The role of item- and category-specific information in the discrimination of people versus nonpeople images by pigeons

ULRIKE AUST and LUDWIG HUBER
University of Vienna, Vienna, Austria

Herrnstein and Loveland (1964, pp. 549–551) successfully trained pigeons to discriminate pictures showing humans from pictures that did not. In the present study, a go/no-go procedure was employed to replicate and extend their findings, the primary focus of concern being to reevaluate the role of item- and category-specific information. The pigeons readily acquired the discrimination and were also able to generalize to novel instances of the two classes (Experiment 1). Classification of scrambled versions of the stimuli was based on small and local features, rather than on configural and global features (Experiment 2). The presentation of gray-scale stimuli indicated that color was important for classifying novel stimuli and recognizing familiar ones (Experiments 1 and 2). Finally, the control that could possibly be exerted by irrelevant background features was investigated by presenting the pigeons with images of persons contained in former person-absent pictures (Experiment 3). Classification was found to be controlled by both item- and category-specific features, but only in pigeons that were reinforced on person-present pictures was the latter type of information given precedence over the former.

In a pioneering study, Herrnstein and Loveland (1964) showed that pigeons were able to respond to complex *real-scene* photographs according to the presence or absence of human beings—a class of stimuli so diverse that, according to the authors, it precluded simple characterization. The results—extremely quick discrimination and complete transfer—and the failure to detect any trivial visual clues, such as wavelength, intensity, or frequency (Honig & Urcuioli, 1981; Lubow, 1974; Mostofsky, 1965), indicated that the pigeons readily came to respond to the *target* that we would ourselves call a person. Consequently, Herrnstein and Loveland, as well as the authors of subsequent studies on the discrimination of pictures according to the presence or absence of people (Greene, 1983; Malott & Siddall, 1972; Poole & Lander, 1971; Siegel & Honig, 1970), claimed that classification was due to a strategy beyond discrimination of simple stimulus dimensions or memorizing individual exemplars by rote. However, the lack of evidence regarding which features were actually used by the pigeons to solve the concept task justifies further examination.

The features that control the birds' classification behavior have hitherto been successfully determined mostly in experiments using carefully constructed sets of artificial stimuli such as letters (Blough, 1982), line drawings (Cerella, 1980), pseudoseeds (Lea, Lohmann, & Ryan, 1993), Brunswik faces (Huber & Lenz, 1993), or pseudobutterflies (Jitsumori, 1996). In contrast, attempts to determine which aspects of a complex picture, or class of complex pictures, such as those found in nature, do acquire control over the pigeons' responding have been less successful (Fetterman, 1996). In most discrimination tasks concerned with natural categories, it is difficult, if not impossible, to specify the defining features of the classes, because the separation of classes is based on some set of features that correlate to some extent with the category. A polymorphous feature rule, in which no feature is a singly necessary or jointly sufficient condition, applies to most naturally occurring perceptual categories.

In the *Herrnsteinian* concept discrimination tasks, however, a clearer distinction between category-relevant and category-irrelevant features can be made. Here, the *concept*—for example, *human figure*—is present in one set of images and absent in the second. Nevertheless, in the concept-present class, the critical information is usually embedded in a visual array that contains many conceptually irrelevant features. The critical elements in the stimulus array—that is, the features relevant to the concept—obtain some coherence and constitute the *target* (Greene, 1983). Therefore, the people/nonpeople discrimination problem may be seen as a target search task or a figure–ground separation task.

In general, theories of categorization have always been determined by the specification of the *what* of information that is acquired through encounters with examples of the category. Candidates for the *what* of information in-

clude prototypes, feature lists, complex hypotheses or descriptions, best examples, or even the exemplars themselves. More specifically, exemplars are composed of category-level information plus idiosyncratic, or item-specific, information (Medin, Dewey, & Murphy, 1983). This distinction was already claimed by Restle (1957), who argued that in concept identification tasks, idiosyncratic information is rendered nonfunctional during learning in an abstraction process. Alternatively, knowledge about categories can be acquired implicitly by cumulating information from multiple examples. Category-level knowledge would then have no special status but would emerge naturally from item memory.

Shortly put, two different types of information may enter into the representation that is formed during a concept discrimination task, and categorization can be accomplished by relying on either the one or the other (Medin et al., 1983). On the one hand, the animals may memorize item-specific features of the training stimuli and generalize to novel stimuli on the basis of psychophysical similarity to features of past exemplars of the reinforced and nonreinforced classes (Flannagan, Fried, & Holyoak, 1986). This we may call an *item-specific*, or *nonanalytic*, strategy. On the other hand, they may abstract the category-specific features, perhaps including relational information about common properties of the embedded concept, and then react in the same way to novel stimuli possessing these features (Cook, Wright, & Kendrick, 1990). This we may call a *categorical*, or *analytic*, solution.

In practice, however, few experiments have to date been designed so that the collected data may allow a distinction between item- and category-level information. The present set of experiments addressed that issue. In particular, the aim was to determine which of the two types of information is gathered and stored by the pigeons during people/nonpeople category training and which information is then used to classify novel instances.

Furthermore, we sought to specify the informational content of the internal representation. The term *representation* describes a state of the visual system that stands for a property or object of the environment. It is thus a model of what it represents. A representational system includes the represented world outside the information-processing system (external world, environment), as well as the representing world within the information-processing system (internal representation; see Palmer, 1999, p. 77). We attempted to narrow down, in a stepwise fashion, the possible representations that might be acquired in the classic task of sorting real-scene pictures according to the presence or absence of humans. Our alternative hypothesis to a rather sophisticated classification rule, such as that applied by humans, would be that pigeons use some independent simple features. Given that, inadvertently, some simple physical dimension provides a good basis for discriminating between people-present and people-absent pictures, it is possible that pigeons make use of it and thereby master the task without

responding to the experimenter-defined (semantic) class rule (see Monen, Brenner, & Reynaerts, 1998). In fact, this possibility deserves attention, since pigeons were found to quickly discriminate between male and female human faces on the basis of simple perceptual dimensions (Troje, Huber, Loidolt, Aust, & Fieder, 1999). More specifically, overall brightness and color served as classification cues.

