Original Articles

Testing analogical rule transfer in pigeons (*Columba livia*)

Muhammad A.J. Qadri[a,*], F. Gregory Ashby[b], J. David Smith[c], Robert G. Cook[a]

[a] *Tufts University, Department of Psychology, USA*
[b] *University of California, Santa Barbara, Department of Psychological & Brain Sciences, USA*
[c] *Georgia State University, Department of Psychology, USA*

ABSTRACT

Categorization is an essential cognitive process useful for transferring knowledge from previous experience to novel situations. The mechanisms by which trained categorization behavior extends to novel stimuli, especially in animals, are insufficiently understood. To understand how pigeons learn and transfer category membership, seven pigeons were trained to classify controlled, bi-dimensional stimuli in a two-alternative forced-choice task. Following either dimensional, rule-based (RB) or information integration (II) training, tests were conducted focusing on the "analogical" extension of the learned discrimination to novel regions of the stimulus space (Casale, Roeder, & Ashby, 2012). The pigeons' results mirrored those from human and non-human primates evaluated using the same analogical task structure, training and testing: the pigeons transferred their discriminative behavior to the new extended values following RB training, but not after II training. Further experiments evaluating rule-based models and association-based models suggested the pigeons use dimensions and associations to learn the task and mediate transfer to stimuli within the novel region of the parametric stimulus space.

1. Introduction

Understanding human and non-human animals' categorization abilities has engaged researchers for decades (Cook & Smith, 2006; Ghirlanda & Enquist, 2003; Wasserman, Kiedinger, & Bhatt, 1988). Much research using complex pictorial stimuli has shown that animals can generalize their categorical knowledge to novel situations (Herrnstein, 1990; Herrnstein & Loveland, 1964). Many accounts of categorization have been considered and evaluated using artificial displays to understand this behavior in both humans and non-human animals, generating theories about cue validity, exemplar memorization and prototypes (Beach, 1964; Medin & Schaffer, 1978; Nosofsky & Johansen, 2000; Smith, Redford, & Haas, 2008).

More recently, human categorization has been posited to involve multiple processes (Ashby, Alfonso-Reese, Turken, & Waldron, 1998). The two most heavily studied are an explicit process thought to be rule-based, which operates via high-level cognitive rules or propositions, and an implicit process that uses procedural learning to operate on stimulus-response associations. Evidence for this dichotomy has been found in multiple experimental designs (Ashby & Maddox, 2005; Ashby & Valentin, 2017). The most broadly impactful studies demonstrate learning rate differences during acquisition of different tasks despite similar category structures (Smith, Beran, Crossley, Boomer, & Ashby,

2010). In this design, half of the participants are trained with a "rule-based" (RB) discrimination and the other half are trained with an "information integration" (II) discrimination. In the RB condition, both categories completely overlap along one highly variable dimension, and they are distinct when considering the other less variable dimension (as in Fig. 1, left). In the II condition, the stimuli are structured identically, except the entire stimulus set is rotated about the center of the stimulus space by 45° (as in Fig. 1, right). While this manipulation ensures that the conditions are equally difficult when considering inter-category and intra-category similarity, the RB condition reliably yields faster learning than the II condition (Ashby & Maddox, 2005).

These functional differences in otherwise matched conditions contributed to the development of computational frameworks like COVIS, a neurobiologically-informed computational framework with multiple categorization systems (Ashby et al., 1998). The explicit system in COVIS uses a rule-based learning process, in which rules generate independent decisions about one or more stimulus components to compute a strategy for response selection. Meanwhile, the implicit system in COVIS uses procedural learning to associate responses with regions of stimulus space. This multiple-systems framework and computational implementation has successfully accounted for numerous empirical results, including the learning rate difference noted above, and has successfully predicted many other qualitative differences between
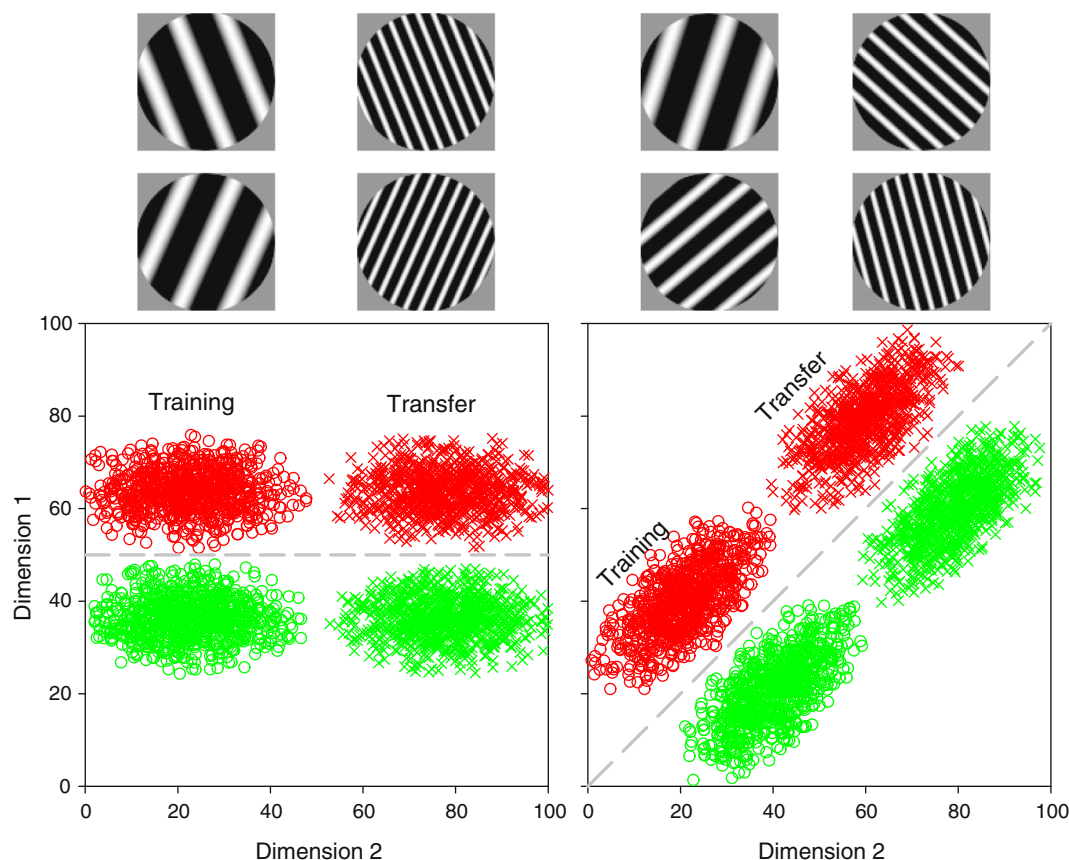
Fig. 1. Distributions of the stimuli investigated in Experiment 1. The left graph depicts one of the Rule-Based training conditions, and the right graph depicts one of the Information Integration training conditions. Sample stimuli depict category means with "Dimension 1" corresponding to orientation and "Dimension 2" to spatial frequency (see Table 1 for exact values). In these examples, the circles denote the training distributions, with the pigeons required to appropriately discriminate between the red and green distributions. The exes denote transfer test stimuli from Experiment 1, although only 72 of these points were tested for each bird. Correct category assignments for the transfer stimuli are based on the dashed gray line separating the distributions. Note that individual birds may have had these setups rotated by 90 degrees (RB) or 180 degrees (RB and II). See the text for more detail. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

learning and performance in RB and II categorization tasks (for a review, see Ashby & Valentin, 2017).

One recent experimental design using "analogical transfer" yielded results that offered further support for this multiple systems categorization theory (Casale et al., 2012). In this design, humans were trained on the traditional RB and II categorization conditions within a restricted region of the stimulus space (e.g., the "Training" distributions in Fig. 1). The observers were then tested with novel regions of the stimulus space during transfer (e.g., the "Transfer" distributions in Fig. 1; cf. McDaniel, Cahill, Robbins, & Wiener, 2014). Observers in the RB condition were able to withstand the shift in the stimulus space, and they showed little if any decrement in performance. This was deemed analogical transfer because the stimuli in the new region of the stimulus space look different from the stimuli in the original training region, but the observers were able to use the learned rule to correctly categorize stimuli. Observers in the II condition, however, showed no savings or benefit of the prior learning, and needed to re-learn the discrimination in the new region of the stimulus space. The authors theorized that the successful transfer to the novel region occurred in the RB condition because participants used rules to solve the task, and these rules were analogically extended to novel portions of the stimulus space. The II condition, however, required use of the implicit system, and consequently, its procedural learning was limited to the region of training, resulting in no savings or transfer to novel items.

The evidence for multiple categorization systems in humans raises questions regarding the evolution of these cognitive mechanisms. Categorization and discrimination of multidimensional stimuli like these have been evaluated numerous times, with clear evidence that pigeons, for example, can simultaneously attend to the conjunction of multiple features or to each feature (Lea et al., 2018; Lea & Wills, 2008; Teng, Vyazovska, & Wasserman, 2015). The implicit learning system that underlies these processes uses basic associative mechanisms that can be found across the animal kingdom. Although it is unclear whether rule-based learning uses recurrent networks, top-down control, or abstract "concepts," rule-based mechanisms seem complex in comparison to well-understood associative mechanisms. In examining categorization by non-human animals, one might expect that those species with more advanced cognitive abilities (tool use, problem solving, etc.) could potentially have both systems, while those with less advanced cognitive abilities might only possess the association-based system.

Smith et al. (2010) tested six macaques in the II and RB conditions, using the same category distributions and stimuli as humans, and the macaques showed faster learning in the RB condition. In Smith et al. (2011), however, two separate sets of pigeons learned the RB and II conditions at the same rate. These comparative results suggest that macaques may possess and use two categorization systems, like humans, while pigeons may possess a single, non-analytic, association-based learning system. Consistent with this thinking, Smith et al. (2015)

recently showed that macaques demonstrated "analogical transfer" when trained in the restricted RB condition, but not so in the II condition. However, when responses to the novel stimuli were not explicitly reinforced, macaques did not extend the apparent "rule" to the novel region (Zakrzewski, Church, & Smith, 2018), suggesting that the macaques' categorization mechanisms may not fully accord with human-like rule-based categorization. In the current investigation, we evaluate such transfer in the pigeon. After an initial replication followed by an extension to novel values, we examine different mathematical models to understand the pigeons' categorization behavior.

