# Effect of between-category similarity on basic level superiority in pigeons

Olga F. Lazareva [a,*], Fabián A. Soto [b], Edward A. Wasserman [b]

[a] *Department of Psychology, Drake University, 316 Olin Hall, Des Moines, IA 50311-4505, United States*
[b] *Department of Psychology, University of Iowa, United States*

## ABSTRACT

Children categorize stimuli at the basic level faster than at the superordinate level. We hypothesized that between-category similarity may affect this basic level superiority effect. Dissimilar categories may be easy to distinguish at the basic level but be difficult to group at the superordinate level, whereas similar categories may be easy to group at the superordinate level but be difficult to distinguish at the basic level. Consequently, similar basic level categories may produce a superordinate-before-basic learning trend, whereas dissimilar basic level categories may result in a basic-before-superordinate learning trend. We tested this hypothesis in pigeons by constructing superordinate level categories out of basic level categories with known similarity. In Experiment 1, we experimentally evaluated the between-category similarity of four basic level photographic categories using multiple fixed interval-extinction training (Astley and Wasserman, 1992). We used the resultant similarity matrices in Experiment 2 to construct two superordinate level categories from basic level categories with high between-category similarity (cars and persons; chairs and flowers). We then trained pigeons to concurrently classify those photographs into either the proper basic level category or the proper superordinate level category. Under these conditions, the pigeons learned the superordinate level discrimination faster than the basic level discrimination, confirming our hypothesis that basic level superiority is affected by between-category similarity.

© 2010 Elsevier B.V. All rights reserved.

Categories are often considered to be organized at multiple hierarchical levels, such as the subordinate, basic, and superordinate levels. The subordinate-level categories *office chair* and *dining chair* are included in the basic level category *chair* which, in turn, is included in the superordinate level category *furniture*.

Starting with the pioneering research of Rosch and Mervis (1975) and Rosch et al. (1976), categorization at the basic level has consistently been found to be preferred to categorization at either the subordinate or superordinate levels. People are faster to name objects at the basic level than at either the subordinate or super-ordinate level (Jolicoeur et al., 1984; Murphy and Brownell, 1985; Murphy and Smith, 1982; Rosch et al., 1976); they also prefer to use basic level terms to name objects (e.g., "chair") rather than super-ordinate level terms (e.g., "furniture") or subordinate-level terms (e.g., "office chair"; Lin et al., 1997; Rosch et al., 1976).

Rosch and Mervis (1975) hypothesized that this preference for basic level categorization is encouraged by high *within-category similarity* (i.e., perceptual coherence within a basic level category) and by low *between-category similarity* (i.e., perceptual coherence across different basic level categories). On the other hand, superordinate categories entail low within-category similarity and low between-category similarity. Consequently,

superordinate level categorization is less preferred than basic level categorization.

Rosch et al. (1976) have also suggested that children exhibit basic level categorization before they exhibit superordinate level categorization. To test this developmental proposal, these researchers presented 3-year-old children with three objects (e.g., two airplanes and a dog or an airplane, a car and a dog) and asked them to identify the two objects that were alike. Children successfully identified two similar objects in the basic level task (e.g., two airplanes), but not in the superordinate level task (e.g., an airplane and a car), suggesting that basic level categorization might indeed precede superordinate level categorization.

Later studies have confirmed this basic-before-superordinate trend (Quinn et al., 1993; Rakison and Butterworth, 1998), although other reports have questioned the generality of these findings. For example, Mandler and Bauer (1988) used a sequential touching procedure in which infants are first presented with an object (e.g., a toy poodle) and are later given a group of other objects with which to play (e.g., a set of four toy dogs or a set of four toy cars). If the infants touched a group of objects more frequently than expected by chance, then they were presumed to group these objects into a single category. Mandler and Bauer (1988) found that 12- and 20-month-old infants more frequently touched the objects within the same superordinate level category (e.g., animals or vehicles) than within the same basic level category (e.g., horses or dogs), indicating the presence of superordinate level categories, but not

* Corresponding author. Tel.: +1 515 271 4801; fax: +1 515 271 1925.
*E-mail address:* olga.lazareva@drake.edu (O.F. Lazareva).

basic level categories. Other studies have confirmed these findings (Bornstein and Arterberry, 2010; Mandler et al., 1991; reviewed by Quinn, 2002).

Several attempts have been made to reconcile these strikingly contradictory results. For example, Mareschal and Quinn (2001) suggested that infants' performance may depend on whether the experimental technique includes a familiarization phase. A familiarization phase usually involves the presentation of several objects that belong to the same category and several objects that do not belong to that category. Consequently, this procedure could affect

infants' perception of category boundaries (see also Mareschal and Tan, 2007; Quinn, 2002).

Another possible explanation may involve the variability of exemplars within a single category as well as the variability of exemplars across different categories. In a recent study by Mareschal and Tan (2007), infants were given the possibility of playing with a set of homogeneous toys (e.g., four toy birds) or a set of heterogeneous toys (e.g., a toy bed, a wardrobe, a spoon and a teapot), and then were given a sequential touching procedure. During this procedure, all of the infants were exposed to toys that
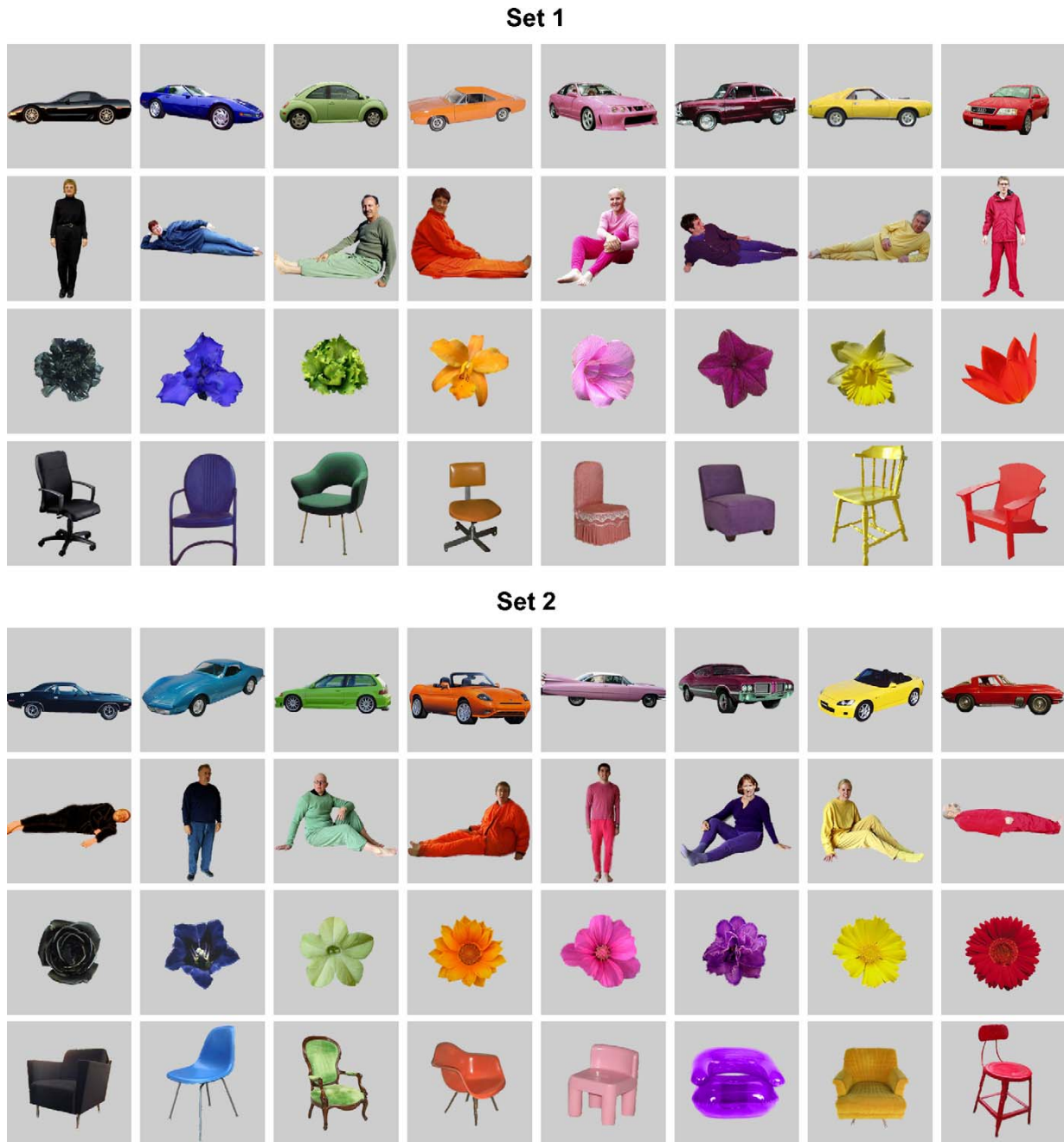


