Category structure and region-specific selective attention

Robert M. Nosofsky[1] · Mingjia Hu[1]

Abstract

A fundamental component of human categorization involves learning to attend selectively to relevant dimensions and ignore irrelevant ones. Past research has shown that humans can learn flexible strategies in which the attended dimensions vary depending on the region of feature space in which classification takes place. However, region-specific selective attention (RSA) is often challenging to learn. Here, we test the hypothesis that RSA is facilitated when individual categories are embedded within single regions of stimulus space rather than dispersed across multiple regions. We conduct an experiment that varies across conditions whether categories are embedded within regions, but in which the same RSA strategy would benefit performance across the conditions. To evaluate the hypothesis, we use measures of overall performance accuracy as well as comparisons among formal computational models that do and do not make allowance for RSA. We find strong support for the hypothesis among the upper-median-performing participants in the tested groups. However, even in the condition that promotes the learning of RSA, performance is considerably worse than in comparison conditions in which a single set of dimensions can be attended throughout the entire stimulus space.

Keywords Categorization

A fundamental component of human categorization involves the process of selective attention to relevant dimensions (Shepard et al., 1961). In general, observers learn to attend to dimensions that are relevant for solving classification problems and to ignore dimensions that are irrelevant. However, even within a fixed domain, the dimensions that are most relevant may vary depending on the specific objects that one is trying to classify. Consider the examples in Fig. 1 involving rock types as defined in the geologic sciences (Marshak, 2019). The top panel contrasts the pair obsidian and anthracite, both of which are shiny black rocks. However, whereas obsidian tends to have a smooth, glassy, scalloped surface, anthracite tends to have a rougher and more layered texture. Hence, dimensions related to surface texture are diagnostic for discriminating between obsidian and anthracite. Alternatively, the bottom panel contrasts the pair breccia and conglomerate, both of which have large-size fragments embedded in a fine-grained groundmass. However, whereas the fragments in breccia are sharp and angular, the fragments

in conglomerate are smooth and rounded. Thus, for this pair, the relevant dimension is fragment shape rather than surface texture.
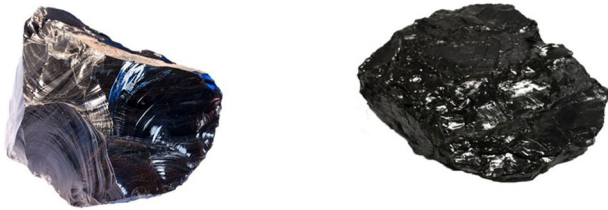
In numerous formal models of categorization, such as exemplar, prototype, and clustering models, the process of dimensional selective attention is formalized in terms of a set of weights that influence the psychological distances along the dimensions that compose the objects (e.g., Kruschke, 1992; Love et al., 2004; Minda & Smith, 2001; Nosofsky, 1986). Selective attention to a dimension, reflected in terms of a high weight, results in "stretching" of the space along that dimension, whereas a low weight results in "shrinking" of the space along an unattended dimension. These early models of category-based attention, however, formalized only a single set of dimension weights for a given task. The same set of weights was presumed to apply for all stimuli and all regions of the psychological space in which the objects were embedded. Following Blair et al. (2009), we will refer to this type of implementation as *task-specific attention*.

Clearly, however, this implementation is inadequate for our Fig. 1 example, where the relevant dimensions change depending on the particular stimuli that need to be classified. Instead, a region- or stimulus-specific form of selective

✉ Robert M. Nosofsky
  nosofsky@indiana.edu

[1] Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, USA

Obsidian vs. Anthracite

Breccia vs. Conglomerate

Fig. 1 Example rock pairs

Table 1 Logical codings for the category structure tested in Blair et al.'s (2009) Experiment 2

Category	Dimension		
	1	2	3
A1	1	1	1
A1	1	1	2
A2	1	2	1
A2	1	2	2
B1	2	1	1
B1	2	2	1
B2	2	1	2
B2	2	2	2

attention seems to be involved, which requires a more flexible dimensional-weighting scheme.

This need for a more flexible region-specific or stimulus-specific attention scheme has been demonstrated in elegant experiments reported by previous researchers. First, in a study reported by Aha and Goldstone (1992), participants learned to classify a set of training stimuli varying along two dimensions into two categories. In one region of the space, the two categories were separated along the horizontal dimension, whereas in a second region, they were separated along the vertical dimension. Following training, participants classified transfer items into the categories. The key result was that participants generalized along the horizontal dimension in the region in which the horizontal dimension was relevant for classifying the training stimuli, whereas they generalized along the vertical dimension in the region in which the vertical dimension was relevant. In addition, Aha and Goldstone (1992) demonstrated that an exemplar model with a stimulus-specific attention-weighting scheme yielded better quantitative fits to the classification data than did the standard version of the model with a fixed set of weights.

Blair et al. (2009) also provided compelling evidence for the operation of a region-specific or stimulus-specific form of selective attention. Rather than relying on patterns of generalization to transfer stimuli, these investigators used eye-tracking methods to provide measures of attention allocation (Rehder & Hoffman, 2005). The category structure used in their Experiment 2 is shown in our Table 1. There were eight stimuli varying along three binary-valued dimensions. The stimuli were divided into four categories (A1, A2, B1, B2)

with two stimuli in each category. As shown in the table, if a stimulus had Value 1 on Dimension 1, then the value on Dimension 2 could be used to classify the stimulus into Category A1 or A2 (with Dimension 3 being irrelevant). Alternatively, if the stimulus had Value 2 on Dimension 1, then the value on Dimension 3 could be used to classify the stimulus into Category B1 or B2 (with Dimension 2 being irrelevant). The results for the set of subjects who met a learning criterion were clear-cut: the eye movement data indicated that, for stimuli belonging to Categories A1 and A2, the least attention was allocated to the irrelevant Dimension 3; but for stimuli belonging to Categories B1 and B2, the least attention was allocated to irrelevant Dimension 2. Again, the attention devoted to the different dimensions varied with the specific stimuli that were being classified.

Other related studies also provide evidence that the nature of category representations may be region-specific or stimulus-specific. For example, Yang and Lewandowsky (2003, 2004) and Little and Lewandowsky (2009) provided evidence that a context cue that is irrelevant on its own can "gate" the use of alternative decision-boundary strategies in different regions of a category space. For example, in Yang and Lewandowsky (2003), when provided with one context cue, observers learned to use a positively sloped decision boundary to separate members of contrasting categories, whereas when provided with a second context cue, observers learned to use a negatively sloped decision boundary. Assuming the context cue indicates different regions of the psychological space, these results provide another example of how category representations can be region specific. Likewise, Erickson and Kruschke (1998, 2002) provided compelling evidence of stimulus-specific and region-specific category representations. In their experiments, observers learned to use "rules" for partitioning large swaths of stimulus space but to rely on similarity to stored examples in other specific regions of the space.