2. Experiment 1

First, we evaluate whether pigeons show transfer to stimuli in novel regions of the stimulus space after RB and II training. We tested for this analogical transfer in eight pigeons using the same procedural tactics as Casale et al. (2012). If the pigeons possess and employ two systems, like humans, then the pigeons in the RB condition should demonstrate transfer, while the pigeons in the II condition should exhibit no transfer. If instead the pigeons have a single, non-analytic, association-based learning system used to learn both types of conditions, as suggested by the previous Smith et al. (2011) findings, then neither group should show any analogical transfer.

We trained two groups of pigeons on a two-alternative forced-choice (2AFC) categorization task to discriminate bi-dimensional sine-wave gratings. These stimuli have been investigated in multiple species, especially when convolved with a Gaussian filter to produce Gabor patches (e.g., Jassik-Gerschenfeld & Hardy, 1979; Tappeiner et al., 2012). These stimuli were used in previous investigations with humans, non-human primates, and pigeons (Smith et al., 2011) and featured two dimensions, grating orientation and grating width. Four pigeons were trained using RB conditions and four pigeons were trained using II conditions. During training, the pigeons were only presented stimuli from a restricted region of the total bi-dimensional stimulus space, so that the remaining portion of the space could be used during novel analogical transfer testing in the same manner as previously tested with humans and macaques.

2.1. Methods

2.1.1. Participants

Eight male pigeons (*Columba livia*) were tested. The pigeons were housed and tested at 80–85% of their free-feeding weights, with ad libitum grit and water in their home cage, and they were experimentally naïve at the time of training. Prior to these experiments, they only received training to peck at a circular white signal for food reinforcement. All animal procedures were reviewed and approved by Tufts University's Internal Animal Care and Use Committee.

2.1.2. Apparatus

A touchscreen (EZ-170-WAVE-USB) operant chamber was used to present video stimuli and record peck responses. Stimuli were displayed on an LCD computer monitor (NEC LCD 1525X; 1024 × 768, 60 Hz refresh rate) situated just behind the touchscreen. Mixed grain reward was delivered via a central food hopper positioned beneath the touchscreen. A houselight in the ceiling was constantly illuminated, except during timeouts.

2.1.3. Stimuli

The stimuli in these experiments were sine-wave gratings that varied in spatial frequency and orientation, designed after those used in Smith et al. (2011). These stimuli were composed of a solid gray square with a circular aperture that contained the sine-wave grating (see Fig. 1, top). Stimuli were generated using ImageMagick (http://www.imagemagick.org). Each image was a 100 pixel × 100 pixel square.

Table 1
Distribution parameters for training and transfer distributions in Experiment 1. The values listed here indicate the means, variances, and covariance between the dimensions in the normalized (0–100) stimulus space. The resulting distributions are visualized in Fig. 1. Note that training and transfer designations here are only representative for a subset of the birds; for the remaining subjects, the data need to be rotated around the point (50, 50) by 90° or 180°.

Distribution		μ_1	μ_2	σ_1^2	σ_2^2	cov_{xy}
RB training	A	36.3	22.5	20.9	91.7	0.0
	B	63.7	22.5	20.9	91.7	0.0
RB transfer	A	36.3	77.5	20.9	91.7	0.0
	B	63.7	77.5	20.9	91.7	0.0
II training	A	20.9	40.3	56.3	56.3	36.7
	B	40.3	20.9	56.3	56.3	36.7
II transfer	A	59.7	79.2	56.3	56.3	36.7
	B	79.2	59.7	56.3	56.3	36.7

The category distributions were defined using the dimensions of spatial frequency and orientation, and they were designed after those used in Smith et al. (2015). These bivariate normal distributions were generated using MATLAB (MathWorks) with fixed mean and covariance parameters as described in Table 1, and the sampling was restricted such that the Mahalanobis distance for all points was less than 7.5. The resulting distributions are depicted in Fig. 1. The dimensions mapped onto spatial frequency with a minimum of 0.3 peaks (i.e., bars) per image (i.e. normalized 0) up to a maximum of 12.3 peaks per image (i.e. normalized 100) and orientation with a minimum of 4.4° and a maximum of 173.3°, with 0° corresponding to horizontally oriented bars and positive angles corresponding to counter-clockwise rotation. A third parameter of these types of functions is phase, which affects the position of the "bar(s)" within the image; this parameter was randomized across stimuli.

2.1.4. Procedure

Pre-training. The pigeons were trained to peck at a centrally located, white, 2.5 cm *ready signal* prior to the start of this experiment. They were then trained to peck at a sample stimulus when it appeared in return for food on a fixed-ratio schedule. Each training trial used a randomly selected stimulus from either training distribution as the sample. The FR to this sample was slowly increased to accommodate the final variable-ratio schedule. After they were pecking reliably to the sample, we began training the choice key response. After completing the FR on a trial, a single red (RGB 255,0,0) or cyan (RGB 0, 255, 255) choice alternative positioned 275 pixels to either side of the sample appeared, and one peck to this alternative resulted in food (the sample was visible during the choice). Once the pigeons were pecking reliably in all phases of the trial, discrimination training began.

Training. On every trial, a centrally-located, white, 2.5 cm *ready signal* appeared. When the pigeon pecked this signal, the signal was replaced with a sample stimulus. The sample stimuli were randomly selected stimuli from the two categories. After pecking at the sample stimulus on a variable ratio schedule that was uniformly distributed between 13 and 15 pecks, choice alternatives appeared on both sides of the sample. The red and cyan choice alternatives corresponded to the category of the stimulus. A single peck at the red choice alternative indicated that the pigeon categorized the sample as a "red" category stimulus, and a single peck at the cyan choice alternative indicated that the pigeon categorized the sample as a "cyan" stimulus (note, we depict and refer to this as "green" in the remainder of the manuscript). Each alternative appeared equally often on either side of the display. Correct choices resulted in access to mixed grain (i.e. food reward) for 2.5 s (for one subject, this was increased to 4 s), and incorrect choices resulted in an 8-s timeout during which the houselight was also turned off. A 3-s inter-trial interval then followed, and then the ready signal would appear to allow the next trial to be initiated. A correction procedure was

used such that incorrect responses resulted in the trial being re-presented until the correct response was given. Only the first trial in this sequence was considered for accuracy metrics.

Four of the pigeons were trained in the RB condition and four in the II condition. The distributions used are listed in Table 1. Half of the birds in each case were trained using the "lower" distributions (i.e., the left distributions marked "Training" in Fig. 1), where the features of interest occupy the lower portion of the total values used, and the other half were trained using the "higher" distributions (i.e., the right distributions marked "Transfer" in Fig. 1). For the II condition, only the depicted positively-correlated distributions were used, and the corresponding negatively-correlated distributions (i.e., 90° rotation from Fig. 1, Right) were not used. For a complete perspective, Supplemental Fig. S1 provides a bird-specific depiction of all the training conditions and stimuli used. Each training session contained 80 total trials (40 from each category). Training was considered completed when the pigeon achieved an accuracy of at least 80% for five sessions (non-consecutively).

Transfer. The pigeons were then given six sessions of testing with the appropriate transfer distributions (i.e. birds trained on "high" distributions were given transfer tests from the corresponding "low" distributions and vice versa). For transfer tests, six stimuli from each transfer distribution were randomly selected and interspersed within a regular session (72 total test trials, 36 for each category). All responses for these test trials resulted in food reward and no time out (i.e. non-differential reinforcement).

2.2. Results

Acquisition. Seven of the eight pigeons learned the discrimination to criterion, with no difference in acquisition rate between the RB and II training conditions. The three successful pigeons in the RB condition required 16, 26, and 50 sessions to complete training, and the four pigeons in the II condition required 24, 25, 26, and 37 sessions to complete training (for complete learning curves, see Supplemental Fig. 2). Altogether, the birds in the RB condition averaged 30.7 sessions to criterion while the birds in the II condition averaged 28 sessions, which is negligibly different ($t(5) = 0.47$, $p = .782$). Despite 100 additional sessions of training, one pigeon in the RB condition failed to reach criterion and its data are not further considered. This result replicates the prior pigeon results that showed no dramatic differences in

learning rates (Smith et al., 2011) and continues to contrast sharply with studies of nonhuman primates and humans, who learn in the RB condition more quickly than in the II condition (Ashby & Maddox, 2005; Smith et al., 2010; Smith et al., 2015).

Analogical Transfer. In the subsequent test for analogical transfer, the pigeons' novel transfer performance was related systematically to their training condition. As scored relative to the extension of their training design, the pigeons in the RB condition showed successful transfer, while the pigeons in the II condition did not. As shown in Fig. 2, the three pigeons in the RB condition performed above chance (63.9%, 63.9%, 80.6%; individual binomial tests, $ps < 0.02$). This accuracy is reduced in comparison to baseline performance for two pigeons (chi square test of independence using accuracy and phase; χ^2s $(1) > 66$, $ps < 0.001$), while the third showed no decrement between baseline and transfer accuracy ($\chi^2(1) = 2.4$, $p = 0.120$). The pigeons in the II condition did not show transfer. Three pigeons performed not significantly different from chance (41.7%, 45.8%, 47.2%; individual binomial tests $ps > 0.14$), whereas the fourth was significantly below chance (31.9%, $p = 0.001$). While the acquisition results fail to support a multiple categorization systems hypothesis, this differential transfer to the novel region of testing contradicts our intuitions and expectations of how a single system would perform.

2.3. Discussion

During analogical transfer testing, the RB pigeons exhibited transfer similar to humans although to a lesser degree (i.e., humans show essentially perfect transfer). Like humans, the pigeons who learned the RB task extended their learning beyond the training portion of the stimulus space. Further, also like humans, the pigeons who were trained with an II task were unable to extend their learning to the novel region. They showed no discrimination during analogical transfer. In the initial investigations with humans and monkeys, this type of differential success following RB training has been suggested to be an example of rule extension or analogical transfer. How should this similar transfer in pigeons then be interpreted, especially in light of the acquisition results?

Two resolutions are possible. The first resolution is to assert that birds, similar to humans and non-human primates, use a rule-based solution in the RB task and a non-analytic solution in the II task. This would allow linear transfer of the rule to novel regions of stimulus space in the RB task, but not so for the II task. It does not easily explain why three independent evaluations find no differences in learning rates for the RB and II tasks. How could dimensions be meaningful to a rule-based system during analogical transfer, but not during learning?