Using a similar experimental design and modern technical equipment, we replicated and extended the experiment described in Herrnstein and Loveland (1964). After training, we carried out two generalization tests (Experiment 1) to obtain first clues regarding the features involved in controlling responding behavior during training. These preliminary experiments were followed by several tests that investigated in more detail the relevance of particular feature domains for classification. First, the relative importance of features that remain unaffected through picture decomposition was examined (Experiment 2). That experiment followed the assumption that if the human figures are progressively fragmented, the critical category-specific information is gradually destroyed as well. Consequently, test performance should suffer to the same extent—on condition that the pigeons' training behavior was exclusively controlled by category-specific information. Also, the role of color, which was supposed to be a likely candidate for governing classification, was investigated in Experiments 1 and 2. Finally, we conducted a test similar to those designed by Greene (1983) to find out whether the pigeons relied on category-irrelevant background features or whether classification was definitely coupled to the presence or absence of people (Experiment 3).

EXPERIMENT 1

Experiment 1 was very similar to the experiment described in Herrnstein and Loveland (1964). The training procedure required the pigeons to classify pictures according to the presence or absence of people. Apart from providing a basis for further tests, the training also served as a control for possible effects of procedural differences between our experimental design and the one used by Herrnstein and Loveland, since comparable levels of performance would point to the irrelevance of differences in the discrimination procedures and the stimulus materials for the acquisition of the present task.

Given the relatively small number of training stimuli in combination with the extraordinary memorial capacities of pigeons (Vaughan & Greene, 1984), the possibility of rote learning cannot be ruled out. A common way to assess the generalizability of the acquired classification is to determine how readily the birds respond to novel exemplars of the two classes. In order to investigate whether the pigeons were able to generalize to novel stimuli, a generalization test with novel, full-color stimuli of both classes was conducted (Test 1). A second generalization

test with gray-scale stimuli was carried out to determine the contribution of color properties to the classification of novel stimuli (Test 2).

Method

Subjects. In contrast to Herrnstein and Loveland (1964), who used 5 racing pigeons, we employed 4 pigeons of a local Austrian race (*Columba livia*, "Strasser") as experimental subjects. In previous experiments (e.g., Huber, 1994; Huber, Aust, Michelbach, Ölzant, & Nowotny, 1999; Huber & Lenz, 1993; Troje et al., 1999) that race of pigeons had been shown to be more balanced, as well as more eager to peck, than racing pigeons, which were generally less tame and more difficult to handle. Water and grit were freely available in the aviary, but food was provided only in the learning boxes and after the experimental sessions. The birds were maintained at approximately 90% of their free-feeding weight. All the pigeons were experimentally naive to the task at the onset of Experiment 1.

Stimuli. In the study by Herrnstein and Loveland (1964), a total of over 1,200 35-mm color slides were presented at a rate of 80 or 81 per session. Approximately half the photographs shown in each session contained at least one human being, whereas the remainder did not. In the present study, 80 full-color photographs served as training stimuli. They were taken from the database of the PHOTODISC Starter Kit and the PHOTODISC collection (watermarked; www.photodisc.com). Half of the stimuli showed one or more persons, whereas the other half did not. As in the original study, the pictures involved a variety of settings, and the people contained in the people-present stimuli varied with respect to number, identity, sex, race, age, size, and apparel. Furthermore, they differed from each other with regard to their position within the picture, their posture and the context in which they were acting, the degree to which parts of their bodies were obscured by intervening objects, and the angle of regard. Some pictures showed close-ups; in others, the people appeared at a distance. Examples of the stimuli used in Experiment 1 are shown in Figure 1A. In the generalization tests, 160 test stimuli were employed, half of which were presented in full color (Test 1), whereas the other half were presented as gray-scale versions (Test 2). In each test, half of the test stimuli contained persons, whereas the other half did not. Examples of the test stimuli are shown in Figure 1B.

Apparatus. The birds entered an experimental indoor chamber (Skinner box) from an outdoor aviary compartment through a passway system (Huber, 1994). The interior size of the wooden chamber was $50 \times 30 \times 40$ cm. In the center of the front panel, there was a clear perspex pecking key (5-cm diameter; ENV-125M, MED Associates). Directly below the key there was a 6×6 cm aperture for a 28-V DC solenoid activated hopper featured by the grain feeder (ENV-205M). A hopper light illuminated the receptacle area whenever grain was accessible. Throughout the experiments (except prior to stimulus presentation), the chamber was weakly illuminated by a 2-W houselight (ENV-215) located in the rear part of the chamber. In contrast to the experiment by Herrnstein and Loveland (1964), stimulus presentation was not accomplished by slide projectors but was controlled by a Pentium PC running a Delphi program that selected the patterns according to a prespecified sequence. Stimuli were presented on a 15-in. (38-cm, diagonal) PC monotor (Panasonic PanaSync 4G) at a distance of 5 cm behind the pecking key. They were presented at a size of 128×128 pixels at a resolution of $1,024 \times 768$ pixels, thus producing a 45×45 mm picture on the monitor. The MED behavioral control package was used to control events in the operant chamber, to signal onset and offset of presentation to the presentation PCs, and to register responses.