**Fig. 1.** Discriminative stimuli (Sets 1 and 2) used in Experiments 1 and 2.

were different from the toys that were presented during the familiarization phase. The infants who played with the heterogeneous toys touched objects forming basic level categories but not superordinate level categories, whereas the infants who played with the homogeneous toys touched objects forming both basic level and superordinate categories. These and other results (e.g., Bornstein and Arterberry, 2010; Bornstein et al., 2010; Mareschal and Tan, 2008) suggest that category variability may indeed affect infants' categorization behavior.

In a thematically related project, we trained pigeons to concurrently classify photographs of cars, chairs, flowers and persons (see Fig. 1) either at the basic level or at the superordinate level (Lazareva et al., 2004). Fig. 2 illustrates the sequence of events in the course of a trial. For example, if a photograph of a car were shown together with a set of four choice keys, then the pigeon had to select the key that was associated with cars. But, if the same photograph were shown together with a different set of two choice keys, then the pigeon had to select the key that was associated with both cars and chairs (car–chair category). A second superordinate category comprised images of flowers and persons (flower–person category). Both basic level and superordinate level trials were equally often shown during training.

Note that this training regime does not provide direct evidence for the hierarchical structure of the superordinate level category. However, prior research has clearly documented that pigeons trained to associate perceptually dissimilar objects with a common response do integrate those objects into a functionally equivalent stimulus class (Wasserman et al., 1992) and our pigeons were trained in the same manner. Thus, it is plausible to assume that our superordinate level categories do bear at least functional resemblance to superordinate level human-language categories.

The pigeons in Lazareva et al. (2004) mastered both component discriminations, flexibly categorizing the same stimuli at two different levels depending on task demands. Detailed analysis of the acquisition patterns (see Fig. 3) revealed an unexpected disparity. Pigeons required significantly fewer training session to master the basic level discrimination than the superordinate level discrimination when they were shown photographs of cars and chairs; but, the same pigeons were slightly, albeit nonsignificantly, faster to acquire the superordinate level discrimination than basic level discrimination when they were shown photographs of flowers and persons.

What is responsible for this unexpected disparity? We hypothesized that there may be differential between-category similarity across these four basic level categories. Highly dissimilar basic level categories may be easy to discriminate at the basic level, but be difficult to group at the superordinate level, resulting in the basic-before-superordinate learning trend. Conversely, highly similar basic level categories may be easy to group at the superordinate level, but be difficult to discriminate at the basic level, producing the superordinate-before-basic learning trend. This hypothesis
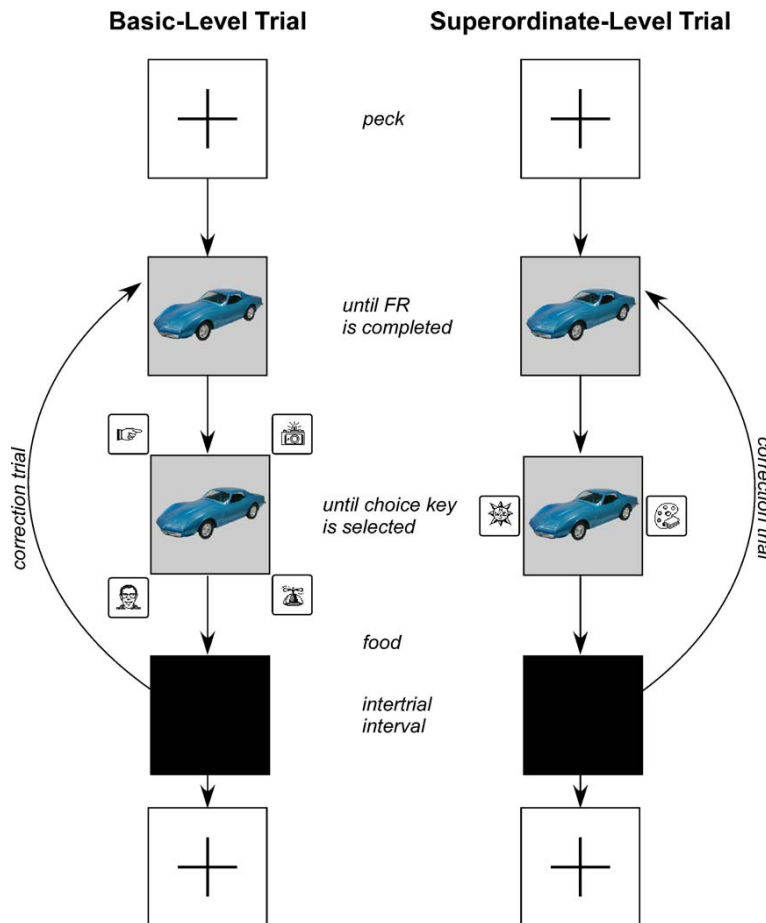