In summary, there is clear-cut evidence that dimensional selective attention and related forms of category
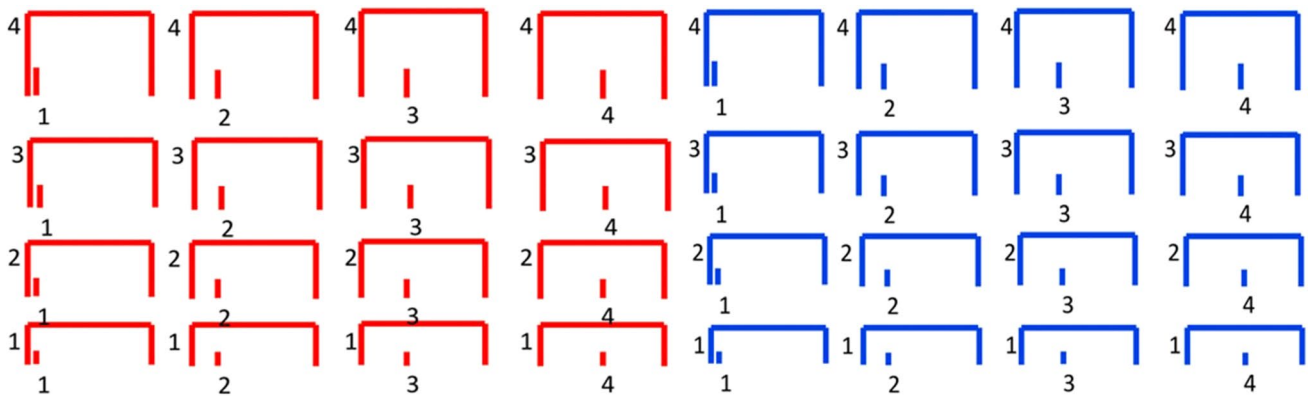
Fig. 2 Illustration of the stimulus set used in all conditions

processing and representation can be region specific and stimulus specific. Nevertheless, our impression is that the evidence also suggests that the development of region-specific attention is highly cognitively demanding and not nearly as psychologically "primary" as the task-specific selective attention formalized in the early foundational models. For example, in Blair et al.'s studies, 9 of 27 participants failed to meet the learning criterion in Experiment 1 and 16 of 38 participants failed in Experiment 2 (see Blair et al., 2009, pp. 1199, 1202). Similarly, in their Experiment 1, Yang and Lewandowsky (2003, p. 667) reported that classification performance in one of the context-gated regions of the stimulus space was close to chance. In a recent article, Braunlich and Love (2022) presented a sampling emergent attention (SEA) model of concept learning that can produce the types of region-specific and stimulus-specific selective attention described above (see also, e.g., Kruschke, 2001; Jones & Love, 2006). Included among their applications were simulations of the Blair et al. (2009) Experiment 2 paradigm. Although an impressive model, Braunlich and Love (2022, pp. 229–230) summarize the modeling outcomes by writing: "Across the simulations considered here, SEA's behavior could be characterized as somewhat idealized in comparison to human participants" (pp. 229–230). (We provide a much fuller discussion of the SEA model and region-specific attention in our General Discussion.)

The question therefore arises what factors may promote the successful development of region-specific selective attention and under what conditions is it most likely to be observed. Here, we propose and test the hypothesis that one important factor involves the correspondence between the category structures and the regions of the stimulus space in which the to-be-classified objects are embedded. In particular, as we elucidate in more detail below, we hypothesize that effective region-specific attention learning is more likely to operate when individual categories are embedded in single regions rather than dispersed across multiple regions. One

basis for this hypothesis is that, under conditions in which it would behoove observers to adopt region-specific selective-attention, each category would be assigned a fixed set of region-specific attention weights if embedded in a single region. By contrast, observers would need to learn multiple sets of region-specific attention weights for individual categories if those categories are dispersed across multiple regions. The need to learn multiple sets of region-specific attention weights seems a cognitively complex requirement for representing individual categories. Hence, we hypothesize that region-specific attention learning will operate more effectively in conditions in which individual categories are embedded within single regions. As described in the next section, a second basis for our hypothesis is that logical rules for stating membership in individual categories tend to be cognitively simple when categories are embedded in single regions but cognitively complex when dispersed across multiple regions.

Experiment

The design of our experiment is illustrated schematically in Figs. 2 and 3. Figure 2 illustrates the structure of the stimulus set used in all conditions. The stimuli were colored open rectangles with an embedded vertical line and they varied orthogonally along three dimensions: color (red or blue), left-right position of the line (four continuous values labeled 1–4), and rectangle height (four continuous values labeled 1–4).

The two major experimental conditions are the *four-category* and *two-category* conditions illustrated in Fig. 3. (The layout of the figure corresponds to the one illustrated in Fig. 2.) In all conditions, there was an initial training phase in which participants learned to classify a set of training stimuli into categories. Following training there was a test phase in which participants classified both the training
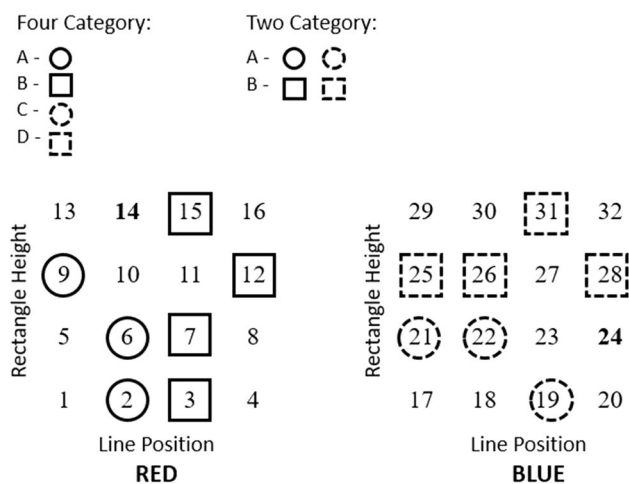
Fig. 3 Schematic illustration of the four-category and two-category conditions. Critical-transfer stimuli 14 and 24 are in boldface font

stimuli as well as novel transfer stimuli. The training stimuli are those enclosed by geometric shapes in Fig. 3, whereas the transfer stimuli are not enclosed by geometric shapes. As shown, in the four-category condition, the training stimuli of Category A were 2, 6, and 9; of Category B were 3, 7, 12, and 15; of Category C were 19, 21, and 22; and of Category D were 25, 26, 28, and 31. In the two-category condition, the training stimuli of Categories A and C were merged into a single Category A (indicated by solid and dashed circles); and the training stimuli of Categories B and D were merged into a single Category B (indicated by solid and dashed squares).

As can be seen, different dimensions are relevant depending on whether stimuli occupy the red versus the blue regions of the stimulus space. In the red region, the dimension of line position is most relevant for separating the training examples of the contrasting categories; but in the blue region, the dimension of rectangle height is most relevant. This same region-specific relevance of the dimensions holds for both the four-category and two-category conditions. To the extent that observers learn to adopt region-specific attention, we expect they will give greater attention weight to line-position in the red region, but to rectangle-height in the blue region.

Our hypothesis is that this form of region-specific selective attention will be learned more effectively in the four-category condition than in the two-category condition. As stated in our introduction, there are multiple bases for this hypothesis. One basis, which we expand upon in the Model-based Analysis section of our article, is that the four-category condition affords the use of a single set of region-specific attention weights for each category, but such is not the case in the two-category condition. Another basis can be stated using the language of logical rules.[1] In the

four-category condition, there is a unique combination of color and a set of values on the contingent second relevant dimension that is associated with each individual category. For example, in the four-category condition, members of Category A can be described as red with left-positioned lines; members of B as red with right-positioned lines; members of C as blue and short in height; and members of D as blue and tall in height. By contrast, there is no single conjunctive combination of dimension values associated with individual categories in the two-category condition. Instead, *within* each category, subjects need to switch their attention to different dimensions depending on the color of the stimulus. For example, in the two-category condition, Category A is composed of stimuli that are red with left-positioned lines OR that are blue with short rectangle heights. Thus, application of region-specific selective attention for stating rules appears to be more cognitively complex in the two-category than in the four-category condition. This added complexity is likely to tax the working memory that is needed for constructing rules more in the two-category than in the four-category condition.