Procedure. The pigeons were first trained to enter the box voluntarily, to find the food hopper, and to peck for food at the illuminated key. Then they were transferred to a standard go/no-go procedure (see, e.g., Vaughan & Greene, 1984), which required them to peck in the presence of positive stimuli and to withdraw from pecking in the presence of negative stimuli. In this respect, we deviated from the original design by Herrnstein and Loveland (1964), who used an intermittent schedule of reward for pecks emitted in the presence of a positive stimulus. Whereas Herrnstein and Loveland had only one experimental group, with all the birds being required to peck in the presence and withhold pecking in the absence of "people," we used a counterbalanced design. The pigeons were arbitrarily assigned to one of two groups, each of which consisted of 2 birds. For one group (Group P: Birds T71-P and T72-P), the stimuli containing at least one person (Class P) were designated positive, and the stimuli that were devoid of persons (Class NP) were designated negative. These contingencies were reversed for the other group of pigeons (Group NP: Birds T73-NP and T77-NP).

In the experiment by Herrnstein and Loveland (1964), slides were changed each day, but many of them were used again in later sessions. The experiments lasted for several months and obviously included more than 70 sessions, but nothing is said about the exact number of sessions that were carried out. In the present experiment, training involved a total of 12 sessions, with each bird running one session per day, 5 days a week. Each session consisted of a single presentation of all 80 training stimuli appearing in random order. The entire sequences were quasirandom, insofar as they never contained more than three stimuli of the same contingency in immediate succession and the first stimulus of each session was always a positive one. Pecks were counted only during the first 10 sec of a trial. The image remained visible (without any further pecks being counted) at least during a further variable interval (mean, 10 sec; range, 1–20 sec) and then until the response requirement was accomplished. If the stimulus was positive, two pecks emitted within 2 sec resulted in 5 sec of food access. By contrast, responding to a negative stimulus was followed by an error delay, which terminated only after no further responses had occurred within 8 sec. In negative trials, no food was delivered. Each trial was followed by an intertrial interval, a dark phase of 4 sec that signaled the forthcoming stimulus presentation (Troje et al., 1999).

In the generalization tests, the test stimuli were interspersed into sequences of normal training stimuli at a rate of 16 per session, thereby replacing an equal number of arbitrarily selected training stimuli. Each test thus consisted of five consecutive sessions. All the test stimuli of Test 1 were presented with neutral contingency. This means that the respective trial resulted neither in food access nor in a delay interval but was terminated immediately after 10 sec of stimulus presentation, during which pecks were counted. Test 1 was followed by a final session, which consisted of a single presentation of each of the 80 transfer stimuli only. Therefore, all the stimuli had training contingency in the final session. For Test 2, we suspected that the absence of color would make the test stimuli very easily distinguishable from the training stimuli. The gray-scale pictures were thus presented with training contingency in order to counteract the pigeons' previously observed tendency to respond similarly to all stimuli recognized as *novel*.

Results

Generally, a sample size of only two in each group poses a serious problem for statistical analyses, which are based on intergroup comparisons. Therefore, we decided to assess performance in the present experiment, as well as in all subsequent experiments, separately for each bird and not to draw any conclusions about differences between the two groups unless *both* birds of each group performed similarly to each other, but different from the birds of the other group. However, for reasons of clearness, we decided to depict performance separately not for
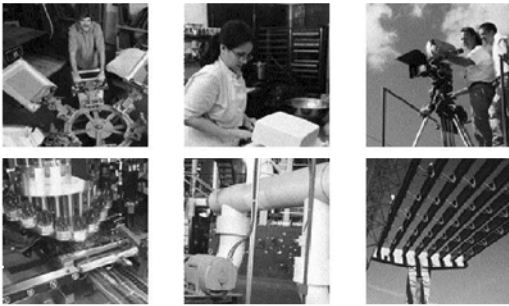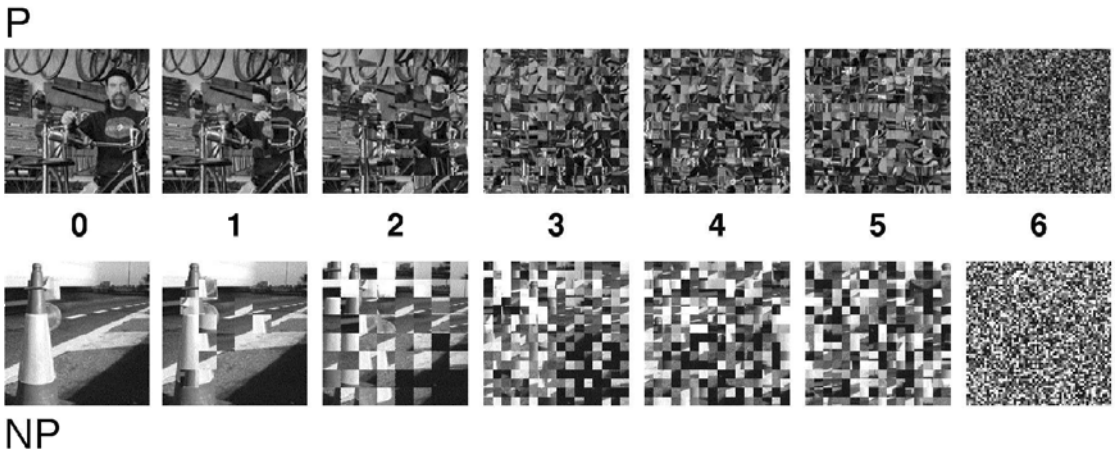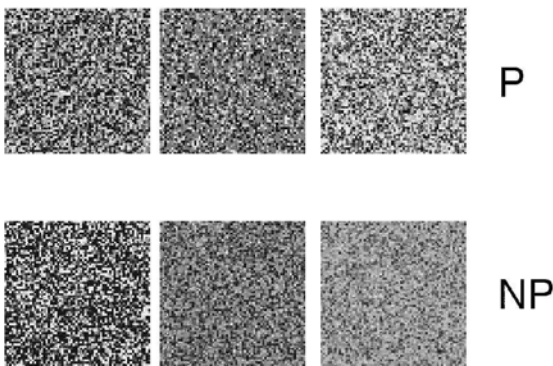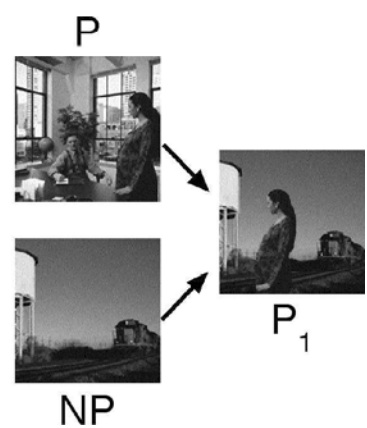
Figure 1. (A) Selection of training stimuli of both classes from Experiment 1. (B) Selection of test stimuli from Test 1 of Experiment 1. (C) One stimulus of each class shown in all scramble degrees (0–6) of Experiment 2. (D) Selection of stimuli of both classes of Degree 6 of Experiment 2. (E) Example of a stimulus of Class P_1 and the original pictures from which it was derived.
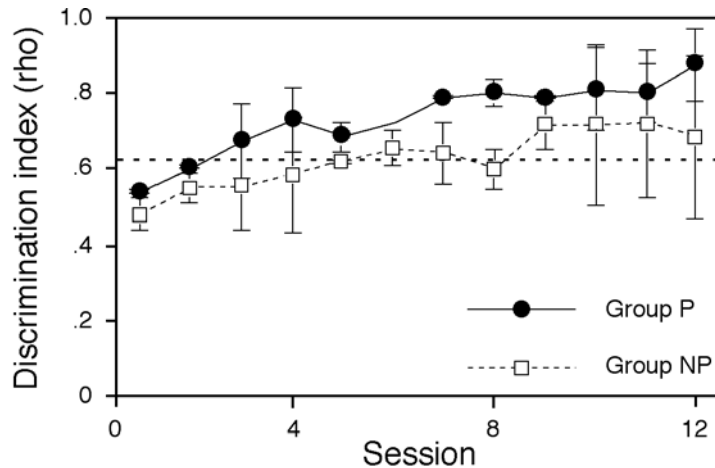
Figure 2. Performance of both groups (P and NP) in training of Experiment 1, shown on a session-by-session basis as ρ values. Means were taken across the 2 birds of each group. The dashed line indicates the limit of significance (.662).