**Fig. 2.** The sequence of events in the course of a trial in Lazareva et al. (2004) and in Experiment 2.
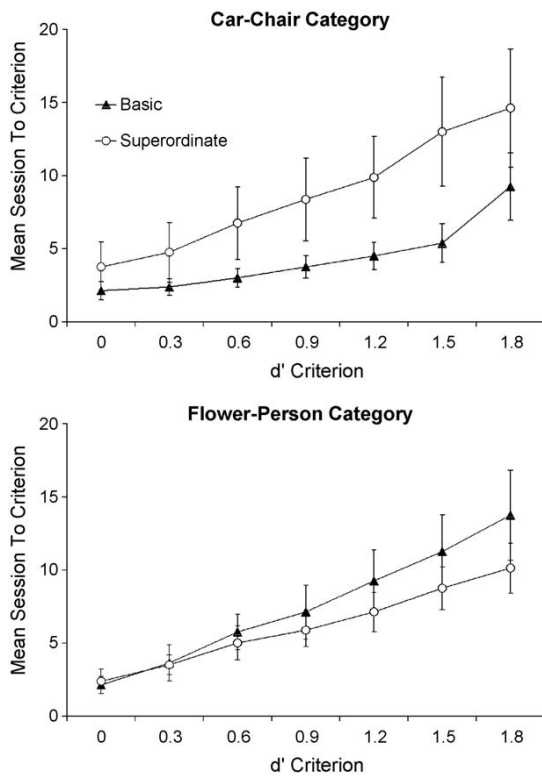
**Fig. 3.** Mean number of sessions required to reach each *d′* criterion throughout training on basic and superordinate tasks. Redrawn from Lazareva et al. (2004).

suggests that pigeons ought to be more apt to confuse flowers with persons than to confuse cars with chairs during basic level discrimination. The confusion errors that we collected in the basic level discrimination task indeed indicated that photographs of flowers and persons were prone to be confused with each other (47.0% of errors compared to 33.3% expected by chance), whereas photographs of cars and chairs were not (30.5% of errors). This pattern of errors was later replicated when the same birds were tested with different stimulus exposure durations and different choice key delays (Lazareva and Wasserman, 2009).

Although these findings support our hypothesis, other factors may have affected our pigeons' categorical discrimination behavior. Prior research has shown that pigeons treat perceptually dissimilar stimuli as members of the same class if these stimuli were previously associated with a common response (Wasserman et al., 1992). Moreover, the prior association with a common response results in decreased discriminability among the stimuli associated with it (Kaiser et al., 1997). Conversely, low perceptual similarity weakens pigeons' ability to form a stimulus class on the basis of association with a common response (Urcuioli and Vasconcelos, 2008). Because our pigeons had to respond in one fashion to chairs or cars and in a second fashion to flowers or persons, it is difficult to disentangle the effects of between-category similarity from common response training.

In the current study, we eliminated this problem by first evaluating pigeons' perception of between-category similarity in the absence of explicit categorical discrimination (Experiment 1). We then used these data to predict other pigeons' acquisition trends in differently constructed superordinate level categories (Experiment 2).

## 1. Experiment 1

In Experiment 1, we used the multiple fixed interval-extinction training procedure devised by Astley and Wasserman (1992) to evaluate pigeons' perception of similarity across four basic level categories in the absence of a categorical discrimination. Specifically, pigeons were presented with a set of eight reinforced stimuli drawn from one basic level category (e.g., cars) and a set of 32 non-reinforced stimuli drawn from each of four basic level categories (cars, flowers, chairs and persons). If pigeons perceived the stimuli from one or more of the basic level categories to be more similar to the reinforced stimuli, then they ought to commit more errors to these categories. We further analysed the pattern of confusion errors by constructing a two-dimensional similarity map and by measuring the distances among the four basic level categories in this similarity space.

## 2. Method

### 2.1. Subjects

Eight feral pigeons (Columba livia) were maintained at 85% of their free-feeding body weights by controlled daily feeding. Grit and water were available ad libitum in the home cages. The pigeons had served in unrelated studies prior to the experiment.

### 2.2. Stimuli

The 64 discriminative stimuli that are shown in Fig. 1 (also viewable in color at http://www.psychology.uiowa.edu/Faculty/Wasserman/graphics.html) were photographs from the World Wide Web plus digital photographs used in our own prior studies (Lazareva et al., 2004). The original background of each photograph was replaced by a solid 20% gray shading using Photoshop® 7.0 (Adobe®, San Jose, CA) and KnockOut 2 (Procreate™, Ottawa, Canada). The stimuli came from four basic level categories: cars, chairs, flowers and persons. We created eight subsets of photographs composed of two stimuli in each category that were primarily black, blue, green, orange, pink, purple, red or yellow. Small areas of other colors (such as the tires of cars, the faces of persons, and the legs of chairs) were allowed as well. All of the target objects had approximately the same area and were presented in different orientations counterbalanced across images. The 64 images were divided into two sets (1 and 2) of 32 and their colors were balanced across black, blue, green, orange, pink, purple, red or yellow.

An additional black-and-white cross-hatching pattern was created for baseline training. All of the images were saved as PICT files with 144 dpi resolution.

### 2.3. Apparatus

The experiment used four 36 cm × 36 cm × 41 cm operant conditioning chambers detailed by Gibson et al. (2004). The boxes were located in a dark room with continuous white noise. The stimuli were presented on a 15-in LCD monitor located behind an AccuTouch® resistive touchscreen (Elo TouchSystems, Fremont, CA), so that the pigeons could view most of the monitor area. A food cup was centered on the rear wall level with the floor. A food dispenser delivered 45 mg food pellets through a vinyl tube into the cup. A houselight on the rear wall provided illumination during the session. Each chamber was controlled by an Apple® iMac® computer. The experimental procedure was programmed in HyperCard (Version 2.4, Apple Computer, Inc., Cupertino, CA). One 10.16 cm square area in the middle of the screen was used to display the photographs and to record observing responses; the rest of the screen was black.

**Table 1**
The counterbalancing of stimulus sets and the order of presentation of the reinforced (S+) category in Experiment 1. Each stage consisted of baseline training and discrimination training.

| Bird | Reinforced (S+) category | | | |
| --- | --- | --- | --- | --- |
| | Stage 1 | Stage 2 | Stage 3 | Stage 4 |
| Set 1 as S+ | | | | |
| 66Y | Car | Flower | Chair | Person |
| 72Y | Chair | Car | Person | Flower |
| 80Y | Flower | Person | Car | Chair |
| 8B | Person | Chair | Flower | Car |
| Set 2 as S+ | | | | |
| 33B | Car | Flower | Chair | Person |
| 66W | Chair | Car | Person | Flower |
| 93W | Flower | Person | Car | Chair |
| 75B | Person | Chair | Flower | Car |

### 2.4. Procedure

We used multiple fixed interval-extinction training to evaluate between-category similarity for each of the four categories for each bird. One set of stimuli (e.g., Set 2) was chosen to serve as the source of the nonreinforced (S−) stimuli; these stimuli were presented during each stage of training. The second set of stimuli (e.g., Set 1) served as the source of the reinforced (S+) stimuli and provided one S+ category for each stage of training. Table 1 illustrates the counterbalancing of stimulus sets and the order of category presentation for all eight birds. Each stage involved baseline training and discrimination training and each bird was exposed to four stages of training; so, for each bird, all four basic level categories were reinforced once.

#### 2.4.1. Baseline training

Baseline training ensured that the birds responded to each stimulus at relatively high and uniform rates before discrimination training began. Each session comprised two blocks of 80 trials, for a total of 160 trials. We used a set of 40 stimuli for each pigeon: 32 stimuli from four different categories that later served as the S−s and eight stimuli from one category that served as the S+s. During baseline training, all 40 stimuli served as S+s to establish a high rate of response; these stimuli were presented once during each 80-trial block. On the remaining 40 trials of each block, the birds were shown a cross-hatching stimulus that served as the S−. This task was designed to encourage the birds to attend to all of the stimuli and to help them adjust to a partial reinforcement schedule.