A second possible resolution is that RB and II tasks are learned by pigeons with a single non-analytic association-based system, as suggested by the learning-rate results. This does not account for why the two groups show different degrees of "analogical" transfer. How does a single learning system show no dimensional benefit in training, but then show it in the analogical transfer? One possibility is that this analogical transfer test is not as diagnostic of rule-based systems as previously suggested. To better understand these alternatives, Experiment 2 further evaluated how these pigeons categorized additional regions of the stimulus space by using strategic, focused evaluations of the rest of the stimulus space.

3. Experiment 2

The previous experiment left open a question about how the pigeons learned the discrimination and then how that training affected their transfer to the untrained and novel portion of the stimulus space. In order to better understand the pigeons' categorization of novel stimuli, we next examined transfer performance to smaller and more distributed clusters of novel stimuli than the larger, diffuse areas used in Experiment 1. We hypothesized this would be helpful in better understanding the nature of pigeons' categorization behavior. These
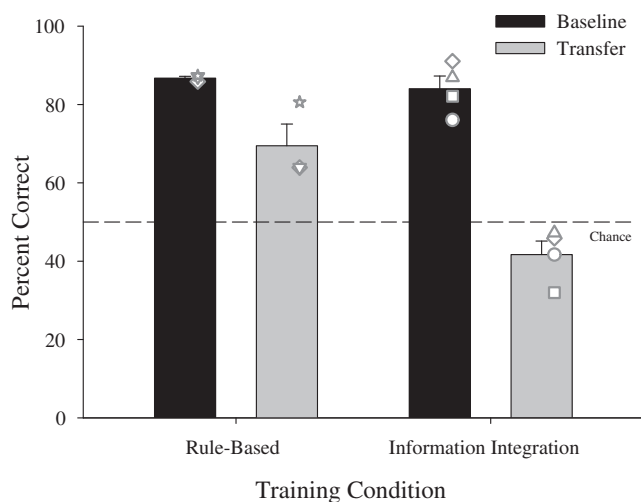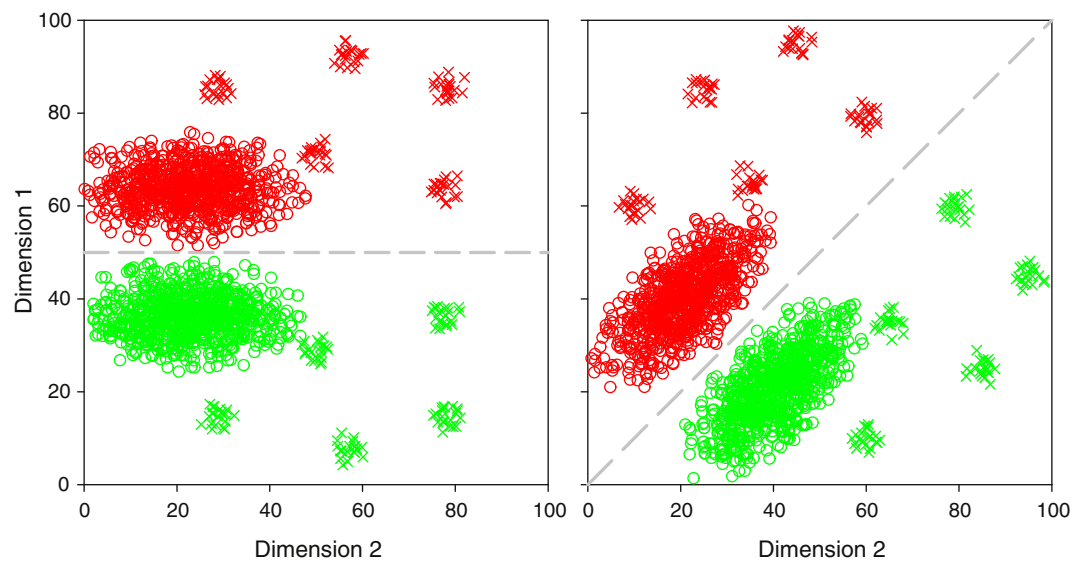

Fig. 2. Analogical transfer performance from Experiment 1. Error bars depict standard error. Individual bird data are depicted slightly offset from the bars using bird-unique symbols.

Fig. 3. Distributions of the stimuli investigated in Experiment 2. The left graph depicts one of the Rule-Based conditions, and the right graph depicts one of the Information Integration conditions. As in Fig. 1, the circles denote training stimuli, and the exes denote the transfer tests. Note that individual birds may have had these setups rotated by 90 degrees (RB) or 180 degrees (RB and II), and that category assignment for transfer stimuli are based on the dashed gray line separating the distributions. See the text for more detail.

additional testing clusters are illustrated in Fig. 3. The specific cluster areas were chosen to elucidate the pigeons' overall response patterns and to examine different possible rule-based and associative mechanisms.
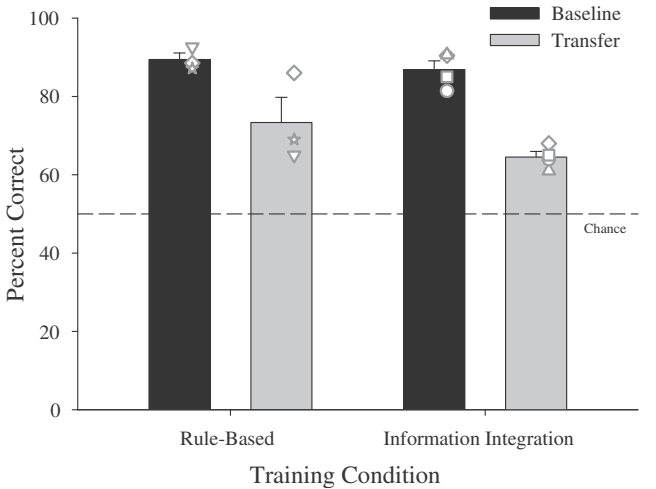
3.1. Methods

3.1.1. Participants and apparatus

The seven successful pigeons from the previous experiment were tested without an intervening break.

3.1.2. Stimuli and procedures

Stimulus values for the ten new clusters tested are listed in Table 2 and graphically depicted in Fig. 3. The standard deviations were fixed to 3.0 in both directions with zero covariance for these clusters.

Each test session contained ten randomly inserted probe trials, one from each of the ten clusters. As before, these test trials were non-differentially reinforced. Ten test sessions were conducted so that each pigeon received a total of 100 test trials, equally divided among the ten new clusters. For analysis, "correct" category assignments for these

Table 2
Distribution parameters for the transfer clusters in Experiment 2. The values listed here indicate the means in the normalized stimulus space. Note that these are only representative for a subset of the pigeons; for the remaining subjects, the data need to be rotated around the point (50, 50) by 90° or 180°.

Distribution #	Information integration		Rule based	
	μ_1	μ_2	μ_1	μ_2
1	59.7	79.2	77.5	63.7
2	79.2	59.7	77.5	36.3
3	45.0	95.0	78.3	85.4
4	95.0	45.0	78.3	14.6
5	35.0	65.0	50.0	71.2
6	65.0	35.0	50.0	28.8
7	25.0	85.0	57.1	92.4
8	85.0	25.0	57.1	7.6
9	10.0	60.0	28.8	85.4
10	60.0	10.0	28.8	14.6

Fig. 4. Analogical transfer accuracy from Experiment 2. Error bars depict standard error. Individual bird data are depicted slightly offset from the bars using the same bird-unique symbols as in Fig. 2.

transfer clusters were determined according to the extension of the linear rule dividing the training categories. Two baseline sessions separated the first test session of this experiment from the last test session of the previous experiment.
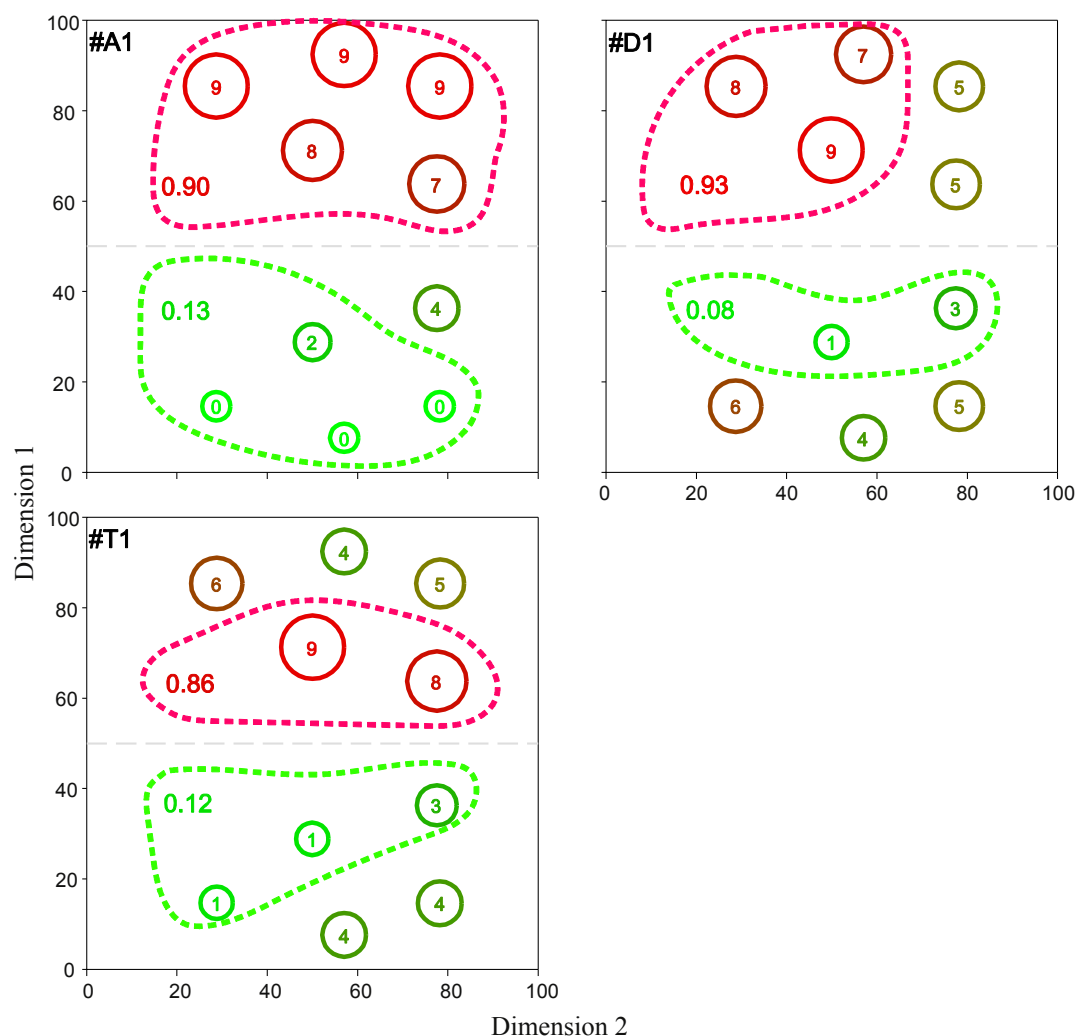
3.2. Results

These data are analyzed in two ways. First, we discuss the pigeons' overall accuracy. This is a naturally meaningful metric and relates to our analysis of Experiment 1. Second, we evaluate the patterns of the pigeons' categorization as it relates to both the location of the training clusters and their overall evaluation of stimuli from that cluster.

