The fact that the extended category learning condition also elicited good performance in the second testing phase is consistent with evidence suggesting that experts are able to transfer to new subcategories of objects similar to those for which their image-based representations have developed. Indeed, the stimulus set in the second category learning phase used the same stimuli, and only the way that they were partitioned was changed (i.e., which dimension was relevant, which items were exceptions, etc.).

## Model-Based Analysis for Extended-Training Experiments

In the modeling for the second set of experiments, we focused on the aggregate data from all conditions in Experiments 2A, 2B, and 3, except for the identification condition in Experiment 2A (in which the subjects were unable to encode the exception items). This includes the extended category learning conditions in Experiments 2A and 3 and the identification conditions in 2B and 3. We used the aggregate data for modeling because the number of subjects in each of these individual conditions and experiments was low. After we present the primary modeling results for these experiments, we show that the model selected on the basis of fits to the aggregate data account for nearly as much variance as did the fits to the individual conditions.

**Fitting procedure**. The model fitting procedure was identical to that described for Experiment 1.

**Modeling results**. The results of the nested model testing are summarized in Figure 10 and the goodness-of-fit values are given in Table 2. The results are similar to those from Experiment 1 and are clear. In every case that a Cx parameter was included, it provided a significant improvement in fit over the nested model for which Cx = 1; this pattern did not hold for any other parameter. An important difference between experiments was that the value Cx was much higher (13.38) in Experiments 2 and 3 than in Experiment 1 (2.67), suggesting that subjects in Experiments 2 and 3 relied more heavily on distinctive feature matches for exception items; this may be a sign of expertise in the task (see Tables 3 and 4 for the predicted/observed probabilities from both experiments and the best-fitting parameter values, respectively).

Finally, even though we did not focus on the individual conditions, in order to show that they were fairly homogeneous we calculated the variance accounted for in each type of condition when they were fit individually, and compared this with the variance accounted for in each of these conditions when the parameters were fixed at the values obtained from the aggregate data. In both cases, the parameter values obtained from fits to the aggregate data accounted for nearly as much variance as did fitting each type of condition individually. In the identification conditions, the parameter values from the Cx + GCM model fit to the aggregate data accounted for 83% of the variance, and the fit to the identification subjects alone fared only slightly better, accounting for 85% of the variance. The

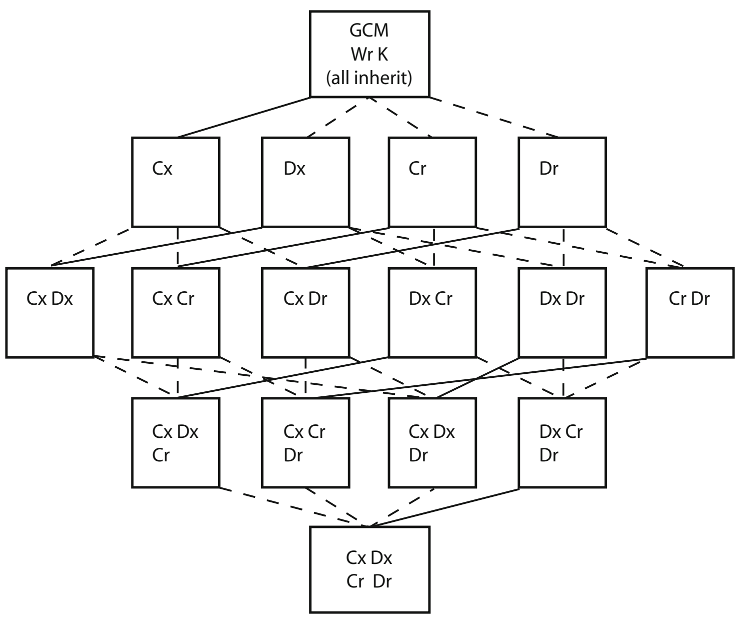


**Figure 10. Model hierarchy for Experiments 2 and 3. Each level of the hierarchy adds an additional free parameter from the possible free parameters discussed in the introduction. Solid connections show significant increases in variance accounted for between levels (general linear test).**

extended category learning condition was much the same, with 78% of the variance accounted for using the parameters from the aggregate model, and only a slight increase to 81% when this condition was fit individually.