Our design includes certain *critical transfer* stimuli to help diagnose the operation of region-specific selective attention. For these experimental conditions, the critical transfer stimuli are 14 and 24 (see Fig. 3). Transfer Stimulus 14 is highly similar to Training Stimulus 15, which belongs to Category B. However, Stimulus 14 is red with a left-positioned line, which is diagnostic of Category A. Thus, the category into which subjects classify Stimulus 14 will provide evidence of the type of strategy they are using. In particular, if they classify Stimulus 14 into Category A, then it provides evidence that, for the red stimuli, they are giving greater attention to the line-position dimension than to rectangle height. We will perform an analogous analysis on Stimulus 24, which is highly similar to Training Stimulus 28 from Category D, but whose value on the relevant rectangle-height dimension matches the training items from Category C. If subjects tend to classify Stimulus 24 into Category C, it provides evidence that, for the blue stimuli, subjects are giving greater attention to rectangle height than to line position. Finally, if subjects tend to classify Stimulus 14 into Category A *and* Stimulus 24 into Category C, it provides evidence that region-specific selective attention is operating. If our hypothesis is correct, then subjects in the four-category condition will

---

[1] As is true of previous work that has sought evidence for the operation of region-specific selective attention, our experiment was not designed to develop formal contrasts between the predictions from selective-attention exemplar models and logical rule models. For previous work illustrating the difficulty of developing such contrasts, see, for example, Nosofsky and Johansen (2000) and Jones and Love (2006).
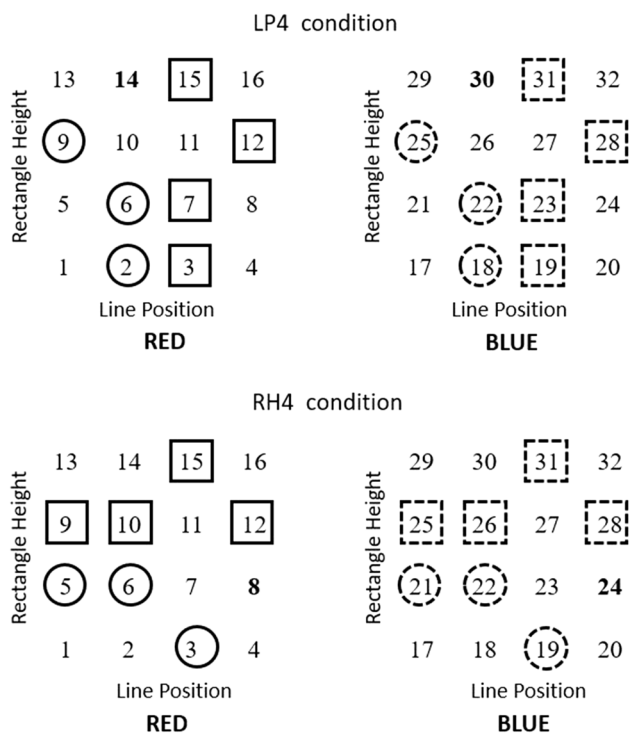
Fig. 4 Schematic illustration of the LP4 and RH4 conditions. Critical transfer stimuli are shown in boldface font

categorize Stimulus 14 into Category A more frequently than in the two-category condition. Likewise, subjects in the four-category condition will classify Stimulus 24 into Category C more frequently than will subjects in the two-category condition classify Stimulus 24 into Category A.

Finally, as a source of comparison, we also tested the four-category conditions illustrated in Fig. 4. In these conditions, participants can perform well by adopting a uniform task-specific set of attention weights that are not region specific. For the top-panel structure, we expect that participants will attend selectively to color and line position while giving little attention to rectangle height; whereas for the bottom-panel structure, they will attend selectively to color and rectangle height while giving little attention to line position. Thus, in the top-panel structure, we hypothesize that the critical Transfer Stimuli 14 and 30 will tend to be classified in Categories A and C, respectively. In the bottom-panel structure, we hypothesize that the critical transfer stimuli 8 and 24 will tend to be classified into Categories A and C, respectively. We expect that because region-specific selective attention is not required, performance in the Fig. 4 structures will be superior to performance in both the four-category and two-category conditions of Fig. 3. For ease of description, we will refer to the top-panel structure as the four-category line-position

condition (LP4) and to the bottom-panel structure as the four-category rectangle height condition (RH4).

Method

Subjects

A total of 180 undergraduates from Indiana University participated in the experiment. There was a total of 62 subjects randomly assigned to the two-category condition and 60 subjects to the four-category condition. In addition, 31 and 27 subjects were randomly assigned, respectively, to the LP4 and RH4 conditions. The sample sizes yield power = .99 for detecting the large-size effects that we anticipated in the main statistical analyses of the performance data. Subjects received credit toward an introductory psychology course requirement in exchange for their participation. A screening requirement was that subjects have normal or corrected-to-normal vision and normal color vision.

Stimuli and apparatus

The stimuli were colored open-ended rectangles with a vertical line within them (Fig. 2). The width of the rectangle was 10.47cm and the height of the line was 2.54 cm. The stimuli varied in color (red or blue), rectangle height (4.13, 5.40, 6.67, or 7.94 cm), and left-right positioning of the line (.79, 2.06, 3.33, or 4.60 cm away from the far-left side of the rectangle). The stimuli were presented in the center of the computer screen on a white background. To assist subjects in completing the tasks, a numerical measuring scale was provided along both the horizontal and vertical axes of the computer display that indicated the levels (1–4) of each of the continuously varying dimensions. This same measuring scale was provided in all four conditions.

Subjects completed the experiment individually on PC desktop computers, with each subject tested in a private, soundproof booth. Subjects sat approximately 20 in. away from the computer screen.

Procedure

An initial instruction message described the three dimensions of the stimuli and showed an example. Subjects were also informed of the number of categories into which they would be classifying the stimuli and which keys on the keyboard corresponded to which category. (Labels were placed on the S, D, K, and L keys to indicate the category responses A, B, C, and D, respectively.) Subjects were instructed that on each trial they would be shown an object and should classify it into one of the categories, after which they would be told the correct category. The instructions also clarified that

at first subjects would be guessing, but by paying attention to the objects and correct answers they could learn to classify the objects accurately.

The experiment consisted of two phases—the training phase and the transfer phase. The training phase consisted of up to 10 blocks of training. Each training block had 14 trials, with each one of the 14 training stimuli shown once in a random order. After classifying each stimulus, subjects were told if their selection was correct or incorrect. If their selection was incorrect, they were also told the correct category. The stimulus remained onscreen during this feedback period. The feedback was presented for 1 s on correct trials and for 2 s on incorrect trials. There was a .5-s intertrial interval with a blank computer screen. After each block, subjects were told their overall percentage of correct responses. Following this training phase there was a break period, and subjects were asked to press the space bar when they were ready to continue to the transfer phase.

During the training phase, it was possible for subjects to complete less than 10 training blocks if they achieved 100% correct classifications for three training blocks in a row. Subjects were informed of this possibility at the start of the experiment to motivate them to learn the categories.

In the transfer phase, subjects were tested on all 32 stimuli. They were tested for two blocks with a total of 52 trials per block. Within each block, each of the original 14 training stimuli was shown twice; each of the two critical transfer stimuli was shown four times; and each of the other 16 transfer stimuli was shown once. The order of presentation was randomized in each block. The critical transfer stimuli were shown more often than the other transfer stimuli because of their high diagnostic value for assessing subjects' classification strategies. Subjects continued to receive corrective feedback on the training stimuli. However, no feedback was given for the transfer stimuli; instead, subjects were simply informed that their response had been recorded.

Results

Based on findings from previous research on region-specific attention and category representations, as well as pilot work involving the present experiment, we expected the four-category and two-category conditions to be highly challenging for numerous participants to learn. Indeed, given the difficulty of the tasks, we expected many subjects to fail to solve completely the classification problems and/or to resort to random guessing. Therefore, rather than analyzing all participants en masse, we decided that a more coherent presentation would arise if we focused our main analyses on the top-performing subjects in each condition, and analyzed separately the results for poor-performing subjects.

Accordingly, in the present Results section and the ensuing Formal Modeling Analyses section, we conducted median splits on the accuracy achieved by subjects for the training stimuli that were presented during the test phase. Using this performance measure as a basis for forming subgroups, we report separately the results for the upper-median and lower-median in each of the conditions.[2]

Note that there are objective correct and incorrect answers associated with the old training stimuli, because subjects received feedback of category assignments for these stimuli. However, no objective feedback was ever presented for the new transfer stimuli. Here we define "accuracy" for the new transfer stimuli as the proportion of a subject's answers that were in accord with the region-specific selective-attention hypothesis. Thus, making reference to Fig. 3, in the four-category condition, the "correct answer" for transfer stimuli that are red with Line Positions 1 and 2 is Category A; for transfer stimuli that are red with Line Positions 3 and 4 is B; for transfer stimuli that are blue with Heights 1 and 2 is C; and for transfer stimuli that are blue with Heights 3 and 4 is D. Analogous definitions of "correct answers" for the new transfer stimuli arise for the two-category condition. In Conditions LP4 and RH4, "correct answers" are defined in terms of consistency with the task-specific selective-attention hypothesis: attention to color and line position for all regions in the LP4 condition; and to color and rectangle height in all regions in the RH4 condition.