A trial began with the pigeon being shown a black cross in the center of a white display screen. Following one peck anywhere on the display, a stimulus was displayed for a fixed interval of 15 s. If responding to the stimulus was to be reinforced, then the pigeon had to peck the stimulus once after the 15 s had elapsed in order to receive food. If responding to the stimulus was not to be reinforced, then the stimulus terminated after 15 s and no reinforcement was delivered. The intertrial interval varied from 5 to 9 s.

Baseline training lasted a minimum of 6 days. On day 6, the pigeon had to meet a "stability" criterion in order to proceed to discrimination training. To ensure that the birds were responding to all four of the pictorial categories at high and similar rates, we calculated the response rates to each of the four pictorial S+ categories and we divided the sum of pecks to all of the S+ stimuli by the total number of pecks to both the S+ and the S− stimuli. If the disparity in response rate to each of the four stimulus categories was less than 20% and the S+/(S+ plus S−) ratio was 0.8 or higher, then the bird proceeded to discrimination training. If the stability criterion was not met on day 6, then the bird continued baseline training until it met criterion.

#### 2.4.2. Discrimination training

Each session comprised four blocks of 40 trials, for a total of 160 trials. Each block consisted of 32 stimuli (e.g., Set 2) that served as the S−s and eight stimuli (e.g., category car from Set 1) that served as the S+s. The sequence of events during a trial was the same as in baseline training, but now only eight out of 40 photographs were reinforced. Discrimination training continued until the birds reliably responded to the S+ category and stopped responding to the S− categories. To ensure that this was the case, we calculated the discrimination ratio for the S+ category by dividing the number of pecks to this category by the total number of pecks during the session; this ratio had to be 0.80 or higher. Additionally, we calculated response rate to each of the four S− categories by dividing the number of pecks to each S− category by the total number of pecks during the session; all of these response rates had to be 0.10 or lower.

### 2.5. Behavioral measures

We used the total number of pecks made during each 15 s of stimulus exposure as our basic dependent measure. To see if the birds' confusion errors were influenced by perceptual similarity, we first computed the percentage of confusion errors for all four categories by dividing the total number of responses to all of the S−s belonging to each category by the total number of responses to all of the S−s from all four categories multiplied by 100 (see Table 2). If the birds' responses were indeed influenced by perceptual similarity, then we should see evidence of within-category similarity defined by a higher percentage of confusion errors to the same S− category as the S+ category.

Additionally, we computed the percentage of confusion errors for only the three non-S+ categories. For example, if the S+ category were car, then we separately computed the percentage of confusion errors only for the chair, flower, and person categories. This percentage of non-S+ confusion errors was calculated by dividing the number of responses to all of the stimuli in each non-S+ category by the total number of responses to all three of the non-S+ categories in each stage of differential training. This percentage should indicate which of the three non-S+ categories was perceived by the pigeons as being most similar to the S+ category thereby providing evidence of between-category similarity. The non-S+ confusion errors were computed primarily for use in multidimensional scaling analysis which involved only the non-S+ categories.

Preliminary statistical analyses revealed no differences related to the reinforced set of photographs (Set 1 vs. Set 2). So, the data of all eight birds were pooled in the following analyses.

### 2.6. Multidimensional scaling

We used the percentage of non-S+ confusion errors to construct a similarity confusion matrix for each bird. Because most of the square matrices were slightly asymmetrical—i.e., the lower half of the matrix was somewhat different from the upper half of the matrix (cf. Table 3)—we created a single triangular matrix by averaging the upper and lower triangular portions of the square matrix. Because of the exponentially decaying relationship between psychological similarity and distance in psychological space (Shepard, 1987), we used an exponential transformation of the confusion errors (Borg and Groenen, 2005; Lee, 2001). We then computed a two-dimensional map for each of the eight triangular confusion matrices using a ratio-level multidimensional scaling algorithm, Euclidean distance, and the SStress badness function (see Borg and Groenen, 2005 for more details). The computations were performed in PERMAP (Version 11.6; Heady and Lucas, 1997). Preliminary simulations revealed that a two-dimensional map produced a better fit than a one-dimensional map; higher-dimensional maps

**Table 2**
Mean percentage and standard deviation of complementary confusion errors for all four nonreinforced categories and for all eight birds in Experiment 1 (chance level at 25%). Bold font indicates the complementary category (S−s from the same category as S+s).

| Reinforced category | Nonreinforced category | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Car | | Chair | | Flower | | Person | |
| | M | SD | M | SD | M | SD | M | SD |
| Car | **44.1*** | **8.0** | 14.8## | 4.0 | 18.9## | 5.6 | 22.3 | 4.6 |
| Chair | 13.5## | 2.9 | **35.3*** | **6.5** | 30.0 | 9.7 | 21.2## | 4.1 |
| Flower | 10.9## | 2.7 | 21.4 | 4.4 | **49.8*** | **7.2** | 17.9## | 6.3 |
| Person | 17.1## | 4.9 | 20.3## | 4.3 | 21.8 | 5.5 | **40.8*** | **8.9** |

  * Significantly above chance, two-tailed $t$-test, $n = 8$, $p < 0.05$.
## Significantly below chance, two-tailed $t$-test, $n = 8$, $p < 0.05$.

**Table 3**
Mean percentage and standard deviation of non-complementary confusion errors for all four nonreinforced categories for eight birds in Experiment 1 (chance level at 33%). Note that non-complementary S−s do not come from the same category as the S+s.

| Reinforced category | Nonreinforced category | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Car | | Chair | | Flower | | Person | |
| | M | SD | M | SD | M | SD | M | SD |
| Car | – | – | 26.0## | 4.4 | 33.3 | 7.1 | 40.7 | 10.9 |
| Chair | 21.0## | 4.4 | – | – | 45.6* | 11.0 | 33.4 | 8.8 |
| Flower | 21.6## | 3.8 | 43.0* | 7.7 | – | – | 35.4 | 9.8 |
| Person | 29.1 | 7.7 | 34.3 | 4.7 | 36.6 | 5.7 | – | – |

  * Significantly above chance, two-tailed $t$-test, $n = 8$, $p < 0.05$.
## Significantly below chance, two-tailed $t$-test, $n = 8$, $p < 0.05$.

could not be tested because of the small number of available categories.

To simplify the visual representation of the data, the eight individual maps were used to find a centroid map using the Procrustean individual differential scaling (PINDIS) procedure. This procedure uses admissible transformations (translation, rotation, reflection, and dilation) to find the centroid map that best represents all eight individual maps. The goodness-of-fit, or the proportion of variance ($r^2$) accounted for by the centroid map, was then calculated for each individual map. These analyses were performed using the PINDIS module in NewMDSX (Version 5.1.3). However, all of the statistical analyses used the individual maps rather than the centroid map.