## GENERAL DISCUSSION

The aim of the present investigation was to examine object-encoding processes in category learning. Models of category learning do not often focus on how stimuli are encoded, and thus tacitly suggest that there is a single encoding mechanism. In contrast, we developed a novel framework to explore the possibility that there are at least two pathways for stimulus encoding in categorization. The first of these pathways, which has received a large amount of attention in the recent expertise/object recognition literature, is through the development of holistic image-based representations. The second pathway, motivated by the part-based representations that are also common in the object recognition literature, involves mapping objects onto discrete symbolic representations.

We used this framework to derive predictions for a series of experiments involving rule-plus-exception designs. Rule-plus-exception designs are well suited for testing assumptions about stimulus encoding because accurate responding requires separate encoding and storage of exception stimuli. We found evidence to suggest that subjects required the ability to individuate items via one of these pathways in order to learn the exception items. This was not the case for individual rule-following items. Indeed, comparing recognition results across experiments (see Figure 6) reveals that none of the manipulations affected performance for the frequency-matched rule-following items, whether these items were presented 9 times prior to the recognition memory test, as in Experiment 1, or 144 times, as in Experiments 2A and 3. Given the consistently high category learning accuracy and low recognition performance across conditions with the frequency-matched rule-following items, we suspect that all conditions in all experiments used similar encoding strategies for these items, regardless of training, but that neither of the conditions consistently used one of the two pathways to encode and store these items as individuals.

In support of the two-path framework, we found in Experiment 1 that subjects could successfully individuate exception items following brief training when stimuli contained tick marks and verbal labels that facilitated part-based encoding. Subjects were not successful when stimuli were purely continuous (i.e., lacking tick marks and verbal labels). This is a potentially important difference between stimuli with continuous and discrete valued dimensions. Discrete valued stimuli are easily mapped onto symbolic representations (i.e., encoded via the part-based pathway), and this facilitates categorization and recognition of these stimuli early in learning. Continuous stimuli do not share this property, so they are not easily mastered in a limited number of learning trials. This result has implications for the category learning literature, because these types of stimuli are often used interchangeably.

In Experiments 2 and 3, we found that subjects trained in a manner stressing either pathway were able to learn to categorize and recognize exceptions in the absence of external discrete cues, such as tick marks and verbal labels. These different training regimens revealed performance profiles that follow from the two-pathway framework. Subjects who were trained in a manner that encouraged the development of image-based representations required less time to encode exception items than did those trained to use the more analytically based and attention demanding part-based pathway. Subjects trained via either pathway were also able to generalize to novel partitionings of the stimulus set. One might expect that the part-based pathway would yield greater generalization; although there was a numerical advantage for conditions that promoted part-based encoding, this advantage did not reach significance, perhaps because of limited power, or because image-based encoding can support generalization within a stimulus set.

Along with the transfer findings from Experiment 3, findings from the expertise literature suggesting that holistic image-based representations can be flexibly extended to novel subcategories reinforce the notion that the representations used by the part-based pathway are compositional (Tanaka et al., 2005; Tarr & Gauthier, 1998). At the same time, the flexibility of subjects in the extended category learning condition in transferring to novel subcategories leaves open the possibility that these subjects had also developed a part-based vocabulary in the course of their training, perhaps via feature learning processes discussed by Schyns et al. (1998). Below, we discuss manipulations that could be used to test this hypothesis.

One of the main conclusions to be drawn from this work is that category learning researchers should consider how stimuli are encoded in categorization tasks. Differences in stimulus properties (such as having continuous or discrete valued features) and expertise are only some of the areas where encoding effects may arise in category learning. We present the two-pathway framework as a potential heuristic for incorporating ideas from the object recognition literature where encoding has been more thoroughly studied. Because this approach is necessarily exploratory, it is useful to consider possible single-system approaches.