All pigeons demonstrated above-chance transfer of their trained discrimination to the new clusters. Fig. 4 depicts categorization accuracy for both groups across all the transfer items. Binomial tests confirmed that all seven pigeons were significantly above chance on these transfer stimuli (all accuracies ≥ 61%, $ps < 0.02$). The above-chance transfer for the RB group is consistent with the findings of Experiment
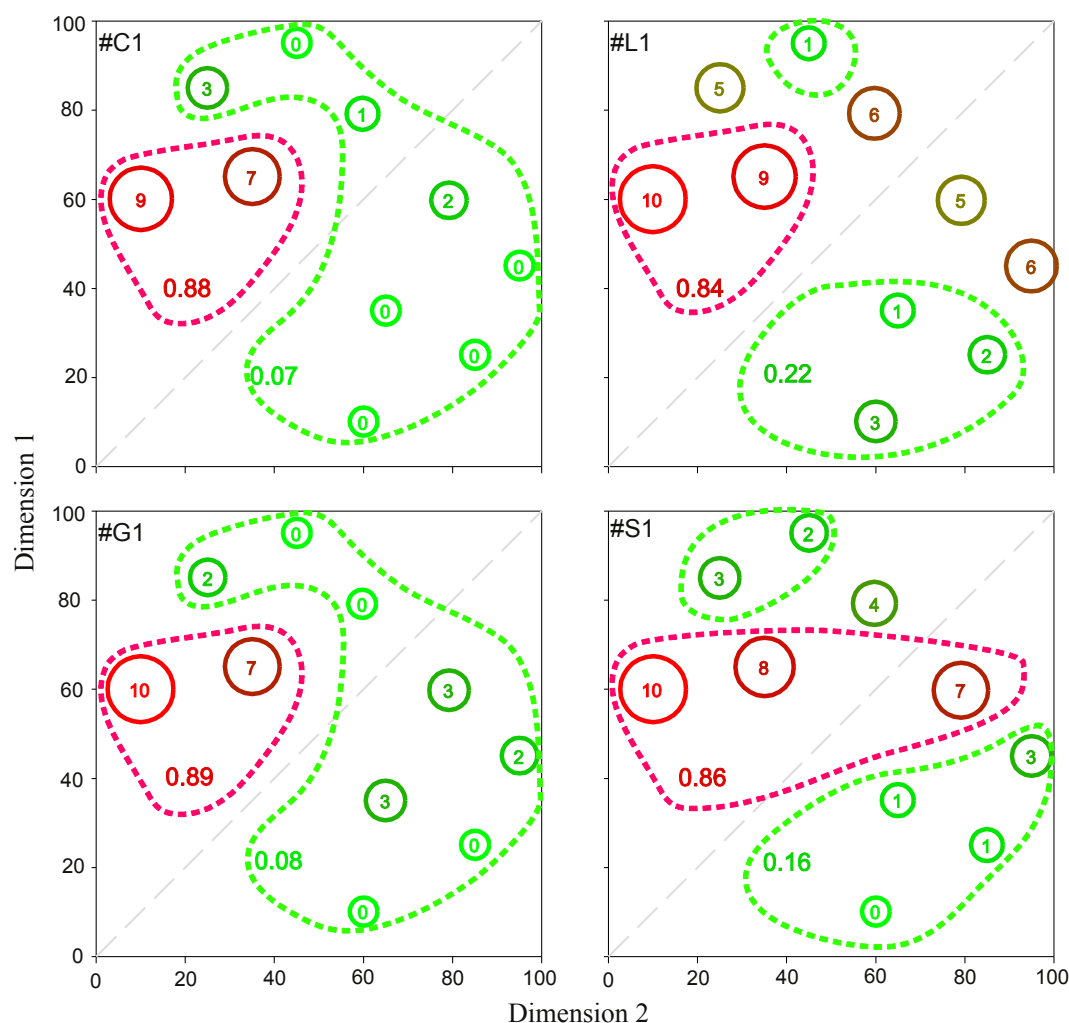
Fig. 5. Analogical transfer performance for the birds in the Rule-Based condition from Experiment 2. Values are positioned where the clusters of interest are positioned. Integer digits, color, and surrounding circle size each depict number of times out of 10 that response category "red" was selected for items from that cluster. Decimal fractions indicate the same but for the proportion of baseline trials that generated a response of category "red". The dashed gray lines indicate the space-dividing category line. The dashed red curve is a manually applied annotation of the figure to highlight the "red" category for each bird, and the dashed green curve is a manually applied annotation to highlight the "green" category for each bird. Note that some of responses and assignment have been rotated and/or flipped to provide a more understandable, uniform appearance to the task, see text in 3.2 for more detail. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

1. It is surprising that the II pigeons, who previously did not show analogical extension, were also able to produce "correct" transfer here. The pattern of categorized and non-categorized clusters reveals the source of this disparity: the location of transfer stimuli within the stimulus space controls the pigeons' performance.

The separate RB and II training conditions resulted in different patterns of transfer. Figs. 5 and 6 show the results of the transfer tests as they relate to the normalized stimulus space for the birds in the RB and II groups, respectively. The dashed light gray line is the ideal extended linear rule that divides the training categories. This is what was used to determine "correct" category assignments. For these displays, the different conditions have been normalized to facilitate inter-bird comparisons, so that all training conditions were placed in the same half of the stimulus space (i.e., pigeons #D1, #L1, #S1, and #S2 were rotated by 180° clockwise, and #A1 by 90°) and all category assignments were made similar (i.e., categories for pigeons #G1 and #L1 were inverted). These transformation make it possible to view the similarities between the birds' categorization behavior without the distraction of counterbalancing or idiosyncratic behavior. The transfer stimuli clusters are shown with the number of (normalized) "red" responses (out of 10)

placed at the cluster center. To better emphasize the patterns as they relate to the overall space, the graphs are annotated to highlight the regions where the pigeons made reliable categorization judgments (defined as greater than six or fewer than four red responses).

The data from the birds in the RB condition depicted in Fig. 5 reveals two prominent patterns. The most obvious pattern, and one that is consistent with analogical transfer, is the clear division in responding across the discriminated dimension. The second pattern concerns how responses vary according to the distance of each cluster from this rule. Pigeon #A1 (top left; 86% overall accuracy) was successfully trained on the spatial-frequency rule, and in this test of analogical transfer, categorization performance is related to the distance from the dashed gray line representing the linearly extended rule. Those transfer distributions distant from the line show the best classification, while those distributions closer to the line engender more mixed responding. This pattern is reversed for pigeons #D1 (top right, 65% overall accuracy) and #T1 (bottom left, 69% overall accuracy). They both had RB discriminations of orientation. While responding is related to the category line, the strength of their classification appears possibly inversely related to the distance from the category line, with items closer to the rule

Fig. 6. Analogical transfer performance for the birds in the Information Integration condition from Experiment 2. Values are positioned where the clusters of interest are positioned. Integer digits, color, and surrounding circle size each depict number of times out of 10 that response category "red" was selected for items from that cluster. Decimal fractions indicate the same but for the proportion of baseline trials that generated a response of category "red". The dashed gray lines indicate the space-dividing category line. The dashed red curve is a manually applied annotation of the figure to highlight the "red" category for each bird, and the dashed green curve is a manually applied annotation to highlight the "green" category for each bird. Note that some of responses and assignment have been rotated and/or flipped to provide a more understandable, uniform appearance to the task. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

supporting the best transfer performance. Thus, these two pigeons' categorization seems more constrained to the training values than #A1.

The II training yielded a different pattern of results. Here the strength of birds' categorization performance was seemingly unrelated to the dashed linear extension of their trained rule. There are again two patterns of results. On the left of Fig. 6, pigeons #C1 (68% overall accuracy) and #G1 (61% overall accuracy) both show responding that looks similar to having learned to respond to the central tendency of just one category. If a transfer item was sufficiently close to the learned "red" category, there was a fair amount of red selection. Accordingly, transfer items not in that region of space were classified as "green." Pigeons #L1 (64% overall accuracy) and #S1 (65% overall accuracy) on the right of Fig. 6 show something similar, but the annotations emphasize the fact that there are isolated clusters that challenge any simple story. For pigeon #L1 (top right), there is a clear band of ambiguity where red responding otherwise dominates, and yet at the top of the graph is a cluster where nine of 10 responses were green, situated beyond the trained "red" distribution and opposite the core "green" distribution. This suggests high confidence that the stimuli from this region were representatives of the "green" category. For pigeon #S1 (bottom right), the rightmost cluster in the annotated region is clearly more distant from the "red" training distribution than the clusters above

the annotated red region, but the clusters above the region received primarily green responses. Some hints of this same pattern may exist in the other two II pigeons but are less clear-cut. Overall, the locations of the transfer stimuli allows for "above-chance" responding, but none of the four II pigeons respond according to a simple extended linear rule from their training. The two groups of pigeons seem to approach their problems in different ways.

3.3. Discussion

The results revealed that all seven pigeons showed systematic behavior to the widely-spaced novel transfer stimuli. The birds in the RB task demonstrated an extension-like behavior from their learned rule, while the birds in the II task did not. Superficially, this pattern is consistent with the differences seen in human and non-human primate tests, which suggests that the pigeons in the two tasks are solving their discriminations in different ways. Perhaps the birds in the RB task were using a rule, while the pigeons in the II task were using simpler associative mechanisms. This conclusion, however, diverges from two sets of acquisition results, which did not find differences in acquisition rates. Such conflicting conclusions require resolution.