For the upper-median subjects, the mean proportions of correct responses to the training stimuli, critical-transfer stimuli, and other-transfer stimuli are shown for each of the four conditions in Fig. 5. The most important result is that, as hypothesized, mean accuracy for the critical-transfer stimuli was significantly higher in the four-category-condition ($M = .706$) than in the two-category condition ($M = .335$), $t(59) = 3.99$, $p < .001$. Indeed, this contrast in results for the critical-transfer stimuli is quite dramatic, with the four-category proportion being substantially greater than .50 and the two-category proportion being substantially less than .50. The mean accuracies for the other-transfer stimuli provide converging evidence that region-specific selective-attention process operated more effectively in the four-category condition ($M = .903$) than in the two-category condition ($M = .786$), $t(59) = 3.78$, $p < .001$. Finally, mean accuracy on the training items themselves was significantly higher in the four-category condition ($M = .943$) than in the two-category condition ($M = .883$), $t(59) = 3.82$, $p < .001$. There are

---

[2] Because our focus is on performance at time of transfer, we use transfer-phase performance as our criterion for forming subgroups rather than learning-phase performance. For a variety of reasons, it is possible that some participants who perform well in the training phase may stop performing well on the training stimuli during transfer.
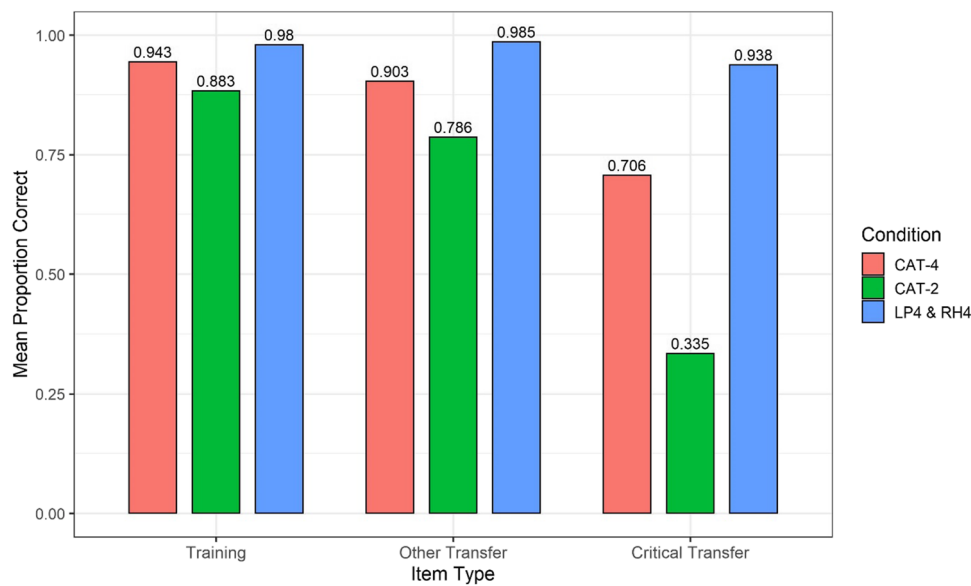
Fig. 5 Mean proportion correct for each item type in each of the conditions

various potential routes to achieving accuracy on the training items. However, a reasonable possibility is that learning the region-specific selective-attention strategy was easier in the four-category condition than in the two-category one and this allowed for higher performance on the training stimuli as well.

The result that accuracy was higher in the four-category than in the two-category condition for all three stimulus types is an impressive finding: Because there are only two response options in the two-category condition but four response options in the four-category condition, chance responding would yield higher accuracy in the two-category condition. The finding that accuracy was higher in the four-category condition is based on a conservative performance measure that does not even involve a correction for guessing.

Despite this evidence that region-specific selective attention operated more effectively in the four-category than in the two-category condition, a comparison with performance in the LP4 and RH4 conditions suggests that, even in the four-category condition, learning of region-specific selective attention is much more cognitively demanding than learning of task-specific selective attention. In the LP4 and RH4 conditions, subjects need attend to only the same two dimensions across all regions of the stimulus space. In the four-category condition, subjects need attend to only two dimensions for any given region, but the relevant dimensions change across regions. As can be seen in Fig. 5, this requirement of learning region-specific selective attention led to much worse performance in the four-category condition compared with the LP4 and RH4 conditions, for all three stimulus types. Combining the data across the LP4 and RH4 conditions, accuracy was significantly higher in these conditions than in the four-category condition

for the training items ($M = .980$), $t(58) = 3.69$, $p < .001$; the other-transfer items ($M = .985$), $t(58) = 3.84$, $p < .001$; and the critical-transfer items ($M = .938$), $t(58) = 3.34$, $p = .001$.

The results for the lower-median subjects in all four conditions are reported in the Appendix Fig. 8. In brief, mean performance in the four-category condition tended to be slightly *worse* than in the two-category condition for all three stimulus types. However, included in the lower median are significant numbers of subjects who performed at near-chance or chance levels, and chance responding is greater in the two-category condition than in the four-category one. This poor performance in the lower median may reflect numerous factors, including failure to understand instructions, insufficient motivation, and the difficulty of learning region-specific selective attention. Excluding the at-chance subjects, we were unable to discover a general performance pattern that characterized a large number of subjects in the lower-median groups and it appears to involve a mix of highly idiosyncratic classification-learning strategies. As expected, even for the lower-median subjects, performance was significantly better in the LP4 and RH4 conditions than in both conditions requiring region-specific selective attention. The individual-subject trial-by-trial learning and transfer data for all participants in this study are available online (https://osf.io/ysb9w/).

Model-based analysis

Overview

To provide converging evidence to support our conclusion that region-specific selective attention operated more

effectively in the four-category (Cat-4) condition than in the two-category (Cat-2) condition, we fit a set of formal models to the classification data. Following Aha and Goldstone (1992), we used Nosofsky's (1986) generalized context model (GCM) as an analytic device, using distance functions with different attention-weighting schemes to assess the hypotheses.

The standard model

According to the standard GCM, the evidence in favor of each category is found by summing the similarity of a test item to all training items in each category; the test item is most likely to be classified into the category that yields the greatest summed similarity. Specifically, the probability that item i is classified in category J is found by summing its similarity to all members of Category J and dividing by the summed similarity of the item to all members of all categories K:

$$\Pr(J|i) = \frac{\sum_{j \in J} s_{ij}}{\sum_K \sum_{k \in K} s_{ik}}, \tag{1}$$

where s_{ij} denotes the similarity of test item i to training item j.[3]

The similarities between test and training items are derived from the distances between the items in the Fig. 3 configuration. We assumed for simplicity that the psychological coordinates of the stimuli on the line-position and rectangle-height dimensions matched the logical codings (1–4) of the physical-dimensions layout depicted in the figure. In addition, we assumed arbitrarily that red stimuli had coordinate value 1 on the color dimension and blue stimuli had coordinate value 4. (Because dimension-weight parameters will be included in the distance function, the assignment of these magnitudes is made without loss of generality.) Following Shepard (1964), in the baseline GCM that assumes task-specific selective attention, we assumed for these highly separable-dimension stimuli that the psychological distance between test item i and training example j was given by a weighted city-block metric:

$$d_{ij} = w_1 |x_{i1} - x_{j1}| + w_2 |x_{i2} - x_{j2}| + w_3 |x_{i3} - x_{j3}|, \tag{2}$$

---

[3] To make allowance for deterministic response strategies, applications of the GCM often include a response-scaling parameter in which the individual category summed similarities are raised to the power γ (Ashby & Maddox, 1993). For the present type of paradigm, however, it is difficult to obtain separate estimates of the response-scaling parameter γ and a sensitivity parameter c used for transforming distance into similarity (Nosofsky & Zaki, 2002). For simplicity in the present analyses and to reduce the number of free parameters, the response-scaling parameter is held fixed at $\gamma = 1$.

where x_{im} denotes the value of item i on dimension m ($m = 1$, Color; $m = 2$, Line Position; $m = 3$, Rectangle Height); and the attention-weight parameters w_m ($0 \leq w_m \leq 1$, $\sum w_m = 1$) represent the relative attention given to the dimensions when making classification responses.