At the outset, the present framework was used to suggest that these two types of encoding, part- and image-based, use fundamentally different pathways. This follows some hybrid accounts in the object recognition literature that use both types of representations in isolated systems designed to fulfill disparate roles in object recognition (e.g., JIM II; Hummel & Stankiewicz, 1996). It could be the case, as other contemporary hybrid accounts suggest (e.g., Ullman et al., 2002), that part-based and holistic representations represent opposite ends of a single hierarchy. It is an empirical question whether these two types of encoding do represent different systems, or whether they are manifestations of a single system; there are likely to be advantages to both perspectives, since they may provide subtly different ways of approaching encoding in category learning. For example, viewing them as separate pathways helps to highlight the probable differences in the compositionality as well as automaticity between representing objects as a set of parts as opposed to a unitary whole. At the same time, viewing them as opposite ends of a hierarchy suggests that there are likely to be representations between these two extremes; this account may be better equipped to characterize how the ability to encode objects changes over time. The present experiments provide both novel and important contributions to the category learning literature, regardless of which (if either) of these two interpretations is ultimately correct. In this spirit, future research should be aimed primarily at discovering further constraints on stimulus encoding in category learning, as well as further integrations of the object recognition and categorization literatures, as opposed to focusing solely on how many pathways or systems are required to account for the extant results. One way to facilitate integration of the present framework and the object recognition literature would be to test predictions for real-world stimuli such as natural scenes. Although the current stimuli are well controlled and typical of category learning experiments, they are visually sparse in comparison with those used in the object recognition literature.

One avenue for future research is to further test possible representational differences between the two encoding pathways that we have proposed by using manipulations more common in the object recognition and expertise literatures. This would help to further strengthen the connections between the fields, as well as provide additional evidence about the nature of the representations in the two pathways. One type of manipulation often used to test whether representations are holistic is testing a subject's recognition sensitivity ($d'$) when the spatial relations between stimulus features have been distorted (e.g., Farah et al., 1998; Tanaka & Gauthier, 1997). Applying these manipulations in the present experiments would involve developing a post-training task requiring subjects to state whether one of the features (height or line-segment position) of a probe stimulus matches the features of a previously seen item. Holistic processing would be consistent with a decreased ability to recognize exception features when removed from the whole, whereas part-based encoding should not be disturbed.

## Relationships With Other Approaches

**Automaticity**. In addition to the clear relationships with work in object recognition and categorization, the present work is readily connected to other topics in cognitive psychology, such as theories of automaticity (e.g., Logan, 1988). These theories distinguish between cases in which people use memory to rapidly retrieve a response (e.g., recalling the result of a previously solved addition problem) and cases in which people effortfully and algorithmically solve a problem (e.g., an addition problem). Similar to these theories, we focus on the possibility that automatic (image-based) and algorithmic (part-based) processes could both be used to accomplish a particular task. We also suggest, in a similar vein to Logan, that the algorithmic processes might still operate and serve a variety of other purposes after expert-level performance via the image-based pathway has been achieved, particularly in the communicative roles that concepts serve. Our work builds on existing efforts in the categorization literature that incorporate notions of automaticity (e.g., Johansen & Palmeri, 2002; Nosofsky & Palmeri, 1997; Palmeri & Tarr, 2008). One unique aspect of our approach is its focus on different encoding formats rather than on the relative processing times of different solution paths.

One potential criticism of the present studies from an automaticity standpoint is that subjects in the two conditions used in Experiments 2 and 3, the extended category learning and identification conditions, were not equal in terms of their skill development at the final-day category learning and identification tests. Because subjects in the extended category learning condition were tested using exactly the same procedures and category structures that they had trained extensively on, it comes as little surprise that in Experiment 2A they tested significantly better than did subjects in the identification condition who had been trained on a different task. Although we agree that this potential criticism affects the interpretation of Experiment 2A, we know of no theory of automaticity that suggests that extending the duration of *feedback* (Experiments 2B and 3) would reverse this deficit and enable subjects in the identification condition to encode the exceptions. Rather, in applications to category learning (Nosofsky & Palmeri, 1997), the speed-up in task performance due to automaticity is described as occurring prior to subject's responses; lengthening feedback should have no effect. Thus, although it is unlikely that the choice of training regimens severely affects the interpretation of the data, future research may benefit from using training regimens that are better matched in terms of ease of transfer.

**Integral versus separable dimensions**. Another area that has an important relationship to the present research is Garner's (1974; see also Maddox, 1992) work on integral versus separable dimension stimuli. Integral dimension stimuli are ones for which the stimulus dimensions combine and are difficult for subjects to consider separately from one another, whereas separable dimension stimuli allow subjects to consider each stimulus dimension in isolation. Likewise, we proposed two pathways for stimulus encoding, one of which (image-based) relies on holistic, nondecomposable representations, and another that uses representations made up of individual stimulus parts (part-based). It is likely that some of the same psychological mechanisms underlie these two frameworks. Separable dimension stimuli are likely to encourage analytic processes such as part-based encoding, and integral dimension stimuli may encourage more holistic processing such as image-based encoding.