each bird, but for each group in the figures. Since in most cases, there turned out to be no significant differences between the birds belonging to the same group, it seemed reasonable to take means across them and indicate interindividual within-group variation by the respective standard deviations only.

Assessment of performance during training was based on rho values (ρ). This is a nonparametric statistic introduced by Herrnstein, Loveland, and Cable (1976) that estimates the probability that the average positive exemplar is ranked above the average negative exemplar, using ranked rates of pecking as the measure of performance. We chose that measure for reasons of comparison, since the results presented in the study by Herrnstein and Loveland (1964) were also based on ranked data. However, in the generalization tests, as well as in all further experiments, calculations were based instead on the mean standardized response rate, which is a more demanding measure than the ρ value. It was obtained by dividing the peck rate on each trial of a session by the average peck rate measured over all the trials (except test trials) of the respective session. Then, pecks on all the trials of the session—training as well as test trials—were divided by the mean peck rate.

Training performance is depicted separately for the two groups on a session-by-session basis in Figure 2. All the pigeons readily learned to discriminate between the two classes (limit of significance, $\rho = .620$) and had reached a reasonable level of performance by the end of training. The results of the generalization tests are illustrated in Figure 3 as mean standardized response rates, shown separately for the two groups and in comparison with the first presentation of the respective training stimuli.

For each test, we carried out a two-sample t test comparing performance on the positive and the negative test stimuli. In Test 1 (full color), all the birds discriminated significantly between the positive and the negative test stimuli (all ps $\leq$ .05). There were significant differences neither between the 2 birds of each group nor between any bird of Group P compared with any bird of Group NP. Although the pigeons showed good transfer to the novel color stimuli, Figure 3 seems to indicate that performance was nevertheless slightly worse than on the training component. This was also reflected by the corresponding ρ values, which were .933 for Group P and .884 for Group NP on the first presentation of the training stimuli and .729 for Group P and .748 for Group NP on the test stimuli (averaged across the birds of each group, since they performed quite similarly; limit of significance, .626). That difference was probably due to long-term learning effects in the case of the training stimuli. Performance further improved in the final session, in which the 80 transfer stimuli were shown with training contingency (all ps $\leq$ .05). The respective ρ values were .924 for Group P and .768 for Group NP (averaged across the birds of each group).