Finally, the similarity between test-item i and training-example j was assumed to be an exponential decay function of their distance (Shepard, 1987):

$$s_{ij} = e^{-c d_{ij}}, \tag{3}$$

where the sensitivity parameter c describes the rate at which similarity declines with distance. Conceptually, the sensitivity parameter provides a measure of overall discriminability among items in the feature space. This "standard" GCM with task-specific selective attention (Eq. 2) makes use of three free parameters: the sensitivity parameter c (Eq. 3) and attention weights w_1 and w_2 (with $w_3 = 1 - w_1 - w_2$) (Eq. 2).

GCM with region-specific attention (GCM-RSA)

The version of GCM with region-specific attention (GCM-RSA) that we test here is a straightforward extension of the standard baseline model just described (see also Aha & Goldstone, 1992; Jones & Love, 2006). In particular, GCM-RSA uses a generalized distance function in which the values of the attention weights are now allowed to vary with the color region of the space that is occupied by test-item i. The specific formalization adopted here assumed that the distance between test-item i and exemplar j is given by

$$d_{ij} = w_{i1} |x_{i1} - x_{j1}| + w_{i2} |x_{i2} - x_{j2}| + w_{i3} |x_{i3} - x_{j3}|, \tag{4}$$

where $w_{im} = w_{m, red}$ if item i occupies the red region; and $w_{im} = w_{m, blue}$ if item i occupies the blue region. Within each color region, the attention weights are still constrained to sum to one.

More specifically, we assumed for simplicity that the attention weight given to the values on the color dimension itself ($m = 1$) were region-invariant, $w_{1, red} = w_{1, blue} = w_1$. In other words, the attention weight given to the color dimension itself does not depend on whether the test item occupies the red region or the blue region. However, the attention devoted to the line-position dimension ($m = 2$) is region specific. Presumably, greater attention weight is given to line position when the stimulus is red rather than blue (see Fig. 3): $w_{2, red} > w_{2, blue}$. Because the attention weights across the dimensions in each region are constrained to sum to 1, the region-specific attention weights on the rectangle-height dimension ($m = 3$) do not enter into the count of the number of free parameters. However, it is expected that the estimated attention weight given to rectangle height will be greater when the stimulus is blue rather than

red (see Fig. 3): $w_{3, blue} > w_{3, red}$. This region-specific version of the GCM adds only one extra free parameter to the standard version. The free parameters are the sensitivity parameter c (Eq. 3); the attention weight given to the color dimension, w_1; and the region-specific attention weights given to the line-position dimension, $w_{2, red}$ and $w_{2, blue}$ (Eq. 4). [Note: $w_{3, red} = 1 - w_1 - w_{2, red}$; and $w_{3, blue} = 1 - w_1 - w_{2, blue}$.]

Standard GCM with response bias

An alternative hypothesis regarding the patterns of classification is that the tendency for participants to classify the critical-transfer stimuli into Categories A and C in the Cat-4 condition may simply reflect a response bias. According to a response-bias explanation, regardless of the stimulus that is presented, there might be an overall bias to respond with Categories A and C, perhaps because there are fewer training examples from these categories than from Categories B and D (see Fig. 3). Although positing such a response bias could account for the patterns of classification for the individual critical-transfer stimuli, such a model predicts global effects on the response probabilities for all stimuli in all categories, which places strong constraints on the model. To test the response-bias hypothesis, we again fitted the standard GCM to the data, except we extended its response rule with a set of response bias parameters:

$$\Pr(J|i) = \frac{b_J \sum_{j \in J} S_{ij}}{\sum_K b_K \sum_{k \in K} S_{ik}}, \tag{5}$$

where b_K $(0 < b_K < 1)$ is the response bias associated with Category K. Without loss of generality, the response biases associated with the categories can be constrained to sum to one. Thus, extending the standard model with response bias adds 3 free parameters to the standard model for the Cat-4, LP4 and RH4 conditions, and adds one free parameter to the standard GCM for the Cat-2 condition.

Target data and method of model fitting

The target data for our modeling were the complete classification-confusion matrices observed in each of the four conditions obtained by aggregating across the upper-median subjects in each condition (available at https://osf.io/ysb9w/). The entry in each cell iJ of each classification matrix gives the frequency with which observers classified each item i into each Category J. Our criterion of fit for the models was to find the free parameters that maximized the likelihood (L) of the observed classification-confusion data in each condition. For the present case, the maximum-likelihood criterion is equivalent to minimizing the quantity

$$-\ln L = -\Sigma \Sigma_{iJ} \ f_{iJ} \ln \left(p_{iJ} \right), \tag{6}$$

where f_{iJ} is the observed frequency in cell iJ; p_{iJ} is the predicted probability from the model that item i is classified in category J; and the sum is across all cells of the classification matrix.[4] Because the alternative models have differing numbers of free parameters, methods are needed to assess the comparative model fits that correct for the number of free parameters. Here we use the Bayesian information criterion (BIC; Schwarz, 1978), given by

$$BIC = -2 \ln(L) + P\ln(N), \tag{7}$$

where L is the (maximum) likelihood of the data, P is the number of free parameters in the model, and N is the total number of observations forming the data set. The term $P\ln(N)$ is a penalty term that penalizes a model for its number of free parameters; the model that achieves a smaller BIC is considered to provide a more parsimonious account of the data. As an auxiliary measure, we also report the sum of squared deviations (SSD) between the predicted and observed classification probabilities for each model in each condition. We fitted the models to the data using the Hooke and Jeeves (1961) parameter-search algorithm, using 10 different random starting configurations for each search. Essentially all parameter searches converged to identical sets of best-fitting parameters for each model.

Model-fitting results

Overview The summary fits of the models to the data in each condition are reported in Table 2, with best-fitting parameters reported in Table 3. To provide insight into the reasons for the patterns of summary-fit results, Figs. 6 and 7 show scatterplots of observed against predicted classification probabilities in the Cat-4 and Cat-2 conditions for the different model versions.

Cat-4 condition As reported in Table 2 (top panel), in the Cat-4 condition, the version of the GCM with region-specific attention (GCM-RSA) yields an enormously better BIC fit than does the standard version of the model without region-specific selection attention. Inspection of the top two panels of Fig. 6 reveals the basis for this result. In the Cat-4 condition, the GCM-RSA yields a reasonably good fit to the entire matrix of classification probabilities (top panel). By comparison, the standard GCM (middle panel) fails on qualitative grounds to account for the results involving the critical-transfer stimuli: It predicts that the critical transfer stimuli

---

[4] This likelihood function assumes that that responses for each stimulus are multinomially distributed into the categories and that the individual distributions are independent.