Configural Activation Model | Dimensional Activation Model
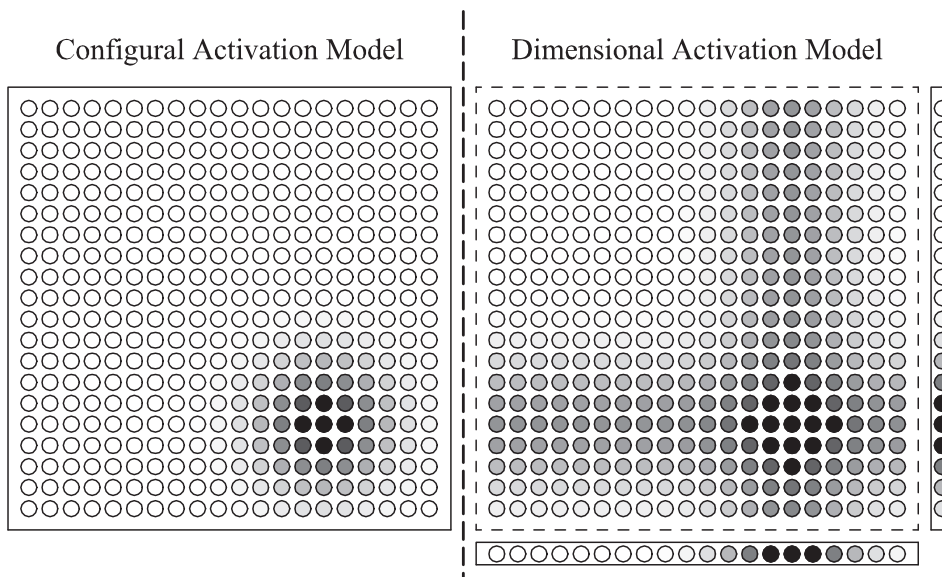
Fig. 7. A depiction of stimulus representation for the configural activation model and the dimensional activation model, using a 20-unit based grid. Each unit is represented by a circle, with highly activated units filled with black, less activated units filled with grey, and inactivated units filled with white. The configural activation model uses a radial basis function to compute the activation of units, while the dimensional activation model uses a Gaussian decay. Both representations are depicting the activation from the same external stimulus. Note that the grid on the right with the dashed boundary is for comparison purposes only – it does not accurately reflect the stimulus representation. The dimensional activation model is best represented by the two lines of units to the right and below the grid on the right.

A hint of consistency comes from a closer examination of the pigeons' results from Experiment 2. Only one pigeon (#A1; in the RB task) demonstrated transfer in a manner consistent with traditional notions of rule-based responding. The other two pigeons in the RB training condition demonstrated better performance for stimuli closer to the hypothetical rule-boundary than farther away. Thus despite the stimuli being *more discriminable* from a rule-based account (i.e., located further from the rule-boundary), these two pigeons showed *poorer* transfer. The four pigeons in the II condition agreed on the categorization of six of the ten transfer clusters. Five of these six agreed-upon clusters were relatively close to the training distributions, but the sixth (the topmost cluster in Fig. 6) was on the other side of the stimulus space from its categorized training distribution. Furthermore, if we consider the two clusters in this experiment that are most similar to the transfer distributions from the previous experiment (top right clusters in Fig. 6), the II birds' responding is either completely biased towards one stimulus or fairly non-discriminate with a slight bias towards a "reversal" of responding (as in Experiment 1). In order to resolve these numerous oddities, we evaluated the pigeons' behavior using several mathematical models with a variety of assumptions to determine if there was a simple and concise explanation of these patterns of transfer.

4. Model fitting

4.1. Overview

Mathematical models of categorization evaluate the likelihood of observing the data as determined by different categorization mechanisms. We considered both parametric and nonparametric classes of models. Parametric classifiers make strong assumptions about the form of the contrasting categories, while nonparametric classifiers make few to no assumptions about category structure (Ashby & Alfonso-Reese, 1995). Prototype models, for example, are parametric models that assume that a category varies around a singular central or average representation. Implementations of this model store only a single prototype that represents the relevant category information. Exemplar models on the other hand are non-parametric in nature, and assume that a memory of every exemplar is stored, and thus, there is no way to describe a category as involving a condensed representation. Following brief explanations of each model type, we discuss the model-fitting used here and then present and discuss the outcomes (more formal treatments of these models are in Appendix A).

4.1.1. Parametric models

There are many parametric models of categorization (Ashby, 1992b), but we will focus on the broad class that derives from general recognition theory (Ashby & Soto, 2015; Ashby & Townsend, 1986; Maddox & Ashby, 1993). This class of models assumes that the observer perceives the stimulus in a dimensionalized perceptual space. They also assume that every point in that perceptual space has some category membership value that informs responding. For example, in a two-category task, some points are associated with one category and some with the contrasting category. The categorization "boundary" is then the set of points separating these two response regions. Stimuli that fall on the boundary are thus maximally uncertain since both responses are equiprobable. This decision boundary can follow a variety of functional forms (e.g., linear or quadratic). Consequently, a parametric model will fit best when the boundary that best separates the categories has the same functional form as the observer's classification rule.

For our purposes, we will evaluate a flexible type of parametric classifier, the general quadratic classifier (GQC), which assumes the category boundaries can be hyperbolic, parabolic, elliptical, or linear. The GQC also includes the prototype model as a special case because prototype models predict that the category boundary is the line of points that are equi-distant from the two category prototypes (Ashby & Gott, 1988). As a result, this includes all models that assume the underlying categories can be represented as multivariate normal distributions.

4.1.2. Nonparametric models

Nonparametric models make much weaker assumptions about the underlying category structure. As a result, the decision boundaries they predict can take on almost any form (e.g., as in exemplar theory). We first consider the striatal pattern classifier (SPC; Ashby & Waldron, 1999), which is the procedural-learning component of the COVIS (Ashby, Paul, & Maddox, 2011) and SPEED (Ashby, Ennis, & Spiering, 2007) models of categorization. The SPC is nonparametric because it can reproduce any (piece-wise linear) decision boundary through the use of numerous decision units.

The SPC model uses a grid of neural network units to represent the perceptual space. When a stimulus is presented, a region of the units in that part of the perceptual space is activated according to a radial basis function (see Fig. 7, left, for a visual depiction of this activation). This grid of units represents a configural activation of features, similar to the configural representation used in other models of animal cognition (George & Pearce, 2012; Pearce, 2002). This entire grid of units is

Table 3
Model fitting results for the pigeon data from sessions containing transfer data in Experiments 1 and 2. The values indicate Akaike Information Criterion (AIC). Definitions of the models and the source of AIC are in the text. The model results are displayed and ranked from 1 (best, top) to 4 (worst, bottom). General Quadratic Classifier (gqc), Dimensional Activation (dim.), Configural Activation (config.), Prototype (prot.).

	Rule-based			Information integration			
	#A1	#D1	#T1	#C1	#G1	#L1	#S1
1st (best)	config. 1095.6	config. 1007.8	config. 1117.3	gqc 918.6	gqc 979.9	dim. 1510.7	dim. 1301.0
2nd	gqc 1098.2	dim. 1032.9	dim. 1128.3	dim. 937.0	dim. 991.9	config. 1518.4	config. 1307.9
3rd	dim. 1128.0	gqc 1123.6	gqc 1241.0	config. 964.5	config. 1002.6	gqc 1520.4	gqc 1314.4
4th	prot. 1151.1	prot. 1288.9	prot. 1331.3	prot. 1034.4	prot. 1251.5	prot. 1594.9	prot. 1485.6

connected to a much smaller set of decision units (maybe even just one unit) whose activity level generates the category response. Over multiple presentations, as a result of feedback-based learning, the weights are adjusted so that ultimately the correct response is produced when the stimuli are presented.

Given the importance of dimensional attention in these experiments, we also examined an alternative version of the SPC that uses a different type of stimulus representation. Specifically, this alternative uses a separate set of units *for each feature* in the stimulus to represent activation in just a single dimension. Thus, instead of a grid of units, we end up with two lines of units, with each unit representing different values along a dimension (see Fig. 7, right). When a stimulus is presented, it generates activity in each dimension. Learning and categorization then proceed as in the SPC. In particular, the two lines of units are connected to a much smaller set of decision units (maybe even just one unit), and through feedback-based learning the weights are adjusted to generate the correct response. To disambiguate the two SPC versions, we will refer to the latter model as the *dimensional-activation* model and the former (original SPC) as the *configural-activation* model in order to identify their critical difference. One illustration of the difference between the two models is shown in Fig. 7. Note that while the units are displayed in the context of the grid of stimuli used in the configural-activation model, this is only to emphasize the differences in activation patterns in the model and not to suggest that there are configural units that become activated in this fashion. If these dimensional units representing the (unbound) activation within each dimension were later bound together for a different task, how their activation would pattern is an open question.

4.2. Model fitting methods

In total, we evaluated how well two parametric and two nonparametric models fit the data from transfer sessions in Experiments 1 and 2 (for an evaluation of model fits from the end of acquisition, see Supplementary Material). The parametric models were the prototype model and the GQC. The non-parametric models were the configural-activation SPC model and the dimensional-activation SPC model. For each model and each bird's individual results, we found parameter estimates that minimized the value of the Akaike Information Criterion (AIC), a common metric used to determine minimally-complex, best-fitting models (see Appendix B for more detail). The AIC is defined as AIC = $2k - 2LL$, where k is the number of free parameters and LL is the log likelihood. So, for example, the GQC has 6 parameters and the

prototype model has 3 parameters, so the GQC must compensate for its extra parameters by providing a higher value of LL. The k component of AIC therefore penalizes complexity, while the LL component of AIC penalizes poor fits to the data. The best model is the one with the lowest AIC (i.e., closest to negative infinity), with variations that can be difficult to judge as meaningful. The rule of thumb that has been widely adopted is that a difference of 2 is considered "meaningful" or "significant" (not statistically; see Burnham & Anderson, 2004). We report AIC for each model and animal.