In case of Test 2 (gray-scale), Figure 3 seems to suggest a difference in responding between the two groups. In fact, both birds of Group P [T71-P, $t(77) = -2.96$, $p \leq .05$; T72-P, $t(77) = -3.56$, $p \leq .001$], but no bird of Group NP discriminated significantly between the positive and the negative test stimuli. There were no significant differences between the 2 birds of either group. The corresponding ρ values were .997 for Group P and .998 for Group NP on the training component and .745 for Group P and .571 for Group NP on the test component (limit of significance, .626). Thus, both groups performed extraordinarily well on the training component, but in the case of the test component, only Group P achieved a discrimination above the criterial level, whereas Group NP clearly performed below the criterion value. Performance on the test component in Group P was comparable with that in Test 1, with peck rates on neither stimulus class differing significantly between Tests 1 and 2. In Group NP, however, performance
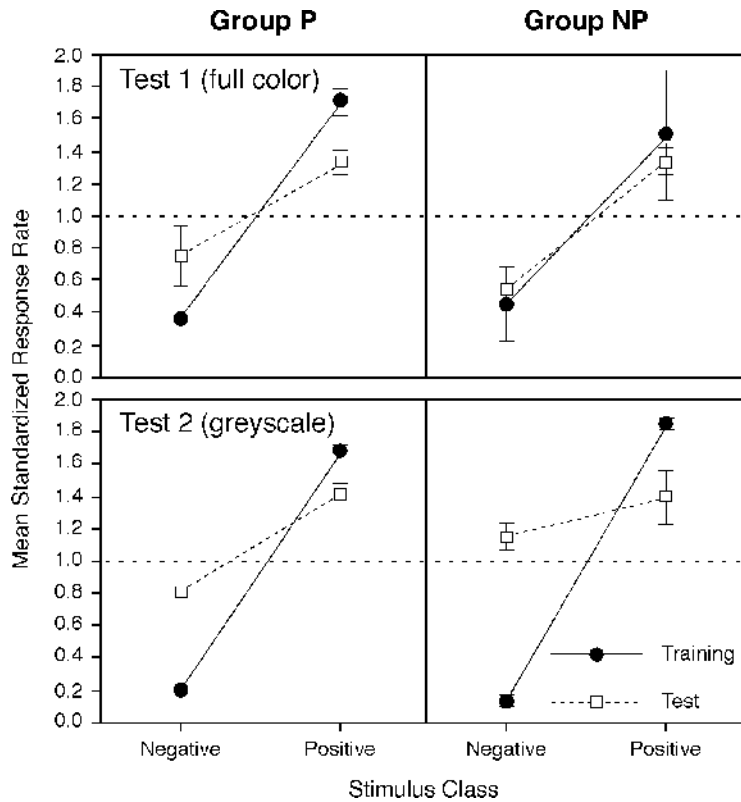
Figure 3. Generalization performance in Experiment 1, shown separately for the two groups and in comparison with the first presentation of the respective training component as mean standardized response rates ($\pm SD$). Means were taken across the birds of each group.

on the test component was worse in Test 2 than in Test 1 in the case of the negative stimulus class [$t\,(77) = -3.44$, $p \leq .001$].

Discussion

Experiment 1 confirmed the finding of Herrnstein and Loveland (1964) that pigeons readily learn to discriminate pictures characterized by the presence or absence of people and also generalize to novel instances of the two classes. The acquisition of the task was obviously not adversely affected by the use of a procedure that differed in several respects from the one employed by Herrnstein and Loveland.

The results of Test 1 show that the birds were well able to classify novel color stimuli correctly, which were characterized by either the presence or the absence of human beings. This seems to indicate that the pigeons exploited category-specific information. However, a considerable amount of positive transfer to novel instances could have been due to stimulus generalization on the basis of item similarity between previously learned exemplars and novel stimuli (Cook et al., 1990; D'Amato & Van Sant, 1988; Greene, 1983; Lea, 1984). Successful transfer per se is not indicative of what information had entered the pigeons' representation of the training stimuli.

Group P performed well on both novel color (Test 1) and gray-scale (Test 2) stimuli, whereas in Group NP performance was worse on the latter than on the former. Given the differential responding behavior of the two groups in Test 2, it is difficult to reach any general conclusions about the importance of color for the present task. The results of Group NP do indicate that color was used as an important source of information for classifying novel stimuli in this group. The birds may, for instance, have exploited color cues typically present in pictures containing humans (e.g., the color of the skin). Alternatively, color in general may have contributed to structuring the pictures by setting boundaries and thereby facilitated singling out the *target* (i.e., a person; see Jacobs, 1993; Mollon, 1989; Wurm, Legge, Isenberg, & Luebker, 1993). As for Group P, it is not quite clear whether color was generally unimportant for classifying novel stimuli or whether the pigeons used a classification code that was sufficiently wide to compensate for the loss of color information.

EXPERIMENT 2

Experiment 1 showed that pigeons are able to classify pictures according to the presence or absence of human

beings and that this ability also generalizes to novel stimuli. However, it is not yet obvious what information a pigeon might be able to extract about such stimuli. In order to elucidate the nature of the representation, we chose an approach of *progressive elimination* in which we stepwisely excluded specific stimulus properties. First, we examined the importance of *global*, as compared with *local*, features contained in the stimuli.

As we will use the term, global features are characteristics of the whole stimulus, such as size, orientation, and brightness. They roughly correspond to the type of features that are computed automatically and in parallel—that is, at every location in the visual field at the same time (see, e.g., Theeuwes, 1996; Treisman & Gelade, 1980). A visual scene may quickly (*preattentatively*) be coded along a number of such dimensions that elicit separate representations.

In contrast, distinct stimulus locations may be processed serially—that is, one item and one location at a time—by driving an attentional focus through the visual field. Sequentially focusing attention on different parts of the display may occur in an attempt to find differences between individual stimulus elements. For instance, by applying a selective, analytic process to the images, the stimulus array may be decomposed into objects and background. In addition, complex objects may themselves be decomposed into distinct parts. In this sense, local features are restricted portions of an image that have semiautonomous, objectlike status in visual perception (Palmer, 1999). A human body, for example, is perceived as being composed of a head, a torso, two arms, and two legs. Object perception includes the spatial relations among the parts—that is, its *coherence*—for a disassembled pile of legs, arms, and heads does not look even remotely the same as an intact human figure.