## 3. Results and discussion

### 3.1. Percentage of confusion errors for all four categories

To see whether the birds' confusion errors were influenced by perceptual similarity within each basic level category, we first analysed the percentage of confusion errors for all four categories. Table 2 shows that the S−s from the same category as the S+s consistently attracted the highest percentage of confusion errors. All of these percentages were significantly above chance (25%). Additionally, errors to the S−s from the other three non-complementary categories either did not differ from chance or were significantly below chance, again demonstrating reliable within-category similarity.

To confirm these observations, we conducted a full-factorial analysis of variance (ANOVA) with reinforced category (car, chair, flower and person) and nonreinforced category ("complementary" or same as S+, "non-complementary" or different from S+) as fixed factors and bird (eight) as a random factor. The ANOVA found a significant main effect of nonreinforced category, $F(1, 7) = 157.79$, $p < 0.0001$, indicating that complementary categories attracted significantly more errors than non-complementary categories. Although the nonreinforced category × reinforced category interaction was significant, $F(3, 21) = 5.08$, $p = 0.008$, planned contrasts found that errors to the complementary categories were significantly higher than errors to the non-complementary cate-

gories for each of the four reinforced categories, $F(1, 21) \geq 14.50$, $p \leq 0.001$. We therefore concluded that pigeons' responses here were strongly affected by the perceptual similarity of the stimuli, as was previously documented by Astley and Wasserman (1992).

### 3.2. Percentage of non-complementary confusion errors

Table 3 shows the mean percentage of non-complementary confusion errors calculated to reflect perceptual similarity between the four basic level categories. To help visualize the pattern of between-category similarity, we computed a two-dimensional centroid map based on the eight individual maps that were derived from the individual confusion matrices (see Section 2.6 for further details). Fig. 4 shows the resultant centroid map. The two most similar (and, therefore, most confusable) basic level categories were chairs and flowers; the two most dissimilar (and, therefore least confusable) basic level categories were chairs and cars. More pertinent to the present theoretical concern, the map also shows that the distance in
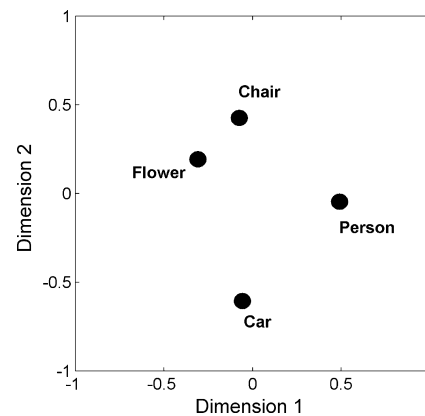


**Fig. 4.** The two-dimensional centroid map derived from the eight individual similarity maps. The correlation of the individual maps with the centroid map equaled: .70 for 8B, .80 for 66Y, 0.82 for 72Y, 0.86 for 80Y, 0.87 for 33B, 0.55 for 66W, 0.86 for 75B, and .68 for 93W.
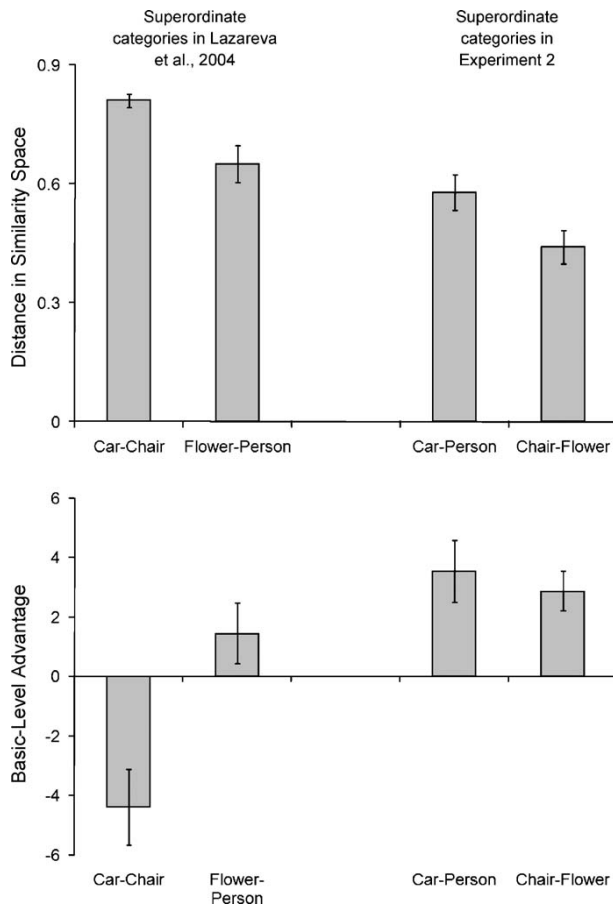
**Fig. 5.** Top panel: mean euclidean distances between different superordinate combinations. These distances were derived from the individual similarity maps rather than from the centroid map that is shown in Fig. 2. Error bars represent the standard error of the mean. Bottom panel: basic level advantage scores for the car–chair and flower–person combinations (from Lazareva et al., 2004) and for the car–person and chair–flower combinations (from present report). Numbers below zero indicate that the basic level discrimination was acquired faster than the superordinate level discrimination, whereas numbers above zero indicate that the superordinate level discrimination was acquired faster than the basic level discrimination.

the similarity space between cars and chairs was larger than the distance between flowers and persons. This disparity could have been responsible for the different acquisition patterns that were found during basic level and superordinate level discrimination learning reported earlier (Lazareva et al., 2004) and that are depicted in Fig. 3.

To explore this possibility, we used the eight individual similarity maps to compute the distances between the basic level car category and chair category and between the basic level flower category and person category. As the top panel of Fig. 5 shows, the car–chair distance (M ± SD: 0.81 ± 0.05) was significantly greater than the flower–person distance (0.65 ± 0.13; two-tailed paired t-test, $t$ (7) = −4.21, $p$ = 0.004). The other distances shown in the top panel of Fig. 5 were used to create new categories for Experiment 2 and will be explained later.

Thus, the present data corroborate our previously reported findings (Lazareva et al., 2004); cars and chairs are indeed the two most dissimilar categories (cf. Fig. 4 and top panel of Fig. 5). Thus, combining them into a superordinate level category produces the basic-before-superordinate trend reported earlier. Likewise, flowers and persons are relatively similar to each other (cf. Fig. 4 and

top panel of Fig. 5), which results in the superordinate-before-basic trend.

## 4. Experiment 2

Our earlier results (Lazareva et al., 2004), plus the data from Experiment 1, suggest that between-category similarity may help to explain the relative speed of pigeons' acquiring basic level and superordinate level discriminations. If this were indeed the case, then we ought to be able to manipulate the relative speeds of acquiring the two tasks by manipulating the similarity of the basic level categories that constitute the superordinate level category. Experiment 2 tested this proposal.

In Experiment 2, we created two different superordinate level categories by combining the basic level categories of chairs with flowers (distance, M ± SD: 0.44 ± 0.12) and cars with persons (distance, M ± SD: 0.58 ± 0.13; cf. Fig. 4). Fig. 5 shows that the mapped distance between flowers and persons (M ± SD: 0.65 ± 0.13) was numerically greater than the distance between cars and persons; however, that difference was not significantly different (two-tailed paired t-test, $t$ (7) = −0.95, $p$ = 0.38). The mapped distance between flowers and persons was significantly greater than the mapped distance between chairs and flowers (two-tailed paired t-test, $t$ (7) = 3.64, $p$ = 0.008). In Lazareva et al. (2004), the superordinate-before-basic trend was observed when flowers and persons were combined. Thus, we expected to observe the superordinate-before-basic trend for both of the categories created in the present study. The rest of the training procedure was the same as reported by Lazareva et al. (2004).