Table 2 Summary fits of the models to the classification-confusion matrices of the Cat-4, Cat-2, line-position, and rectangle-height conditions

Cat-4 condition

Model	-lnL	BIC	SSD	P
GCM-RSA	1054.4	2141.0	0.327	4
Standard GCM	1253.0	2530.2	1.163	3
Standard GCM with Bias	1193.4	2435.0	1.046	6

Cat-2 Condition

Model	-lnL	BIC	SSD	P
GCM-RSA	1242.8	2517.9	0.392	4
Standard GCM	1266.7	2557.6	0.381	3
Standard GCM with Bias	1257.0	2546.2	0.424	4

Line-Position Condition

Model	-lnL	BIC	SSD	P
GCM-RSA	239.6	508.8	0.040	4
Standard GCM	239.6	501.4	0.040	3
Standard GCM with Bias	233.8	512.1	0.030	6

Rectangle-Height Condition

Model	-lnL	BIC	SSD	P
GCM-RSA	166.3	361.7	0.018	4
Standard GCM	166.3	354.4	0.018	3
Standard GCM with Bias	165.6	374.9	0.017	6

P = # free parameters

will tend to be classified into the "incorrect" Categories B and D with probability substantially greater than .50, whereas in the observed data they are classified into those categories with probabilities substantially less than .50. Instead, they are classified into the "region-specific attention" Categories A and C. In addition, inspection of the middle panel of Fig. 6 reveals that the standard GCM tends to substantially underpredict accuracy for many of the other transfer stimuli. Although there is a slight tendency for the GCM-RSA to also underpredict accuracy for the other transfer stimuli, it is far less pronounced than for the standard GCM. Finally, as can be seen in the bottom panel of Fig. 6, although adding response-bias parameters improves the ability of the standard GCM to account for the classification probabilities involving the critical-transfer stimuli, it comes at the expense of worse predictions for many of the remaining stimuli. Overall, the standard GCM with bias still yields a substantially worse fit to the classification data than does the GCM-RSA (see Table 2).

Inspection of the best-fitting parameters for the GCM-RSA in Table 3 indicates that, as hypothesized, in the Cat-4 condition participants gave greater attention weight to the line-position dimension ($m = 2$) in the red region, but to the rectangle-height dimension ($m = 3$) in the blue region. This region-specific switch in attention is what allows the model to capture reasonably well the overall structure of the

Table 3 Best-fitting parameters for the models in each condition

GCM-RSA

	Parameters					
Condition	c	w_1	$w_{2,red}$	$w_{3,red}$	$w_{2,blue}$	$w_{3,blue}$
Cat-4	4.405	0.390	0.474	(**0.136**)	**0.177**	(0.433)
Cat-2	4.261	0.063	0.635	(**0.302**)	**0.435**	(0.502)
Line Position	5.324	0.294	0.683	(**0.022**)	0.683	(**0.022**)
Rect. Height	6.440	0.269	**0.114**	(0.616)	**0.117**	(0.613)

Standard GCM

	Parameters							
Condition	c	w_1	w_2	w_3				
Cat-4	4.591	0.373	0.331	(0.295)				
Cat-2	4.099	0.097	0.538	(0.365)				
Line Position	5.327	0.294	0.683	(**0.023**)				
Rect. Height	6.445	0.269	**0.116**	(0.615)				

Standard GCM with Bias

	Parameters							
Condition	c	w_1	w_2	w_3	b_1	b_2	b_3	b_4
Cat-4	4.862	0.352	0.345	(0.303)	0.336	0.209	0.301	(0.153)
Cat-2	4.255	0.087	0.541	(0.372)	0.558	(0.442)	---	---
Line Position	5.396	0.299	0.661	(**0.040**)	0.347	0.270	0.255	(0.128)
Rect. Height	6.548	0.266	**0.132**	(0.602)	0.245	0.191	0.316	(0.248)

Parameter values in parentheses are not free to vary. Cell entries with dashes indicate that the parameters do not exist for that model and condition. Bold-faced cell entries are for region-specific or task-specific dimensions that have low diagnosticity for that condition. Dimension 1 = color, Dimension 2 = line position, Dimension 3 = rectangle height.
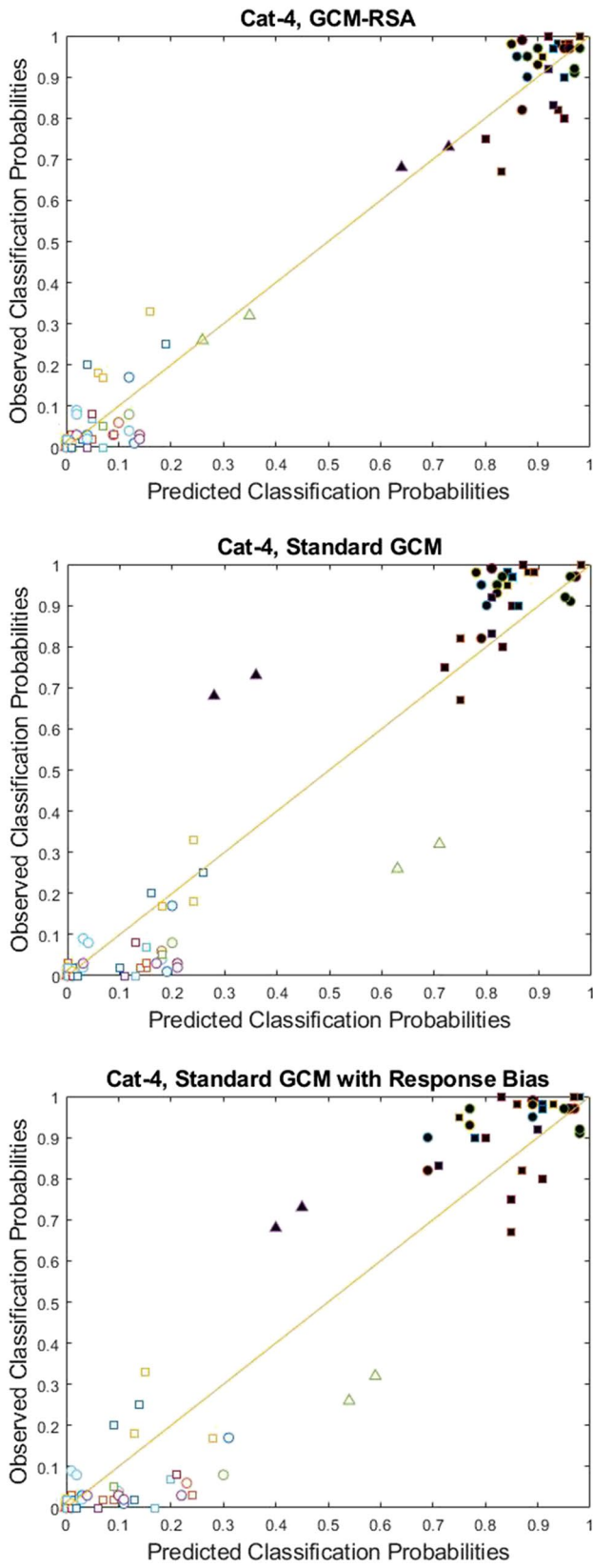
Fig. 6 Observed-against predicted classification probabilities for each model in the Cat-4 condition. Solid = correct, open = incorrect. Squares = training, circles = other-transfer, triangles = critical-transfer
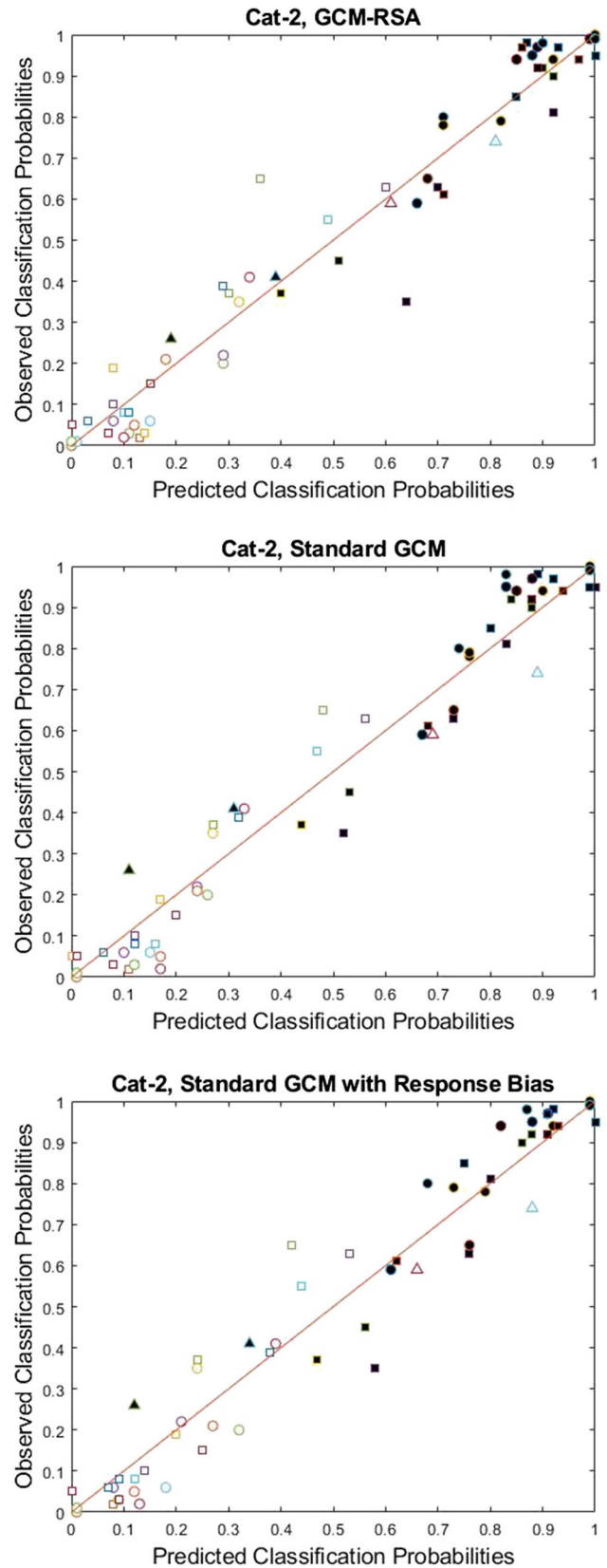
Fig. 7 Observed-against predicted classification probabilities for each model in the Cat-2 condition. Solid = correct, open = incorrect. Squares = training, circles = other-transfer, triangles = critical-transfer