4.3. Model predictions

A key test of the parametric and nonparametric models is in how they would classify new stimuli outside the trained region, just like those tested in these analogical transfer tests conducted above. The GQC posits that the response to these stimuli would be based on how far they are from the decision boundary regardless of their distance to the original training distributions. Stimuli far from the boundary should elicit clear classification, and stimuli close to the boundary should be near chance. In contrast, the configural-activation SPC model predicts that confidence should decrease with the distance (i.e., bound, two-dimensional distance) from the trained distributions. Transfer stimuli that fall in regions near the training distributions will mostly activate trained units, so the model's response will be systematic and correct. For stimuli that fall in regions further from the training distributions, the weights that connect those units to the decision layer were never trained or modulated to any serious degree, resulting in arbitrary (although potentially not-chance-level) responding. Thus, for the space between the two category distributions, the GQC and the SPC models would predict similar results, but in the regions outside the two training distributions, their predictions clearly diverge.

How the analogical transfer results would vary between the dimensional-activation and the configural-activation SPC models is not self-evident. Examining Fig. 7 supports the intuition that the models would agree on how to classify stimuli from the region between the training distributions. Important differences emerge, however, with increased distance from the training distributions. In the dimensional-activation model, the within-dimension distances (i.e., unbound, one-dimensional distances) from the trained values would control responding. This means that outside of the training region, while using the same parameter values, responding could be quite different from that predicted by the configural-activation model. Thus, this dimensional-activation model could readily account for some of the "aberrations" found in Experiment 2. The "distant" cluster that was readily and systematically categorized by the pigeons was only distant from the training distributions along a single dimension while remaining relatively close in the second dimension. These two models predict different outcomes as a result: the configural-activation model will likely falter (i.e., predict chance-level behavior or behavior like the nearest category) while the dimensional-activation model predicts more systematic responding tied to the dimensions.

4.4. Model results

Table 3 shows the results of the different models. It reveals that the nonparametric activation-based models were the best fit to the birds' general results, as the configural-activation and the dimensional-activation models took the first two spots in 11 of 14 cases (each model's average rank 1.86 of 4). The most serious parametric competitor was the GQC. It was the best model in two cases and the second best in one case (average rank 2.29 of 4). The prototype model provided the poorest fit in all birds.

The results from the birds trained in the RB task were best fit with the configural activation model, as found for all three pigeons, #A1, #D1, and #T1. The dimensional activation model was second best for two pigeons #D1 and #T1, while the general quadratic model was second for pigeon #A1. The pigeons trained with the II task showed less consistency. Pigeons #C1 and #G1 were best fit by a general quadratic model, with the dimensional activation model being second best. In contrast, pigeons #L1 and #S1 were best fit by the dimensional activation model, and the configural activation model was second best.

4.5. Discussion

The nonparametric dimensional-activation and configural-activation models provided the best descriptions of the pigeons' behavior in both tasks. Between these two models, the dimensional-activation model generally described the pigeons' behavior better than the configural-activation model for the four pigeons in the II task. This model readily accounts for the differential transfer effects from Experiments 1 and 2, including the seemingly counterintuitive partial-reversals observed in the set of widely-spaced transfer tests. The configural-activation model could fit the results about as well, especially in the case for the three birds in the rule-based task. Between the experimental results and model fits, the pigeons appear to be using some form of a dimensionally-oriented, associative mechanism to solve the task.

These outcomes tentatively bring the results of Experiments 1 and 2 to a sharper and more coherent resolution. As supported by the rates of learning during acquisition, both RB and II groups of pigeons seem to solve this task using the same classification mechanism, one based around associative processing. The differential analogical transfer results for these two groups initially created a problem for this unified account. The inclusion of dimensional activation to an associative approach helped to reveal a resolution for this tension. This generalization mechanism extended learning well beyond the trained region of space. The unusual reversal observed during the II transfer in the widely-spaced test regions offers revealing and confirmatory evidence of this type of dimensional activation and allows it to explain the pigeons' results so far.

If the dimensional activation model truly has the most merit, then the current results raise an apparent anomaly. If the pigeons' representation of these stimuli is fundamentally dimensional in nature, with potentially few configural units to permit joint representation, then why does the pigeons' learning of the two tasks proceed at roughly equivalent rates? Would not dimensional representations yield an advantage when learning the RB task in comparison to the II task? Simulations of the dimensional activation model learning suggest that the difference is so minute as to be nearly undetectable. In one analysis, training 10,000 simulated networks on each of the RB and II tasks to criterion as in Experiment 1 produced an average of 60.6 ($SD = 18.4$) "sessions" for the RB task and 61.5 ($SD = 19.0$) "sessions" for the II task. A Wilcoxon signed rank test was able to detect this difference ($z = -4.1$, $p < 0.001$), though the effect size is minute ($d = 0.05$). Thus, assuming a dimensional representation instead of the extant configural representation model comes at little explanatory cost. The model succinctly captures the results in these experiments and does not contradict this or previous low-power acquisition results (cf. Smith et al., 2011).

Given the explanatory power and success of the dimensional activation model, the apparent success of the configural activation model is unexpected. If the activation were restrained to a region around the stimulus percept, how does the model allow for transfer to novel regions of the stimulus space? An examination of the best-fitting parameters for the configural activation models reveals the key: the variance (i.e., spread of the radial basis function) in one dimension is two-to-eight times larger than the variance in the other dimension.

Thus, by tweaking the definition of "around the stimulus," this version of the configural activation model can create the partial transfer observed in Experiment 1 and the tripartite response pattern in Experiment 2. This is not to say that the variances are much more aligned in the dimensional activation model (in fact, there is no clear relationship in the variance ratios for the two model fits), although whether a less relevant or ignored dimension should have broadly or narrowly tuned variances is somewhat unclear. However, training simulations with this imbalance contradict the acquisition results. In simulations with the dimensional variances at a 2.5 ratio, if the skew is greater in the irrelevant dimension, the discrimination is learned much more quickly in the RB condition than the II condition ($z = 15$, $p < 0.001$), and if the skew is in the relevant dimension, the neural network struggles to learn even once in the RB condition. Therefore, the apparent success of the configural activation model should perhaps be taken with a modicum of caution.

These experimental results and the subsequent models should also be considered in light of the stimuli employed. For example, the orientation dimension of the sine-wave stimulus is periodic, meaning that the most extreme values (i.e., 4° and 173°) are not as extremely different as the stimulus space implies. Perhaps, then, the more robust discrimination generalization in Experiments 1 and 2 are the result of the periodic nature of the stimulus. While this fact can possibly explain some of the dramatic effects seen in the information integration condition, it is not a perfect explanation. The successful transfer effects seen in Experiment 1, for example, were more robust when orientation was the relevant discrimination dimension. In Experiment 2, the periodicity only explains the transfer patterns in one dimension, not both. Future attempts to fit the configural and dimensional activation models may need to account for this non-linearity or use dimensions that avoid this periodicity.

Another important factor is that sine-wave grating stimuli are known to be "separable" dimensions, in which attention to a particular dimension is unaffected by the other available dimension(s). In contrast, "integral" dimensions seem to be bound together in their processing, such that, for example, the hue of a color patch cannot be processed without also processing its intensity (Garner & Felfoldy, 1970). In models of categorization, the use of a Euclidean distance metric, like that used in the configural activation model here, would imply that the stimuli are integral and not separable, and in that vein, perhaps the distance method known as "city-block" would be more appropriate for separable stimuli. Given how the networks function, however, the relative fit of the dimensional versus configural activation models may better address the separability versus integrality as compared to the particular distance metric used. Nevertheless, determining how different distance metrics alter the configural activation model's fit merits further investigation.

5. General discussion

The current experiments generated two critical empirical results. First, we documented differential transfer to novel regions of the stimulus space conditional on the pigeons' RB or II training. The pigeons in the RB task demonstrated transfer in Experiment 1 similar to previous human results, inconsistent with the prior associative account and suggestive of a rule-based categorization mechanism. This apparent contradiction required resolution. This yielded our second critical result, a partially "reversed" transfer to widely-spaced stimuli from pigeons in the II condition. Our modeling suggests systematic, dimensional, associative generalization accounts for both transfer patterns. Furthermore, the associative dimensional model we developed is potentially well-matched by a configural model that utilizes dimensionally-unequal generalization. We altogether conclude that these dimensions are salient, independent, and meaningful in the pigeons' categorization.

The use of differential analogical transfer as evidence of separate categorization mechanisms consequently needs to be revised. Casale et al. (2012) first demonstrated this operation in humans in several contexts, the most dramatic of which is comparable to the method used here. Non-reinforced testing of the novel region in their Experiment 3 demonstrated that the humans showed almost no decrement with the novel stimuli, while II performance crashed. In contrast, Smith et al. (2015) used differential reinforcement during transfer, and demonstrated similarity between human and macaque performance. While perhaps the degree of transfer during non-reinforced tests indicates the operation of multiple categorization systems, the apparent differential transfer with properly reinforced testing could have masked the generalization of an associative mechanism. Subsequent tests with macaques using this method of non-reinforced transfer yielded contradictory results from the first investigation, suggesting that the macaques' potential rule-use was not as generalizable as the humans' (Zakrzewski et al., 2018). Specifically, the authors note that there was a decrement during transfer, although they make no comment on the monkeys' above-chance performance. It remains an open question if these associative mechanisms can account for those results, but that data has similarities to our pigeons' outcomes. Thus, differential analogical transfer in this paradigm may be indicative of rule use only when the testing is conducted under uninformative (i.e., nondifferential) reinforcement conditions.

In order to account for the current data, any successful model of categorization needs the capacity to divide the stimulus space into three distinct areas after being trained on only the two areas of the basic categorical task. Most models can divide and associate these two areas effectively. The third clearly associated area, however, is in an untrained region of the stimulus space, and critically, the categorization systems of pigeons (and according to limited preliminary testing, some humans) associate it with the more distant training distribution. The hyperbolic version of the general quadratic classifier was able to account for these results by employing the under-utilized second branch of the hyperbola to effectively create three regions in the stimulus space. Traditional configuration-based associative models may suffice, if the perceptual variance of the dimensions are highly unequal; however, this does not provide a parsimonious account of the pigeons' acquisition results. We found that an associative model that used independent dimensions intuitively and accurately accounts for the whole of the pigeons' behavior.