## 5. Method

### 5.1. Subjects

Four new feral pigeons (Columba livia) different from those in Experiment 1 were maintained at 85% of their free-feeding body weights by controlled daily feeding. Grit and water were available ad libitum in the home cages. The pigeons had served in unrelated studies prior to the experiment.

### 5.2. Stimuli and apparatus

The same stimuli (Fig. 1) and apparatus were used as in Experiment 1. One 10.16 cm square area in the middle of the screen was used to display the photographs and to record observing responses; the rest of the screen was black. Six black Macintosh icons on white backgrounds served as the report buttons and were 3.30 cm wide × 2.79 cm high. The basic level report buttons were placed at the corners of the square center button. For two birds, the superordinate level buttons were above and below the central display; for the other two birds, the superordinate level buttons were left and right of the central display (see Fig. 2).

### 5.3. Procedure

#### 5.3.1. Training

During training, the pigeons' task could be either four-alternative forced-choice report of the four basic level categories (cars, chairs, flowers or persons) or two-alternative forced-choice report of the two superordinate level categories (cars plus persons or chairs plus flowers), as shown in Fig. 2. A trial began with the pigeon being shown a black cross in the center of the white display screen. Following one peck anywhere on the display, a training photograph appeared. The bird had to complete an observing response requirement (gradually increasing from 3 to 26 pecks)

to the stimulus; then, the report buttons appeared. The observing report requirement depended on individual pigeons' performance. If the pigeon was completing the session but not meeting criterion, then the number of required pecks was increased. If the pigeon ceased to complete the session, then the number of required pecks was decreased.

On a basic level trial, four report buttons were presented; on a superordinate level trial, two different report buttons were shown. basic level and superordinate level trials were randomly presented throughout a session. For two pigeons, the training stimuli came from Set 1; for the other two pigeons, the training stimuli came from Set 2.

If the pigeon's report response was correct, then food was delivered and the intertrial interval ensued. If the pigeon's report response was incorrect, then the house light and the monitor screen darkened and a correction trial was given. Correction trials continued to be given until the correct response was made. All responses were recorded, but only the first report response of each trial was scored in data analysis.

A session consisted of three blocks of 64 trials, for a total of 192 trials. Each block consisted of 32 basic level trials and 32 superordinate level trials, so that each stimulus was shown once during each task. The pigeons were trained until they reached a criterion of 75% correct for each basic level category and 88% correct for each superordinate level category, so that both percentages yielded equal $d'$s of 1.8 (see below).

### 5.4. Behavioral measures

Because the two- and four-alternative forced-choice tasks involve different levels of chance performance (50% and 25%, respectively), direct comparison of accuracy scores was inappropriate. So, we transformed the percentage of correct choices to the signal detection measure $d'$ (Algorithm 1, Smith, 1982). Signal detection theory has been used to interpret performance in several perceptual and cognitive tasks, and $d'$ can be calculated for many measurement techniques and psychophysical procedures, including percentage correct and forced-choice procedures (Swets, 1964; Tanner and Birdsall, 1964). The chance level of discrimination corresponded to a $d'$ of 0.00 for both tasks. For all statistical tests, alpha was set at 0.05.

## 6. Results and discussion

Training to criterion took a mean of 61.75 days, with a standard deviation of 29.57 days (a minimum of 37 days and a maximum of 99 days). To analyse acquisition performance, we calculated the number of sessions required to reach $d'$s of 0.0, 0.3, 0.6, 0.9, 1.2, 1.5 and 1.8. A $d'$ of 0.0 corresponded to chance; a $d'$ of 1.8 was the highest level that all of the birds met to the individual basic level and superordinate level categories.

Previously, we reported that, during acquisition, pigeons more readily confused photographs of cars and chairs than photographs of flowers and persons (Lazareva et al., 2004; Lazareva and Wasserman, 2009). However, pigeons tend to group perceptually dissimilar stimuli into a single class when these stimuli have previously been associated with a common response; in our prior experiments, photographs of cars and chairs were associated with one key and photographs of flowers and persons were associated with another key. If, as Experiment 1 suggested, the pattern of confusion errors was influenced by between-category similarity rather than by response key assignment, then we ought to observe the same pattern of confusion errors in Experiment 2 (i.e., high confusion errors for flowers and persons, but low confusion errors for cars and chairs). Alternatively, if the pattern of confusion errors
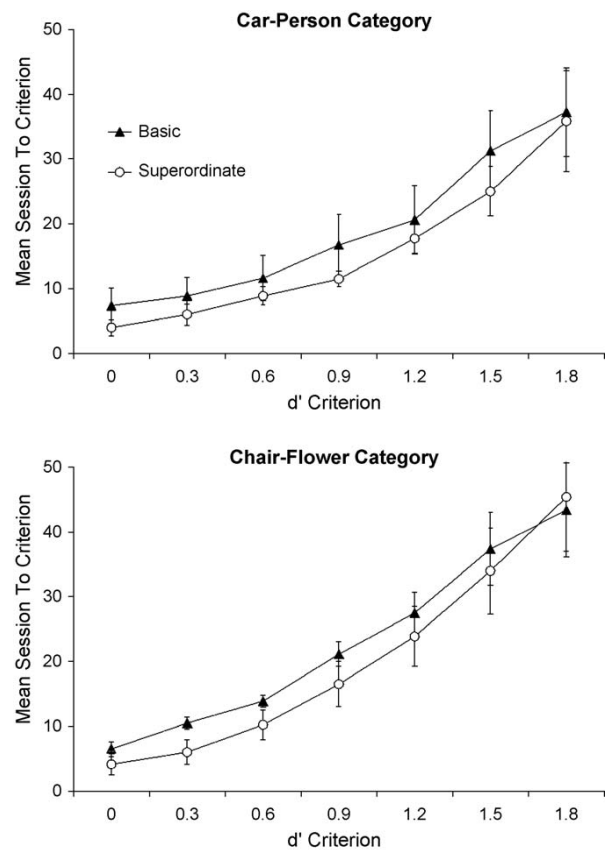


**Fig. 6.** Mean number of sessions required to reach each $d'$ criterion throughout training on basic and superordinate tasks for the car–person category (top panel) and the chair–flower category (bottom panel). A $d'$ of 0.0 corresponds to 25.0% and 50.0%, a $d'$ of 0.3 corresponds to 33.1% and 58.7%, a $d'$ of 0.6 corresponds to 42.2% and 66.8%, a $d'$ of 0.9 corresponds to 51.9% and 74.1%, a $d'$ of 1.2 corresponds to 61.5% and 80.2%, a $d'$ of 1.5 corresponds to 70.3% and 85.2%, and a $d'$ of 1.8 corresponds to 77.8% and 89.1%, for basic and superordinate tasks, respectively.

was influenced by response key assignment, then we should no longer observe greater confusion errors between flowers and persons because they were now assigned to two different response keys.