complete classification-confusion matrix in the Cat-4 condition. (In interpreting the parameter estimates, the reader is reminded that the magnitude of the weight on Dimension 1 [color] cannot be meaningfully compared with the magnitude of the weights on the other dimensions because the dimensions have not been psychologically scaled.)

Cat-2 condition The model-fitting results from the Cat-2 condition are reported in the second panel of Table 2 and in the scatterplots in Fig. 7. Although these Cat-2 results show some similarities with those from the Cat-4 condition, they mainly show dramatic differences. In our view, the main take-home message is that comparison of the model scatterplots in Fig. 7 suggests that the overall quantitative fit of the three versions of the GCM is roughly the same in the Cat-2 condition. Although the BIC fit of the GCM-RSA continues to be better than that of the standard GCM and biased-GCM in the Cat-2 condition (see Table 2), the fit differences are vastly reduced compared with the Cat-4 condition. Inspection of the best-fitting parameter estimates for the GCM-RSA in the Cat-2 condition (see Table 3) indicates that, overall, subjects continued to give greater attention weight to line position than to rectangle height in the red region, but greater attention to rectangle height than to line position in the blue region. However, the magnitude of the switch in attention is much less in the Cat-2 condition than in the Cat-4 condition. Overall, the outcomes from the model-fitting analyses converge with those reported in our Results section in suggesting that the degree to which region-specific selection attention operated in the Cat-2 condition was far less than in the Cat-4 condition.

Line-position and rectangle-height conditions The model-fitting results from the LP4 and RH4 conditions are easy to summarize: Because most of the correct-classification probabilities were near ceiling in those conditions (and most of the error-classification probabilities were near floor), all models provided nearly perfect accounts of the classification data in those conditions (see Table 2). The standard GCM yields the lowest BIC because it uses the fewest free parameters, and the extra free parameters afforded the GCM-RSA are superfluous for those conditions. According to all three versions of the GCM (see Table 3), in the line-position condition, participants devoted little attention to rectangle height; whereas in the rectangle-height condition, participants devoted little attention to line position. For the GCM-RSA, these attention-weight parameter estimates were nearly invariant across the red and blue regions of the space, which is a sensible result because the relevant dimensions in each condition are no longer region dependent.

Modeling guided by cluster analysis A limitation of the modeling analyses reported in this section is that they were applied to data that were aggregated across individual observers, and the presence of individual differences may be obscured by averaging. Therefore, in an attempt to identify homogeneous subgroups, we conducted a variety of additional modeling analyses that made use of different forms of cluster analysis. In one approach, we defined idealized response vectors that would be produced by use of the optimal attention strategies in each condition (RSA in the Cat-4 and Cat-2 conditions, and task-specific attention in the LP4 and RH4 conditions). We then clustered together those participants in each condition whose empirical response vectors were close in distance to the idealized vectors. As it turned out, in the Cat-4, LP4 and RH4 conditions, there was high overlap between the upper-median-performing participants and the clusters of participants who were chosen by this clustering method. Hence, the patterns of upper-median individual-participant classification behavior were apparently relatively homogeneous in these conditions and the aggregated data were representative of the performance patterns at the individual-participant level. By contrast, in the Cat-2 condition, the clustering analysis identified a small group of participants with empirical response vectors close to the idealized RSA response vector but with the vast majority of participants having response vectors relatively far away. Recall that the best-fitting parameter estimates from the GCM-RSA in the Cat-2 condition suggested a tendency toward use of RSA, but one that was much reduced in magnitude compared with the Cat-4 condition. These parameter estimates were most likely reflecting the behavior of this small subgroup of participants in the Cat-2 condition who had learned to use the RSA strategy. Future research might attempt to identify specific strategies applied by individual participants through use of model fitting conducted at the individual-participant level (e.g., Donkin et al., 2015).

General discussion

Consideration of the structure of various real-world categories suggests that effective category learning may sometimes require observers to learn to use region-specific selective attention (RSA). In RSA, observers attend selectively to certain dimensions in some regions of a stimulus space but to other dimensions in other regions of that space. Past research has provided convincing evidence that at least some participants can learn to use RSA (as well as closely related forms of region-specific category representation).

However, the process seems considerably more cognitively demanding than learning of a single, fixed selective-attention strategy across an entire stimulus space. Thus, an important question concerns factors that may facilitate the learning of this challenging form of selective attention. In the present research, we investigated one such

factor—namely, the relation between individual categories and the regions in which they are embedded. Our hypothesis was that effective region-specific attention is more likely to be learned when individual categories are embedded in single regions rather than dispersed across multiple regions. When embedded in single regions, observers can learn a single fixed set of region-specific attention weights for each individual category; but if the individual categories are dispersed across multiple regions—each defined by differing diagnostic dimensions—then observers would need to learn multiple sets of region-specific weights for each individual category. The latter seems a considerably more challenging cognitive task than the former.

The present research provided evidence in support of our hypothesis, at least for the upper-median-performing participants in each condition. In particular, for these participants, learning of region-specific attention took place more effectively in the four-category condition (in which each category was embedded in a single region) than in the two-category condition (in which each category was dispersed across separate regions). We reported two lines of evidence supporting this conclusion. First, participants in the four-category condition achieved higher "accuracy" on the critical transfer stimuli and the other transfer stimuli than did participants in the two-category condition (where "accuracy" was defined in terms of consistency with the hypothesized region-specific selective-attention strategy). Second, in the four-category condition, the version of the exemplar model with region-specific attention weights provided dramatically better quantitative accounts of the complete matrix of classification data than did the comparison standard exemplar model (without region-specific attention weights). By comparison, the RSA-exemplar model showed only a small advantage compared with the standard model in fitting the complete matrix of classification data from the two-category condition, suggesting that the operation of RSA was greatly diminished in that condition. An additional result was that performance in both the LP4 and RH4 conditions—which did not require RSA—was dramatically better than performance in both the region-specific Cat-4 and Cat-2 conditions, suggesting that learning of RSA is indeed more cognitively demanding than learning of a single, fixed set of task-specific attention weights across an entire stimulus space.