It would be interesting to know whether the features that controlled discrimination and generalization in Experiment 1 were global or local, in the usage of the terms given above. Without this information, inferences about the nature of the representation formed during people/nonpeople discrimination are severely limited. Arguably, finding that pigeons master the categorization task by using only global features, instead of responding to the target *human figure*, would also have important implications for the processes underlying people/nonpeople categorization.

The strategy we used in Experiment 2 was to examine the effects of stimulus fragmentation. Pictures of both classes were increasingly scrambled so that many global properties of the images (e.g., overall brightness) remained intact, whereas the spatial relationships among stimulus elements, as well as coherent objects (e.g., human figures) or objectlike parts (e.g., faces), were gradually destroyed.

The effects of fragmentation were determined by presenting the pigeons test stimuli that increased in the degree of scramble. In the extreme case, the pictures were divided into 4,096 two-by-two-pixel squares, which were then distributed arbitrarily over the whole picture area. To investigate the possibility that transfer to the scrambled pictures was mediated by item similarity to previously learned exemplars, the birds were shown scrambled versions of familiar stimuli, as well as scrambled versions of completely novel pictures. The role of color for classifying scrambled familiar pictures was investigated by presenting the pigeons with scrambled gray-scale versions of such stimuli.

Method

Subjects. The experimental subjects were the same as those in Experiment 1.

Stimuli. The training stimuli were the same as those in Experiment 1. As test stimuli, we generated pictures of six degrees of progressive scrambling. Up to Degree 2, the pictures were generated manually in Photoshop (Adobe Systems). Higher scramble degrees were generated by means of a MATLAB (The Math Works) computer program. Figure 1C shows one stimulus of each class on all scramble degrees.

The individual scramble degrees were defined by the following rules. The unscrambled originals, presented prior to tests with their scrambled derivatives in Sets 3 and 4, had a scramble degree of 0. For Scramble Degrees 1–6, the pictures were divided into squares of 16×16 pixels (Degrees 1 and 2), 8×8 pixels (Degrees 3, 4, and 5), or 2×2 pixels (Degree 6). In addition to square size, the scramble degrees were defined by the distribution of the squares over the picture. Thus, degrees characterized by the same square size differed from each other with respect to the rules according to which the squares were arranged. On Degree 1, 8 squares were displaced arbitrarily, but the 28 marginal squares were always excluded from dislocation. On Degree 2, every second square of every second row changed place with its right diagonal neighbor from the row directly below. On Degree 3, the squares were arranged arbitrarily within each picture quadrant only, and on Degree 4 only within the upper and the lower half of the picture. On Degrees 5 and 6, all the squares were arranged arbitrarily.

Apparatus. The apparatus was the same as that in Experiment 1.

Procedure. We presented the pigeons with scrambled versions of stimuli belonging to four different stimulus sets. First, they were shown scrambled derivatives of the 80 training stimuli (Set 1, *familiar*) and of the 80 test stimuli from Experiment 1 (Set 2, *semifamiliar*). In both sets, the scrambled stimuli were interspersed into sequences of normal training stimuli from Experiment 1 at a rate of 10 per session in the cases of Degrees 1 and 2 and at a rate of 16 per session in the cases of Degrees 3–6. (Since it turned out that the presentation of scrambled stimuli did not disrupt performance, the number of test stimuli per session was increased, in order to accelerate the progress of the experiment). Set 3 (*novel*) involved the presentation of completely unfamiliar stimuli. The pigeons were shown 20 stimuli of each scramble degree (1–6), as well as 20 unscrambled stimuli (this condition was labeled as Degree 0). The test set therefore consisted of 140 different stimuli, each of which occurred only once and in only one particular scramble degree (0, 1, 2, 3, 4, 5, or 6). The test stimuli were interspersed into sets of training stimuli at a rate of 10 per session. Finally, we presented the pigeons with gray-scale versions of a random assortment of 20 stimuli from the training set used in Experiment 1 (Set 4, *familiar gray-scale*). Those 20 stimuli were presented unscrambled (Scramble Degree 0) and in all six scramble degrees (1–6) at a rate of 10 per session. As the characteristic appearance of scrambled stimuli probably made them easily distinguishable from the originals, all the scrambled stimuli of all the sets were subjected to the same re-

inforcement schedule as the training stimuli. Figure 1D shows examples of the stimuli of Class P and Class NP of Degree 6. The depicted stimuli were taken from the training set and were presented in color in the case of Set 1 and in gray-scale in the case of Set 4.

Results

Figure 4 shows the results of Experiment 2 as mean standardized response rates to the positive and the negative test stimuli as a function of scramble degree, separately for the two groups. Variations between the 2 birds of each group are indicated by the standard deviations. In the case of Set 3 (*novel*) and Set 4 (*familiar grayscale*), the pigeons were also presented with unscrambled originals of the same type as the test stimuli (Degree 0). To provide a baseline value, performance on the first presentation of the respective training stimuli (tr) is depicted as well.

Performance was assessed by computing for each set and scramble degree a two-sample t test comparing performance on the positive and on the negative test stimuli. In Set 1 (*familiar*), 2 birds performed at a highly significant level up to Degree 5 [Bird T71–P, $t(77) = -4.84$, $p \leq .0001$; Bird T72-P, $t(66) = -5.36$, $p \leq .0001$]. At Degree 6, discrimination of Birds T71-P and T72-P was no longer significant. Birds T73-NP and T77-NP performed significantly at all degrees (with the exception of Degree 5 in the case of Bird T77-NP), even at Degree 6 [Bird T73-NP, $t(76) = -2.21$, $p \leq .05$; Bird T77-NP, $t(77) = -3.01$, $p \leq .001$]. In conclusion, all the pigeons were well able to classify even strongly scrambled familiar stimuli correctly, although progressive scrambling resulted in a reduction of discrimination (see Figure 4).