We found the same pattern of confusion errors as previously reported: the flower and person categories were more often confused with one another than would be expected by chance (41.4% instead of 33%, two-tailed $t$-test, $t$ (8) = 1.97, $p$ = 0.045; 1 out of 3 choice keys represented a confusion error). The percentage of confusion errors for car and chair categories did not differ from chance (31.9%, $t$ < 1). These results again suggest that the previously reported high confusability of the flower and person categories was due to perceptual similarity, not to response key assignment.

To create Fig. 6, we combined the $d'$ scores from the car and person basic trials to create a score for the basic car-person category and we combined the $d'$ scores from the chair and flower basic trials to create a score for the basic chair–flower category. We then compared these scores to the pigeons' responding to the superordinate car–person category and the superordinate chair–flower category. The same process was used to derive the scores in Lazareva et al. (2004) that are depicted in Fig. 3. Fig. 6 shows that both the car–person category and the chair–flower category were more readily classified at the superordinate level than at the basic level. Analysis of variance (ANOVA) with task (basic, superordinate) and

category (car–person, chair–flower) as fixed factors, criterion (7) as a nested factor within both fixed factors, Bird as a random factor, and with mean sessions to criterion as the dependent measure revealed a significant main effect of Task, $F(1, 3) = 26.62$, $p = 0.014$, but no significant task × category interaction, $F < 1$, indicating that the superordinate level task was mastered faster than the basic level task for both the chair–flower category and the car–person category. The ANOVA also revealed a significant main effect of Category, $F(1, 3) = 15.65$, $p = 0.029$, indicating that, on average, birds acquired both discriminations within the car–person category faster than within the chair–flower category ($17.3 \pm 1.8$ versus $21.5 \pm 2.1$ sessions, respectively).

In summary, the results of Experiment 2 supported our proposal. We were able to determine the relative speeds of pigeons' acquiring the basic level and superordinate level discriminations by manipulating the similarity of the basic level categories constituting the superordinate level category. This finding suggests that between-category similarity determines the emergence of the basic-before-superordinate or the superordinate-before-basic trends.

## 7. General discussion

We earlier found that, when presented with photographs of cars and chairs, pigeons learned the basic level discrimination significantly faster than the superordinate level discrimination (Lazareva et al., 2004); but, the opposite pattern arose when pigeons were presented with photographs of flowers and persons (Fig. 3). We suggested that this disparity could be due to differential between-category similarity among the four basic level categories (cars, chairs, flowers and persons).

In Experiment 1 of the present paper, we explicitly explored pigeons' perception of between-category similarity by presenting them with a set of reinforced stimuli selected from one basic level category (e.g., cars) and a set of nonreinforced stimuli that included all four basic level categories (cars, chairs, flowers and persons). By systematically rotating through all four sets of reinforced stimuli, we were able to construct a between-category confusion matrix (Table 2) that was subsequently used to derive a two-dimensional similarity map (Fig. 4). Analysis of the distances among the four basic level categories depicted in the top panel of Fig. 5 disclosed that cars and chairs were indeed the most dissimilar categories. Therefore, they should be easy to discriminate at the basic level, but they should be difficult to group at the superordinate level. Flowers and persons were closer to each other in the similarity space than were cars to chairs; consequently, they should be easy to group at the superordinate level, but they should be difficult to discriminate at the basic level. These were the same trends that were reported by Lazareva et al. (2004); (Fig. 3).

In Experiment 2 of the present paper, we used the similarity map that we obtained in Experiment 1 to construct two different superordinate categories from basic level categories with relatively high between-category similarity: cars plus persons and chairs plus flowers. Indeed, the distance between cars and persons and between chairs and flowers in two-dimensional similarity space was either slightly smaller or much smaller than the distance between flowers and persons (top panel of Fig. 5) that had earlier produced the superordinate-before-basic trend (bottom panel of Fig. 3). Our new categories, cars plus persons and chairs plus flowers, each produced the superordinate-before-basic trend (Fig. 6), yielding the first experimental demonstration that between-category similarity affects the relative speeds of learning the basic level and superordinate level discriminations.

Although in this report we concentrated on between-category similarity, one might suggest that within-category similarity could

also explain our results. That is, perceptually coherent categories with high within-category similarity might be easier to discriminate at the basic level but be more difficult to group at the superordinate level, whereas categories with low within-category similarity might be more difficult to discriminate at the basic level but be easier to group at the superordinate level. Under this within-category similarity hypothesis, finding a basic level advantage for cars and chairs, but the opposite trend for persons and flowers (Lazareva et al., 2004) leads to the prediction that cars and chairs should have higher within-category similarity than persons and flowers. Therefore, we should see a higher percentage of complementary confusion errors for cars and chairs than for persons and flowers because their members will be more perceptually cohesive. In Experiment 1, we found a very different pattern of results (Table 2): namely, flowers had the highest percentage of confusion errors (49.8%) and chairs had the lowest percentage of confusion errors (35.3%), whereas cars and persons occupied intermediate positions (44.1% and 40.8%, respectively). Thus, within-category similarity is unlikely to provide a convincing explanation for our results.

To compare the results of Experiment 2 with our previously reported data, we calculated the number of sessions that were required for pigeons to reach $d's$ from 0.0 to 1.8 for each bird and for each category (cf. Figs. 3 and 6). We then subtracted the number of sessions required to reach criterion at the superordinate level from the number of sessions required to reach criterion at the basic level to obtain a measure of the basic level advantage. Negative values of the basic level advantage indicate that the basic level discrimination was acquired faster than the superordinate level discrimination, whereas positive values indicate that the superordinate level discrimination was acquired faster than the basic level discrimination.

The bottom panel of Fig. 5 illustrates the discrimination learning results of Lazareva et al. (2004) and Experiment 2 of the present study; for comparative purposes, the top panel of Fig. 5 illustrates between-category similarity from Experiment 1 of the present study. As the results of Experiment 1 suggested, cars and chairs were the two most dissimilar categories (cf. top panel of Fig. 5). Consistent with this result, the basic level advantage for the car–chair category was significantly below zero (two-tailed $t$-test, $t(27) = -3.46$, $p = 0.002$), documenting that the pigeons were faster to learn the basic level discrimination than the superordinate level discrimination between these categories. Experiment 1 further found that flowers and persons were significantly more similar to each other than cars were to chairs. And, for the flower–person category, the basic level advantage did not differ significantly from zero ($t(27) = 1.42$, df = 3, $p = 0.17$); here, the birds acquired the basic level discrimination and the superordinate level discrimination at similar speeds. Finally, cars and persons were slightly more similar to each other than were flowers and persons, whereas chairs and flowers were much more similar to each other than were flowers and persons. For both the car–person category and the chair–flower category, the basic level advantage was significantly greater than zero ($t(27) \geq 2.88$, $p \leq 0.0002$), documenting that pigeons learned the superordinate level discrimination before the basic level discrimination. It is perhaps surprising that, according to this final analysis, the car–person category yielded a significant basic level disadvantage, because its between-category similarity was only slightly greater than the flower–person category, which did not. Nonetheless, between-category similarity and the basic level advantage do indeed appear to be related in a systematic fashion as can be seen by comparing the top and bottom portions of Fig. 5.