In the present research, our representative from the class of RSA-exemplar models was highly parsimonious as it involved the addition of just a single free parameter (a region-specific attention weight) to the standard exemplar model. This analytic device was sufficient to provide convincing evidence of the operation of RSA in the Cat-4 condition. Future research, however, should examine the extent to which even more stimulus-specific forms of selective attention may be operating in these paradigms.

For example, Aha and Goldstone (1992) described the use of a trial-by-trial learning model that produced unique attention weights associated with each individual stimulus in the space. Indeed, future work is needed that can specify in more rigorous terms what constitutes a "region" of a stimulus space. Here we have proceeded with what seems to us an intuitively compelling assumption that the regions are defined in terms of the color of the stimulus (see Figs. 2 and 3), but we acknowledge that deeper and more rigorous formalizations are needed. Depending on how one defines "regions," the distinction between "region-specific" and "stimulus-specific" forms of selective attention may become blurred.

Although the present research provided evidence concerning certain experimental factors that may promote RSA, it was silent with respect to specifying the dynamic cognitive processes that underlie RSA. In particular, as an analytic device for measuring the operation of RSA, we endowed the GCM with a set of region-specific selective-attention-weight parameters and used the best-fitting parameter estimates from the model to draw inferences about the operation of RSA. By comparison, modern theories of attention in categorization instead specify dynamic processes in which category decisions unfold in time based on different patterns of information sampling (Braunlich & Love, 2022; Weichart et al., 2022). For example, Braunlich and Love's (2022) sampling emergent attention (SEA) model combines two interacting components. The first reflects an observer's beliefs about the structure of the environment, and the second estimates the value of different knowledge states that would arise by sequentially sampling alternative information sources. Within a trial in which a specific stimulus is presented, SEA tends to sample dimensions in orders that are expected to give rise to gains in utility. Thus, consider again the Blair et al. (2009) category structure shown in Table 1. Here, the value on Dimension 1 (D1) acts as an "indicator" for whether Dimension 2 (D2) or Dimension 3 (D3) is relevant and should be sampled. SEA predicts that observers will first sample D1; then, if D1 = 1, they will sample D2, whereas if D1 = 2, they will sample D3. Using eye-tracking methods, Blair et al. (2009) demonstrated precisely this predicted pattern of information sampling behavior. Analogously, for the category structure tested in the present experiment, one would expect observers to first sample color and then to sample either vertical-line location or rectangle height depending on the sampled color value. Future work might be aimed at testing this prediction involving the dynamics of sampling. Nevertheless, although the dynamic attention models developed by Braunlich and Love (2022) and by Weichart et al. (2022) provide elegant accounts of such forms of within-trial information sampling, it remains an open question whether they predict the fundamental phenomenon reported in this article: RSA was

learned more effectively when individual categories were embedded within single regions rather than dispersed across multiple regions of psychological space.

There are a number of other limitations of the present research that need to be pursued in future work. For example, we have investigated only a single stimulus type and task configuration in testing our hypothesis, so the generality of our findings remains to be demonstrated. Likewise, other factors beyond whether individual categories are embedded in single regions may also promote RSA. For example, observers may be more prone to adopt RSA in paradigms that explicitly encourage sequential information sampling strategies. Finally, although we have analyzed our data within the framework of an exemplar-similarity model, we have acknowledged that logical-rule-based models can likely provide an equally viable account of the data. To distinguish between such models, one might test alternative category structures in which rule-based horizontal or vertical decision boundaries fail to perfectly classify all the training items in the space, but in which graded use of RSA would still lead to benefits in classification performance. Through adjustment of RSA weights, exemplar models might provide reasonable accounts of such data that logical-rule-based models might struggle to handle.
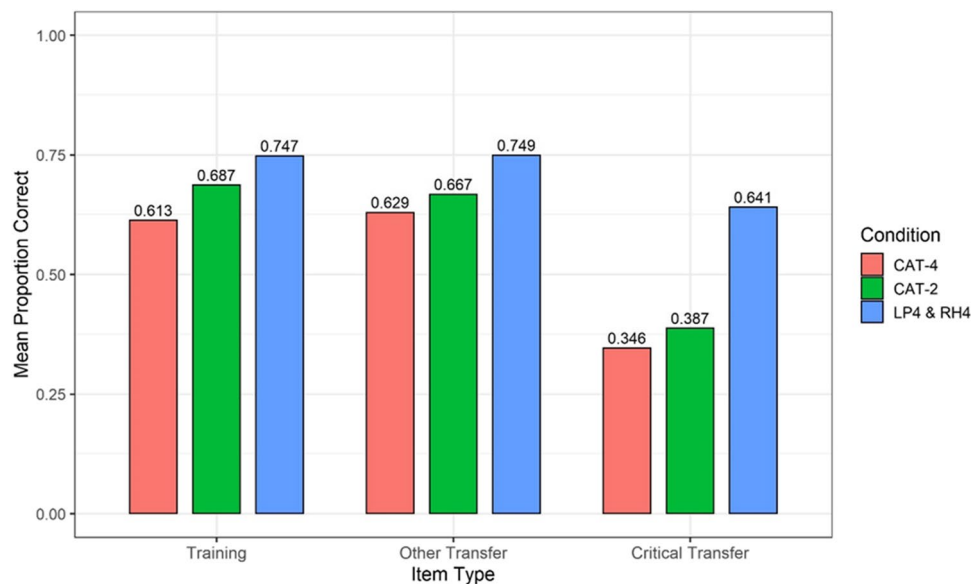
Declarations

Conflicts of interest The authors have no relevant financial or non-financial interests to disclose.

Ethic approval The study was approved by the Indiana University Institutional Review Board.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent for publication Not applicable.

References

Aha, D. W., & Goldstone, R. L. (1992). Concept learning and flexible weighting. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 534–539).

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology, 37*(3), 372–400.

Blair, M. R., Watson, M. R., Walshe, R. C., & Maj, F. (2009). Extremely selective attention: Eye-tracking studies of the

Appendix

Fig. 8 Mean proportion correct for each item type in each of the conditions (lower-median subjects)

dynamic allocation of attention to stimulus features in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1196.

Braunlich, K., & Love, B. C. (2022). Bidirectional influences of information sampling and concept learning. *Psychological Review, 129*(2), 213.

Donkin, C., Newell, B. R., Kalish, M., Dunn, J. C., & Nosofsky, R. M. (2015). Identifying strategy use in category learning tasks: A case for more diagnostic data and models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(4), 933.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127*(2), 107.

Erickson, M. A., & Kruschke, J. K. (2002). Rule-based extrapolation in perceptual categorization. *Psychonomic Bulletin & Review, 9*(1), 160–168.

Hooke, R., & Jeeves, T. A. (1961). "Direct Search" Solution of Numerical and Statistical Problems. *Journal of the ACM (JACM), 8*(2), 212–229.

Jones, M., & Love, B. C. (2006). The emergence of multiple learning systems. In: *Proceedings of the Annual Meeting of the Cognitive Science Society, 28*(28).

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*(1), 22.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45*(6), 812–863.

Little, D. R., & Lewandowsky, S. (2009). Beyond nonutilization: Irrelevant cues can gate learning in probabilistic categorization. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 530.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review, 111*(2), 309.

Marshak, S. (2019). *Earth: Portrait of a planet* (6th ed.). W. W. Norton and Company.

Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: the effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(3), 775.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39.

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin & Review, 7*(3), 375–402.

Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(5), 924.

Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology, 51*(1), 1–41.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology, 1*(1), 54–87.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237*(4820), 1317–1323.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied, 75*(13), 1.

Weichart, E. R., Galdo, M., Sloutsky, V. M., & Turner, B. M. (2022). As within, so without, as above, so below: Common mechanisms can support between-and within-trial category learning dynamics. *Psychological Review, 129.*

Yang, L. X., & Lewandowsky, S. (2003). Context-gated knowledge partitioning in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 663.

Yang, L. X., & Lewandowsky, S. (2004). Knowledge partitioning in categorization: constraints on exemplar models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(5), 1045.