A similar pattern of results was found for Set 2 (*semifamiliar*). Bird T71-P performed significantly up to Degree 6 [$t(70) = -4.30, p \leq .0001$], Birds T72-P and T77-NP up to Degree 5 [T72-P, $t(65) = -4.94$, $p \leq .0001$; T77-NP, $t(77) = -2.52$, $p \leq .05$]. Discrimination performance of Bird T73-NP was significant at all degrees (except at Degree 4), up to Degree 6 [$t(77) = -2.12, p \leq .05$]. Discrimination performance was higher for both birds in Group P than for both birds of Group NP, except at Scramble Degree 6. Overall, the pigeons mastered the task quite well, with the birds of Group P doing somewhat better than those of Group NP. Nevertheless, as in Set 1, performance decreased as scramble degree increased (see Figure 4).

The pigeons performed considerably less well when tested with scrambled versions of novel stimuli (Set 3) or familiar gray-scale stimuli (Set 4). In the case of Set 3, all the birds performed poorly from low scramble degrees onward. Discrimination performance of Birds T71-P, T72-NP, and T77-NP was (with the exception of Degree 1 in the case of Bird T77-NP and of Degree 2 in the case of Birds T71-P and T72-P) significant up to Degree 3 [Bird T71-P, $t(17) = -1.77$, $p \leq .05$; Bird T72-P, $t(13) = -5.22, p \leq .001$; Bird T77-NP, $t(13) = -2.34, p \leq .05$]. Performance of Bird T73-NP was never significant, except at Degree 5 [$t(17) = -3.03, p \leq .01$].

In Set 4, statistical analysis failed to reveal significant discrimination at any scramble degree for Birds T71-P, T72-P, and T73-NP. Bird T77-NP performed significantly at least on the unscrambled originals (Degree 0) and at Degree 1 [$t(13) = -3.40, p \leq .01$]. There were hardly any inter- and intragroup differences between the individual birds.

Discussion

The key findings to emerge from Experiment 2 were the following. (1) Overall, there was an effect of scrambling; progressive scrambling resulted in a reduction of discrimination ability in all four stimulus sets. (2) The effect of scrambling depended on the familiarity of the test stimuli. Although the pigeons were well able to discriminate even strongly scrambled versions of familiar (Set 1) and semifamiliar (Set 2) stimuli, they performed rather badly when presented with scrambled novel stimuli (Set 3). (3) Performance was worst when the test stimuli were gray-scale versions of scrambled training stimuli.

What do these results tell us about the pigeons' representation of the people-present or the people-absent class? First, it cannot be ruled out completely that the decline in performance was due to some kind of alienation effect caused by scrambling itself. However, we found no evidence for such an effect. It should, for example, be noted that one finding common to all four sets was the fact that deterioration of performance with increasing scramble degree was mainly due to increasing peck rates on the negative stimuli, whereas peck rates on the positive stimuli dropped only slightly or not at all. It is thus evident that the pigeons were at least not taken aback by the unusual appearance of the stimuli.

Second, the decline in performance indicates that the pigeons' identification of positives could not depend on the computation of global features that remained intact across all six scramble degrees. This finding supports Herrnstein and Loveland's (1964) basic assumption that the precision of the pigeons' classification performances did not arise from some trivial and unsuspected visual clue in the slides. However, whereas Herrnstein and Loveland only checked the possibility of some correlation between the presence of a human being and color distribution in the slides, we expanded the proof to a wide range of global stimulus properties.

An alternative to relying on global features would be the use of local stimulus properties. This would account for correct classification of moderately scrambled versions of familiar stimuli (Sets 1 and 2), in which the crucial features were still intact. Such *particulate perception* would also be in agreement with Cerella's (1980) finding that pigeons treated partial views, as well as scrambled versions, of a Charlie Brown figure in a way that was similar to that for complete, intact instances of the respective category.

However, if the pigeons' representation of people/nonpeople consisted of local features, why were they able to correctly classify the test stimuli of Sets 1 and 2 not

Set 1 ("familiar")

Group P   Group NP

S+
S−

Mean Standardized Response Rate

Scramble Degree

Set 2 ("semifamiliar")

Group P   Group NP

Mean Standardized Response Rate

Scramble Degree

Set 3 ("novel")

Group P   Group NP

Mean Standardized Response Rate

Scramble Degree

Set 4 ("familiar greyscale")