The success of the dimensional activation model raises questions about the representation underlying perception and categorization. Numerous comparative studies have used multidimensional stimuli, many of which suggest that animals can "analytically" access the underlying dimensional features (Blough, 1972). A series of experiments on attentional tradeoffs in multidimensional displays shows not only attention to the features but attenuation in that attention over time (Teng et al., 2015; Vyazovska, Teng, & Wasserman, 2014). Burgering, ten Cate, and Vroomen (2018) report how zebra finch categorize novel auditory stimuli in a partially analytic manner, which could be consistent with a dimensional representation. Similarly, Wills et al. (2009) report experiments with pigeons and squirrels that suggest they attend to a single diagnostic feature when multiple features are available. Numerous reports, however, contrastingly suggest that animal behavior is guided by the complete stimulus configuration (Lea et al., 2018; Smith et al., 2011). In line with this, one of the more influential models of animal perception successfully posits configural representations, which is consistent with the configural activation model (Pearce, 2002). The current results provide evidence that suggests the pigeons' categorization behavior operates on a dimensional basis, belying the configural representation and suggesting that the dimensions or features are represented separately. In a dimensionalized representation, feature integration in the II task could derive from the processing of these separate representations (Ashby & Townsend, 1986). Higher-level processes may still utilize configural features, but models without dimensional access or control would be incomplete. If further testing with humans reveals parallel function and behavior with the pigeons, it would suggest that the pigeons' singular categorization system and humans' procedural learning system operate similarly.

The potential discovery of a dimensional association system raises interesting questions about the analytic, rule-based system. Dimensional access was previously a distinctive feature of the latter system, achieved by decomposing a configuration into its parts (Smith et al., 2012). However, the current research shows that birds' procedural system could use dimensional representations that are not inherently configural or bound (see also Burgering et al., 2018). Consequently, additional distinctions should be considered between the two systems. Perhaps the rule-based system can process "rules" more effectively in a hypothesis- or model-testing fashion. In this case, the system may be able to consider and evaluate outcomes with respect to multiple hypotheses simultaneously, resulting in faster and more robust learning. This is consistent with the recurrent processing style of models that implement rule-based discrimination (Ashby et al., 2011). Alternatively, the rule-based system may have access to different, more expansive dimensional representations than the association-based system. Perhaps, the activation of a value of 60° may simultaneously activate the representations "less than 75°," "less than 80°," etc., and "more than 45°," "more than 40°," etc. Utilizing such comparison relations could potentially allow for the sort of robust rule learning that underlies the analogical transfer in this task. Yet again, perhaps the rule-based system has the ability to use attention to change the salience of irrelevant dimensions to zero. By applying this sort of attentional hyper-modulation, rule-based discriminations would be acquired more quickly and information-integration discriminations would not. These possibilities need to be considered in the scope of larger datasets and models.

Finally, the differences found between the strategies employed during RB and II tasks could be attentional in nature. Attentional strategies during perceptual categorization would require an organism to possess neural structures with the ability to modulate incoming sensory signals or bottom up processes in order to alter behavior. In primates, COVIS assumes that the rule-learning system has access to a form of attention mediated by the prefrontal cortex (Ashby et al., 1998, 2007; Nomura et al., 2007), though Posner and Petersen (1990) attribute these effects to an anterior cingulate-based attentional system. In the pigeon, modulatory attentional structures have been difficult to identify, but there is some recent evidence that nidopallium caudolaterale processing relates to human prefrontal cortex processing (Lengersdorf, Pusch, Güntürkün, & Stüttgen, 2014). Pigeons may therefore have the capacity for similar two-system categorization processing of these stimuli, but these procedures may not tap into those cognitive abilities. Alternatively, the strength of modulation by the nidopallium may not rival the strength of modulation in the primate structures. Some research suggests that configural perceptual mechanisms effectively prohibit this type of visual attentional modulation in pigeons (Pearce, Esber, George, & Haselgrove, 2008). Further knowledge of the neural structures and behavioral abilities involved, in pigeons especially, may help identify the critical difference between the procedural learning system that is common to both pigeons and humans and the rule-based system that appears to be absent in these birds.

Acknowledgments

Appendix A. Mathematical formalizations

A.1. Parametric models

The prototype model and the GQC both state that a stimulus (S) can be represented in a (veridical) stimulus space and an internal perceptual space. We assume two categories, A and B, which would correspond to the red-correct and green-correct stimuli from Experiments 1 and 2. Specific stimulus i has a two-dimensional stimulus representation and will be designated $S_i = (x_{1i}, x_{2i})$. Its representation in perceptual space is a fair approximation of the veridical stimulus, plus some noise (ε) from sensors: $P(S_i) = P_i = (p_1, p_2) = (x_{1i} + \varepsilon_{1i}, x_{2i} + \varepsilon_{2i})$. Given research on generalization as well as examining neural decay functions, we will assume that the noise is Gaussian distributed with zero mean and uncertain variance (Ghirlanda & Enquist, 2003). Points within the perceptual space can be considered a distance apart, designated as D (P_1, P_2), where D functions appropriately for the dimensions involved (i.e., city-block distance or Euclidean distance). In the perceptual space, there is a function that partitions the space into distinct "A" and "B" regions, which we will designate F(P). Prototype theory also posits prototypes E_A and/or E_B, which are points in perceptual space that represent the categories, and the function F(P) is defined as F(P) = D $(P,E_A)/D(P,E_B) - 1$. If F(P) evaluates to a negative value, response A is produced, and if F(P) evaluates to a positive value, response B is produced. The decision boundary is the set of points for which F(P) = 0. In the prototype model, this is a line with the property that every point on the line is equidistant from the two prototypes (i.e., the line that bisects and is orthogonal to the line that passes through the two prototypes).

The GQC postulates that $F(P) = \beta_5 p_1^2 + \beta_4 p_1 p_2 + \beta_3 p_2^2 + \beta_2 p_1 + \beta_1 p_2 + \beta_0 > 0$. The categorization decisions function the same way, responding A if F(P) > 0 and B if F(P) < 0. For more details on parametric models, see Ashby (1992a).

A.2. Nonparametric models

In some regards, the nonparametric models do not need a further formalization due to their development by neural network generation. However, simulations show that the configural-activation and dimensional-activation models can result in learning of the basic task and that they result in very different patterns of generalization to other parts of the space, even with only one unit in the hidden layer. Because only one unit is needed in the hidden layer, if we wanted to describe the final neural network, we would not need to define the connection weight between each input node and the hidden layer explicitly. Instead, we only need to consider the parameters that define the input surface. We start again with stimulus S_i, but now its representation in perceptual space is no longer provided by a simple formula. Instead, the perceptual representation of a stimulus is the activation it generates in the neural network units of the model. If we assume Euclidean distance for the distance metric and Gaussian decay for all units' sensitivity, the activity levels for each of these activated units can be described using Gaussian distributions. In the configural activation model, each configural unit W_{ab} has optimal responding to dimension 1 value a and dimension 2 value b. Its activation in response to S_i would be computed by $A_{ab}(S_i) = \varphi(D(S_i, [a,b]), 0, \sigma_{AB}^2)$, where $\varphi(X, \mu, \sigma^2)$ is a Gaussian function with mean μ and variance σ^2 evaluated at X. Contrastingly, in the dimensional activation model, each dimension unit U_a or U_b is sensitive only to the values within their dimension. Their activations for $S_i = (x_1, x_2)$ are given by $A_a(S_i) = \varphi(D(x_1, a), 0, \sigma_A^2)$ and $A_b(S_i) = \varphi(D(x_2, b), 0, \sigma_B^2)$; if we wanted to continue representing this in the grid space of the COVIS model (i.e., as in Fig. 7), one combination method would provide us with $A_{ab}(S_i) = \varphi(D(x_1, a), 0, \sigma_A^2) + \varphi(D(x_2, b), 0, \sigma_B^2)$. The output of the activations from all units would map onto a category unit that converts the unbounded activations into a limited response representation. This unit will be activated according to a cumulative distribution function of the standard normal Gaussian (Φ), and the evaluation of this will correspond to the likelihood of emitting (e.g.) a "red" response.

Critically for the purposes of this article, because of the assumptions of our model and the functional simplicity of our distributions, we can model the neural network outcome without having to instantiate the hundreds of connections posited by the networks. With a single unit in the hidden layer, which is as functionally useful as no hidden layer at all, what the neural networks reduce to are simple association networks. Output activations in simple association networks should be proportional to the relative predicatability of the input units. For the networks provided here, the predictability of a given unit depends on the relative activation by stimuli from one category versus the stimuli from the other category. We assume that stimuli from both categories activate units using a Gaussian decay function. The categories and stimuli are distributed according to a normal (i.e., Gaussian) distribution. Given stimuli distributed normally and Gaussian decay functions, the overall activation by all the stimuli of a single category will also be Gaussian. The mean of the activation distribution will be the perceived mean of the stimuli (assuming an unbiased perception, equal to the true mean of the distribution). The variance of the activation distribution will be composed of the true variance in the stimulus category as well as any added variance from the perceptual or decisional processes; however, without further careful experimentation it will not be possible to separate the processing variance perfectly from the distributional variance. Nevertheless, computationally identifying data generated by this model can be accomplished by determining these six values (four means, two variances) underlying the activation distributions (combining the variance from perceptual and decisional processes). Finally, we included a scaling parameter for the final decision process so that the absolute activation levels could be attenuated. The result for a single stimulus in the dimensional activation model was given by the following equation $\Phi(s * [\varphi(D(x_1, \mu_{1A}), 0, \sigma_1^2) - \varphi(D(x_1, \mu_{1B}), 0, \sigma_1^2) + \varphi(D(x_2, \mu_{2A}), 0, \sigma_2^2) - s_1 \varphi(D(x_2, \mu_{2B}), 0, \sigma_2^2)])$, where Φ is the cumulative density function of the standard normal, $\varphi(X, \mu, \sigma^2)$ is a Gaussian function with mean μ and variance σ^2 evaluated at X, D is the distance function, s is the scaling parameter, (x_1, x_2) is the stimulus values in dimensions 1 and 2, μ_{1A} and μ_{1B} are the means of the A and B distributions in the first dimension, μ_{2A} and μ_{2B} are the means of the A and B distributions in the second dimension, and σ_1 and σ_2 are the variances of the activation distributions for the first and second dimensions. Note that for the configural activation model, the function only changes by the joining ("binding") of the dimensional activations by using a multi-variate radial basis function instead of a univariate Gaussian function.