To summarize, we found that between-category similarity affects the relative speeds of pigeons' learning basic level and superordinate level discriminations. It is possible that the basic-before-superordinate trend (or the superordinate-before-basic

trend) may be a byproduct of differential similarity between the basic level categories constituting a superordinate level category.

We believe that these data have important implications for understanding the contradictory results that have been reported in the developmental literature (e.g., Mandler and Bauer, 1988; Mareschal and Quinn, 2001; Oakes and Rakison, 2003; Rosch et al., 1976). Although there are many procedural disparities that could have affected the experimental outcomes of those studies (see Mareschal and Quinn, 2001 for a discussion), between-category similarity may be another important factor contributing to the occurrence of either the basic-before-superordinate or the superordinate-before-basic trend. Experimental results clearly demonstrate that basic level categorization in infants is affected by within-category similarity: in general, infants are better at basic level categorization when within-category similarity is reasonably low (Mareschal and Tan, 2007; Quinn and Eimas, 1998; Quinn et al., 1993; Ribar et al., 2003). The role of between-category similarity, however, has not yet been systematically explored. We hope that our data and analyses will generate fresh empirical and theoretical investigation into this important issue.

## Acknowledgments

## References

Astley, S.L., Wasserman, E.A., 1992. Categorical discrimination and generalization in pigeons: all negative stimuli are not created equal. Journal of Experimental Psychology: Animal Behavior Processes 18, 193–207.

Borg, I., Groenen, P.J.F., 2005. Modern Multidimensional Scaling: Theory and Applications, 2nd ed. Springer, New York, NY.

Bornstein, M.H., Arterberry, M.E., 2010. The development of object categorization in young children: hierarchical inclusiveness, age, perceptual attribute, and group versus individual analyses. Developmental Psychology 46, 350–365.

Bornstein, M.H., Arterberry, M.E., Mash, C., 2010. Infant object categorization transcends diverse object–context relations. Infant Behaviour and Development 33, 7–15.

Gibson, B.M., Wasserman, E.A., Frei, L., Miller, K., 2004. Recent advances in operant conditioning technology: a versatile and affordable computerized touch screen system. Behavior Research Methods, Instruments and Computers 36, 355–362.

Heady, R.B., Lucas, J.L., 1997. PERMAP: an interactive program for making perceptual maps. Behavior Research Methods, Instruments, & Computers 29, 450–455.

Jolicoeur, P., Gluck, M.A., Kosslyn, S.M., 1984. Pictures and names: making the connection. Cognitive Psychology 16, 243–275.

Kaiser, D.H., Sherburne, L.M., Steirn, J.N., Zentall, T.R., 1997. Perceptual learning in pigeons: decreased ability to discriminate samples mapped onto the same comparison in many-to-one matching. Psychonomic Bulletin and Review 4, 378–381.

Lazareva, O.F., Freiburger, K.L., Wasserman, E.A., 2004. Pigeons concurrently categorize photographs at both basic and superordinate levels. Psychonomic Bulletin and Review 11, 1111–1117.

Lazareva, O.F., Wasserman, E.A., 2009. Effects of stimulus duration and choice delay on visual categorization in pigeons. Learning and Motivation 40, 132–146.

Lee, M.D., 2001. Determining the dimensionality of multidimensional scaling representations for cognitive modeling. Journal of Mathematical Psychology 45, 149–166.

Lin, E.L., Murphy, G.L., Shober, E.J., 1997. The effects of prior processing episodes on basic level superiority. Quarterly Journal of Experimental Psychology 50A, 25–48.

Mandler, J.M., Bauer, P.J., 1988. The cradle of categorization: is the basic level basic? Cognitive Development 3, 247–264.

Mandler, J.M., Bauer, P.J., McDonough, L., 1991. Separating the sheep from the goats: differentiating global categories. Cognitive Psychology 23, 263–298.

Mareschal, D., Quinn, P.C., 2001. Categorization in infancy. Trends in Cognitive Sciences 5, 443–450.

Mareschal, D., Tan, S.H., 2007. Flexible and context-dependent categorization by eighteen-month-olds. Child Development 78, 19–37.

Mareschal, D., Tan, S.H., 2008. The role of context in the categorization of hybrid toy stimuli by 18-month-olds. Infancy 13, 620–639.

Murphy, G.L., Brownell, H.H., 1985. Category differentiation in object recognition: typicality constraints on the basic category advantage. Journal of Experimental Psychology: Learning, Memory, and Cognition 11, 70–84.

Murphy, G.L., Smith, E.E., 1982. Basic level superiority in picture categorization. Journal of Verbal Learning and Verbal behavior 21, 1–20.

Oakes, L.M., Rakison, D.H., 2003. Issues in the early development of concepts and categories: an introduction. In: Rakison, D.H., Oakes, L.M. (Eds.), Early Category and Concept Development: Making Sense of the Blooming, Buzzing Confusion. Oxford University Press, New York, pp. 3–23.

Quinn, P.C., 2002. Early categorization: a new synthesis. In: Goswami, U. (Ed.), Blackwell Handbook of Childhood Cognitive Development. Blackwell Publishing, Malden, MA, pp. 84–101.

Quinn, P.C., Eimas, P.D., 1998. Evidence for a global categorical representation of humans by young infants. Journal of Experimental Child Psychology 69, 151–174.

Quinn, P.C., Eimas, P.D., Rosenkrantz, S.L., 1993. Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. Perception 22, 463–475.

Rakison, D.H., Butterworth, G., 1998. Infants' use of object parts in early categorization. Developmental Psychology 34, 49–62.

Ribar, R.J., Oakes, L.M., Spalding, T.L., 2003. Infants can rapidly form new categorical representations. Psychonomic Bulletin and Review 11, 536–541.

Rosch, E., Mervis, C.B., 1975. Family resemblances: studies in the internal structure of categories. Cognitive Psychology 7, 573–605.

Rosch, E., Mervis, K.B., Gray, W.D., Johnson, D.M., Boyes-Braem, P., 1976. Basic objects in natural categories. Cognitive Psychology 8, 382–439.

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

Smith, J.E.K., 1982. Simple algorithms for M-alternative forced-choice calculations. Perception and Psychophysics 31, 95–96.

Swets, J.A., 1964. Indices of signal detectability obtained with various psychophysical procedures. In: Swets, J.A. (Ed.), Signal Detection and Recognition by Human Observers: Contemporary Readings. John Wiley & Sons, New York, pp. 165–171.

Tanner, W.P., Birdsall, T.G., 1964. Definitions of $d'$ and $\eta$ as psychophysical measures. In: Swets, J.A. (Ed.), Signal Detection and Recognition by Human Observers: Contemporary Readings. John Wiley & Sons, New York, pp. 147–163.

Urcuioli, P., Vasconcelos, M., 2008. Effects of within-class differences in sample responding on acquired sample equivalence. Behavioural Processes 89, 341–358.

Wasserman, E.A., DeVolder, C.L., Coppage, D.J., 1992. Nonsimilarity-based conceptualization in pigeons via secondary or mediated generalization. Psychological Science 3, 374–378.