There are, however, important aspects of the present framework that go beyond the classic integral/separable distinction. Whereas Garner (1974) was interested in whether stimulus dimensions could be analyzed independently from one another, we focus on how subjects individuate values along a dimension or images within a psychological space. Furthermore, the typical solution for modeling differences between integral and separable dimension stimuli is to use different distance metrics; integral stimuli are fit using the Euclidean metric, whereas separable dimension stimuli are fit using city-block (e.g., Nosofsky, 1986). This would not help to explain any of the findings presented here, but insights from research on the integral/separable distinction may provide directions for future research. Indeed, all stimuli used in the present experiment were separable dimension stimuli, and it is possible that tasks using integral dimension stimuli, for which holistic representations are the default approach, may yield different results. Like the encoding differences we observed between continuous and discrete stimuli, considering how object properties such as integrality versus separability affect encoding will be a necessary step toward providing a more thorough understanding of the relationships between object perception and categorization.

**Dual coding theory and perceptual symbols**. An additional class of models that share important relationships with the present framework are the perceptual representation theories proposed by dual coding theory (DCT; Paivio, 1969, 1991) and the language and situated simulation theory (LASS; Barsalou, Santos, Simmons, & Wilson, 2008), which is formulated from Barsalou's (1999) work on perceptual symbol systems. Like the present approach, these theories suggest that there are multiple ways in which objects can be encoded and represented. However, unlike the present framework, these theories explicitly reject the idea of amodal propositional representations, and suggest that all representations are stored as analogues in the brain's various perceptual and linguistic systems. Our present framework, in contrast, remains uncommitted to a precise representational format for each pathway, and leaves open the possibility that images could be stored in an amodal "psychological space" or object parts stored in an amodal language of thought (e.g., Fodor, 1975), or in modal systems as described by DCT and LASS.

Another crucial difference between perceptual symbol theories and the present framework is that in DCT and LASS, both the linguistic and perceptual systems (e.g., vision) are proposed to have symbolic capabilities (e.g., compositional structure), whereas we have contrasted our two pathways along this dimension. Like the description of hierarchical image-based representations described above, DCT or LASS might make distinctions between symbols that represent both whole objects and object parts, but both would have the same basic representational format (in this case, symbolic).

## CONCLUSION

In the present investigation, we examined different ways in which stimuli can be encoded in categorization, particularly in cases in which certain objects must be stored separately and require different responses from perceptually similar objects. We proposed a two-pathway framework, motivated by theories in the object recognition literature, that suggests that people can learn to encode objects either automatically, from having extensive experience with them (image-based encoding), or by mapping the object's features onto a compositional vocabulary (part-based encoding). This led us to successfully predict processing differences between discrete and continuous-valued stimuli, as well as differences between subjects who received extensive experience in categorizing exceptions, or learned to assign labels to parts of the stimuli. The findings are important for the category learning literature, in which encoding has not been a thoroughly investigated topic. Although these findings follow from the present framework, they provide insight into category learning performance independent of any theoretical perspective. It is our hope that the research presented here may help to open up broader discussion between the categorization, object recognition, automaticity, and language and thought literatures.

### AUTHOR NOTE

This work was supported by AFOSR Grant FA9550-04-1-0226 and NSF Grant 0349101 to B.C.L. and AFOSR Grant FA9550-06-1-0204 and NIMH Grant MH077708 to W.T.M. We thank Matt Jones, John Dennis, and Darrell Worthy for helpful comments on a previous draft. Correspondence concerning this research should be addressed to T. Davis, University of Texas at Austin, 1 University Station A8000, Austin, TX 78712 (e-mail: thdavis@mail.utexas.edu).

### REFERENCES

Ahissar, M., & Hochstein, S. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, **36**, 791-804.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, **105**, 442-481.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences*, **22**, 577-660.

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245-283). Oxford: Oxford University Press.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, **115**, 107-117.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, **127**, 107-140.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. W. (1998). What is "special" about face perception? *Psychological Review*, **105**, 482-498.

Foard, C. F., & Kemler Nelson, D. G. (1984). Holistic and analytic modes of processing: The multiple determinants of perceptual analysis. *Journal of Experimental Psychology: General*, **113**, 94-111.