Appendix B

The prototype model and the general quadratic classifier have been dealt with fairly thoroughly in the literature (Ashby, 1992a) as well as in Appendix A1, and the association models and their parameters for fitting have been described in Appendix A2. Consequently, we will not expend too much time here to reiterate their differences. Both classifiers use a distance function to evaluate a given stimulus. For our models, in both cases, this function value was subtracted from a threshold and divided by a scaling factor and the resulting value was mapped to the likelihood of the two responses using the inverse standard normal function. This yielded a tractable probabilistic version of these analytic models, allowing us to compare their efficacy against the association based models using log likelihood, which we used as a step of evaluating model fit. Thus, for each response the pigeon made, we evaluated the likelihood of seeing that response for each model.

All model fitting was conducted in MATLAB. Each categorization method was implemented as a separate function that generated a probability of seeing a given response value. These probabilities were compared against the pigeons' response data to compute the log likelihood. In order to determine the optimal set of coefficient values for each of these models, we replicated this process at least 10,000,000 times for each problem using a grid that encompassed all likely values.

We then searched for the parameter sets that minimized the negative log likelihood, using the best parameter sets from the previous step as the starting point for the GlobalSearch solver from the MATLAB optimization toolbox. The final log likelihood was used in the AIC computation.

Appendix C. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2018.11.011. The raw data that underlies the graphs and analyses reported in this manuscript has been deposited in FigShare at https://figshare.com/articles/Data_for_Testing_Analogical_Rule_Transfer_in_Pigeons_Columba_livia_/7361417.

References

Ashby, F. G. (1992a). Multidimensional models of categorization. In F. G. Ashby (Ed.). *Multidimensional models of perception and cognition* (pp. 449–483). Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.

Ashby, F. G. (Ed.). (1992). *Multidimensional models of perception and cognition*. Hillsdale, NJ, England: Lawrence Erlbaum Associates Inc.

Ashby, F. G., & Alfonso-Reese, L. A. (1995). Categorization as probability density estimation. *Journal of Mathematical Psychology, 39*(2), 216–233. https://doi.org/10.1006/jmps.1995.1021.

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*(3), 442–481. https://doi.org/10.1037/0033-295X.105.3.442.

Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review, 114*(3), 632–656. https://doi.org/10.1037/0033-295X.114.3.632.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 33–53.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56*, 149–178. https://doi.org/10.1146/annurev.psych.56.091103.070217.

Ashby, F. G., Paul, E. J., & Maddox, W. T. (2011). COVIS. In A. J. Wills, & E. M. Pothos (Eds.). *Formal approaches in categorization* (pp. 65–87). Cambridge: Cambridge University Press.

Ashby, F. G., & Soto, F. A. (2015). Multidimensional signal detection theory. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.). *The oxford handbook of computational and mathematical psychology* (pp. 13–34). New York: Oxford University Press.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93*(2), 154–179. https://doi.org/10.1037/0033-295X.93.2.154.

Ashby, F. G., & Valentin, V. V. (2017). Chapter 7 - Multiple systems of perceptual category learning: Theory and cognitive tests A2 - Cohen, Henri. In C. Lefebvre (Ed.). *Handbook of categorization in cognitive science* (pp. 157–188). (2nd ed.). San Diego: Elsevier.

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review, 6*, 363–378. https://doi.org/10.3758/BF03210826.

Beach, L. R. (1964). Cue probabilism and inference behavior. *Psychological Monographs: General and Applied, 78*(5–6), 1–20. https://doi.org/10.1037/h0093853.

Blough, D. S. (1972). Recognition by the pigeon of stimuli varying in two dimensions. *Journal of the Experimental Analysis of Behavior, 18*, 345–367.

Burgering, M. A., ten Cate, C., & Vroomen, J. (2018). Mechanisms underlying speech sound discrimination and categorization in humans and zebra finches. *Animal Cognition, 21*(2), 285–299. https://doi.org/10.1007/s10071-018-1165-3.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(2), 261–304. https://doi.org/10.1177/0049124104268644.

Casale, M. B., Roeder, J. L., & Ashby, F. G. (2012). Analogical transfer in perceptual categorization. *Memory & Cognition, 40*(3), 434–449. https://doi.org/10.3758/s13421-011-0154-4.

Cook, R. G., & Smith, J. D. (2006). Stages of abstraction and exemplar memorization in pigeon category learning. *Psychological Science, 17*(12), 1059–1067. https://doi.org/10.1111/j.1467-9280.2006.01833.x.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1*(3), 225–241. https://doi.org/10.1016/0010-0285(70)90016-2.

George, D. N., & Pearce, J. M. (2012). A configural theory of attention and associative learning. *Learning & Behavior, 40*(3), 241–254. https://doi.org/10.3758/s13420-012-0078-2.

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour, 66*(1), 15–36. https://doi.org/10.1006/anbe.2003.2174.

Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition, 37*(1–2), 133–166.

Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science, 146*(3643), 549–551. https://doi.org/10.1126/science.146.3643.549.

Jassik-Gerschenfeld, D., & Hardy, O. (1979). Single-neuron responses to moving sine-wave gratings in the pigeon optic tectum. *Vision Research, 19*, 993–999.

Lea, S. E. G., Pothos, E. M., Wills, A. J., Leaver, L. A., Ryan, C. M. E., & Meier, C. (2018). Multiple feature use in pigeons' category discrimination: The influence of stimulus set structure and the salience of stimulus differences. *Journal of Experimental Psychology: Animal Learning and Cognition, 44*(2), 114–127. https://doi.org/10.1037/xan0000169.

Lea, S. E. G., & Wills, A. J. (2008). Use of multiple dimensions in learned discriminations. *Comparative Cognition & Behavior Reviews, 3*, 115–133. https://doi.org/10.3819/ccbr.2008.30007.

Lengersdorf, D., Pusch, R., Güntürkün, O., & Stüttgen, M. C. (2014). Neurons in the pigeon nidopallium caudolaterale signal the selection and execution of perceptual decisions. *European Journal of Neuroscience, 40*(9), 3316–3327. https://doi.org/10.1111/ejn.12698.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics, 53*(1), 49–70. https://doi.org/10.3758/bf03211715.

McDaniel, M. A., Cahill, M. J., Robbins, M., & Wiener, C. (2014). Individual differences in learning and transfer: Stable tendencies for learning exemplars versus abstracting rules. *Journal of experimental psychology. General, 143*(2), 668–693. https://doi.org/10.1037/a0032963.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207–238. https://doi.org/10.1037/0033-295x.85.3.207.

Nomura, E., Maddox, W., Filoteo, J., Ing, A., Gitelman, D., Parrish, T., ... Reber, P. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex, 17*(1), 37–43. https://doi.org/10.1093/cercor/bhj122.

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

Pearce, J. M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning & Behavior, 30*(2), 73–95. https://doi.org/10.3758/BF03192911.

Pearce, J. M., Esber, G. R., George, D. N., & Haselgrove, M. (2008). The nature of discrimination learning in pigeons. *Learning & Behavior, 36*(3), 188–199. https://doi.org/10.3758/lb.36.3.188.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325.

Smith, J. D., Ashby, F. G., Berg, M. E., Murphy, M. S., Spiering, B., Cook, R. G., & Grace, R. C. (2011). Pigeons' categorization may be exclusively nonanalytic. *Psychonomic Bulletin & Review, 18*(2), 414–421. https://doi.org/10.3758/s13423-010-0047-8.

Smith, J. D., Beran, M. J., Crossley, M. J., Boomer, J., & Ashby, F. G. (2010). Implicit and explicit category learning by macaques (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes, 36*(1), 54. https://doi.org/10.1037/a0015892.

Smith, J. D., Berg, M. E., Cook, R. G., Murphy, M. S., Crossley, M. J., Boomer, J. T., ... Ashby, F. G. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience & Biobehavioral Reviews, 36*(10), 2355–2369. https://doi.org/10.1016/j.neubiorev.2012.09.003.

Smith, J. D., Redford, J. S., & Haas, S. M. (2008). Prototype abstraction by monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: General, 137*, 390–401. https://doi.org/10.1037/0096-3445.137.2.390.

Smith, J. D., Zakrzewski, A. C., Johnston, J. J. R., Roeder, J. L., Boomer, J., Ashby, F. G., & Church, B. A. (2015). Generalization of category knowledge and dimensional categorization in humans (Homo sapiens) and nonhuman primates (Macaca mulatta). *Journal of Experimental Psychology: Animal Learning and Cognition, 41*(4), 322–335. https://doi.org/10.1037/xan0000071.

Tappeiner, C., Gerber, S., Enzmann, V., Balmer, J., Jazwinska, A., & Tschopp, M. (2012). Visual acuity and contrast sensitivity of adult zebrafish. *Frontiers in Zoology, 9*(1), 1–6. https://doi.org/10.1186/1742-9994-9-10.

Teng, Y., Vyazovska, O., & Wasserman, E. (2015). Selective attention and pigeons' multiple necessary cues discrimination learning. *Behavioural Processes, 112*, 61–71. https://doi.org/10.1016/j.beproc.2014.08.004.

Vyazovska, O. V., Teng, Y., & Wasserman, E. A. (2014). Attentional tradeoffs in the pigeon. *Journal of the Experimental Analysis of Behavior, 101*(3), 337–354. https://doi.org/10.1002/jeab.82.

Wasserman, E. A., Kiedinger, R. E., & Bhatt, R. S. (1988). Conceptual behavior in pigeons: Categories, subcategories, and pseudocategories. *Journal of Experimental Psychology: Animal Behavior Processes, 14*(3), 235–246. https://doi.org/10.1037/0097-7403.14.3.235.

Wills, A. J., Lea, S. E. G., Leaver, L. A., Osthaus, B., Ryan, C. M. E., Suret, M. B., ... Millar, L. (2009). A comparative analysis of the categorization of multidimensional stimuli: I. Unidimensional classification does not necessarily imply analytic processing; evidence from pigeons (Columba livia), squirrels (Sciurus carolinensis), and humans (Homo sapiens). *Journal of Comparative Psychology, 123*(4), 391–405. https://doi.org/10.1037/a0016216.

Zakrzewski, A. C., Church, B. A., & Smith, J. D. (2018). The transfer of category knowledge by macaques (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 132*(1), 58–74. https://doi.org/10.1037/com0000095.