Fodor, J. A. (1975). *The language of thought*. New York: Crowell.

Garner, W. R. (1974). *The processing of information and structure.* New York: Wiley.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 431-446.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training "Greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, **38**, 2401-2428.

Hummel, J. E., & Stankiewicz, B. J. (1996). An architecture for rapid, hierarchical structural description. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 93-121). Cambridge, MA: MIT Press.

Johansen, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology*, **45**, 482-553.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, **19**, 577-621.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, **111**, 309-332.

Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 147-180). Hillsdale, NJ: Erlbaum.

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

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of “multiple-system” phenomena in perceptual categorization. *Psychonomic Bulletin & Review*, **7**, 375-402.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, **104**, 266-300.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, **101**, 53-79.

Nosofsky, R. M., & Zaki, S. R. (2003). A hybrid-similarity exemplar model for predicting distinctiveness effects in perceptual old–new recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 1194-1209.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, **76**, 241-263.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, **45**, 255-287.

Palmeri, T. J., & Gauthier, I. (2004). Visual object understanding. *Nature Reviews Neuroscience*, **5**, 291-303.

Palmeri, T. J., & Nosofsky, R. M. (1995). Recognition memory for exceptions to the category rule. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 548-568.

Palmeri, T. J., & Tarr, M. J. (2008). Visual object perception and long-term memory. In S. Luck & A. Hollingsworth (Eds.), *Visual memory* (pp. 163-207). New York: Oxford University Press.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, **343**, 263-266.

Regehr, G., & Brooks, L. R. (1993). Perceptual manifestations of an analytic structure: The priority of holistic individuation. *Journal of Experimental Psychology: General*, **122**, 92-114.

Sakamoto, Y., & Love, B. C. (2004). Schematic influences on category learning and recognition memory. *Journal of Experimental Psychology: General*, **133**, 534-553.

Schyns, P. G., Goldstone, R. L., & Thibaut, J.-P. (1998). The development of features in object concepts. *Behavioral & Brain Sciences*, **21**, 1-54.

Stewart, N., Brown, G. D., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, **112**, 881-911.

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science*, **16**, 145-151.

Tanaka, J. W., & Gauthier, I. (1997). Expertise in object and face recognition. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 36, pp. 83-125). San Diego: Academic Press.

Tarr, M. J., & Gauthier, I. (1998). Do viewpoint-dependent mechanisms generalize across members of a class? *Cognition*, **67**, 73-110.

Tversky, A. (1977). Features of similarity. *Psychological Review*, **84**, 327-352.

Ullman, S., Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use in classification. *Nature Neuroscience*, **5**, 682-687.

Zhang, L., & Cottrell, G. W. (2005). Holistic processing develops because it is good. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum.

## NOTES

1. One exception to this is the hybrid similarity exemplar model (Nosofsky & Zaki, 2003), which was designed to make use of both continuous and discrete features of stimuli. We will discuss applications of this model to the current framework below.

2. Averaged data were used because the most complex model (Cx Dx Cr Dr + standard GCM) contained more free parameters (six) than an individual subject provided unique data points (four).

3. Due to the symmetry of the design, target category, foil category, and relevant dimension did not provide unique similarity relations, and were therefore collapsed across.

4. A factorial condition (discrete or continuous) × item type (exception or frequency-matched rule item) × exception placement (near the decision bound or far; see Figure 4) ANOVA found significant interactions for the category learning [$F(1,99) = 15.586$, $MS_e = .046$, $p < .001$, $\eta_p^2 = .136$] and recognition [$F(1,99) = 11.211$, $MS_e = .029$, $p < .001$, $\eta_p^2 = .10$] phases.

5. Unfortunately, bow effects are not examined in Experiments 2 and 3 because the relatively small number of subjects in these multisession studies provided an insufficient number of recognition and final-block category learning trials to conduct the bow effect analyses.

6. Failure to reach significance in this case is largely due to higher error variance for exception performance in the identification condition. A nonparametric bootstrap analysis was conducted on the difference between means for this comparison. Over 10,000 simulations, only 33 resulted in identification condition performance with exception items exceeding that of the extended category learning condition ($p = .0033$).

(Manuscript received May 14, 2008;
revision accepted for publication December 18, 2008.)