Group P   Group NP

Mean Standardized Response Rate
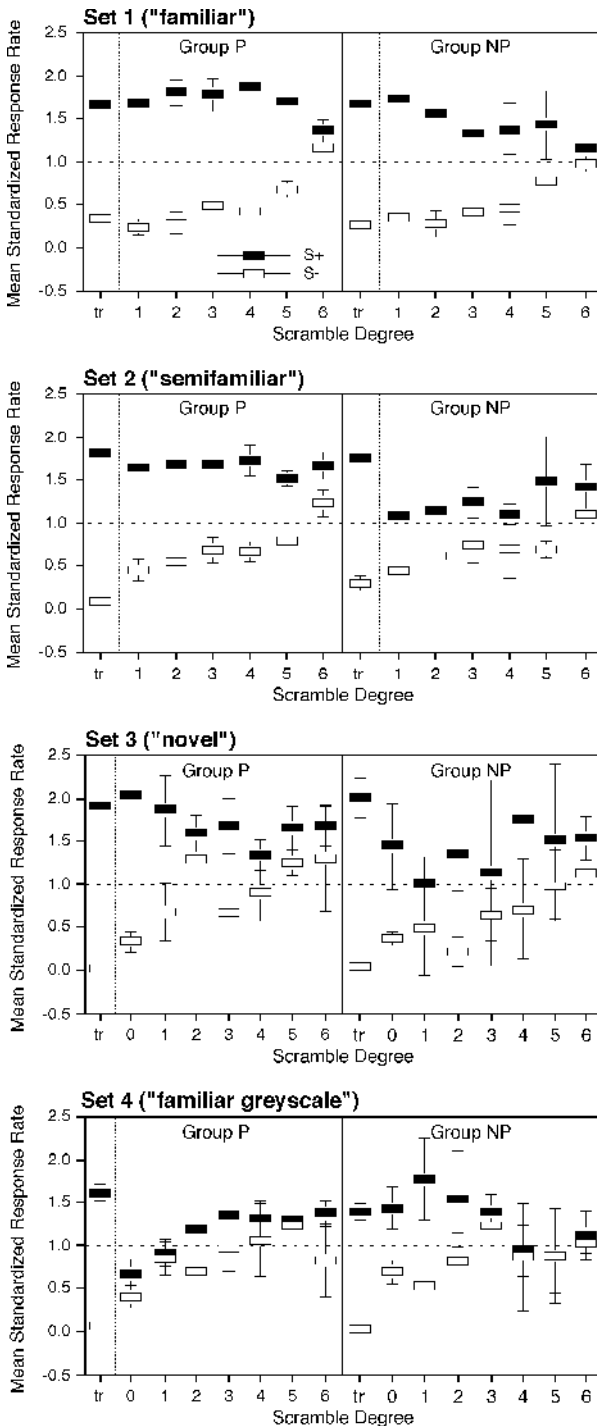
Scramble Degree

Figure 4. Performance on the positive and the negative test stimuli in the four sets of Experiment 2, shown separately for the two groups. Mean standardized response rates ($\pm SD$) are depicted as a function of scramble degree. A scramble degree of 0 relates to the unscrambled originals that were shown in Sets 3 and 4. To provide a baseline value, performance on the first presentation of the training stimuli is depicted as well (tr). Means were taken across the 2 birds of each group. S+, positive test stimuli; S−, negative test stimuli.

only on moderate, but up to very high degrees of scrambling? For instance, the pigeons of both groups succeeded even with stimuli from Experiment 1 that consisted of 256 arbitrarily mixed 8×8 pixel squares. Such tiny fragments hardly involved objectlike parts, such as faces, hands, or legs. For example, the rectangle that roughly enclosed the inner part of the face (i.e., eyes, nose, and mouth), averaged over all 52 faces present in the 40 training stimuli of Experiment 1, consisted of about 560 pixels. We therefore conclude that correct classification on high scramble degrees was mediated by recognition of tiny identification elements of the previously learned exemplars that survived the fragmentation of the stimulus. Such information units are likely to be characterized by the specific appearance of individual pixels or small pixel blocks corresponding to distinct colors or brightness values.

Insensitivity to scrambling also suggests that the absolute or relative placement of items within the image is unimportant. Thus, stimulus representation formed during training could not be stored in the form of exact *internal copies* of individual pictures—that is, templates that consist of the idiosyncratic properties of complete stimuli, together with their spatial relations.

The fact that performance on Set 3 (*novel*) was worse, less stable, and subjected to wider variations than performance on Sets 1 and 2 is further evidence that the pigeons had used item-specific features (tiny identification elements) as the main source of information when confronted with scrambled versions of novel images. If they had been able to exploit global features that correlated with the presence or absence of humans, they should have done better on Set 3.

In Experiment 1, Test 2, we had investigated the relevance of color for classifying novel stimuli and considered the color of the skin to be a possible cue for identifying a novel stimulus as a member of Class P. However, the fact that the pigeons failed on Set 3 of the present experiment, although those pictures still contained skin-colored fragments, indicates a less important role of skin color as a category cue. Considering the difference in performance between Sets 1 and 3, we assume that, in the case of scrambled familiar stimuli, pigeons used item-specific color elements—that is, absolute color aspects of stored exemplars—as cues for identifying them.

The fact that the birds of Group NP performed better on Degree 0 of Set 4 (unscrambled familiar gray-scale stimuli) than on Test 2 of Experiment 1 (novel gray-scale stimuli; see Figures 3 and 4) shows that in the case of familiar stimuli, the pigeons were obviously able to make use of item-specific properties other than color (e.g., the distribution of brightness values or the presence of familiar objectlike stimulus elements)—provided that the stimulus as a whole was intact. The failure of all birds (regardless of group membership) on scrambled gray-scale versions of familiar stimuli suggests that in the scrambled condition, the loss of item-specific color cues could not be compensated for by other features.

EXPERIMENT 3

So far, our experiments do not permit general conclusions about the importance of item-specific information, as compared with category-specific information. The good transfer performance to novel images in Experiment 1 could be accounted for either by physical similarities between familiar and novel exemplars (i.e., item-specific information) or by the acquisition of a category-specific representation of the people concept. The findings of Experiment 2 suggested that item-specific information was used to correctly classify severely scrambled images. One possible way out of this dilemma is to assume that both types of information were stored during training but only item-specific information was retrieved to master the scramble test.

The present experiment, which we called a *reversed contingencies* test, offered a possibility to test this assumption. Similar to the experiments reported by Greene (1983), the key idea was to bring information about the presence or absence of people into conflict with information about the background. First, the test provided a means for comparing the control exerted by any item-specific background feature (be it global or local) with that exerted by the experimenter-intended class rule by pitting the two against each other. For instance, we showed images of familiar people in front of familiar backgrounds from nonpeople images. Because item-specific information contained in the targets could have been stored as well, we also presented familiar backgrounds from nonpeople images, onto which novel human figures had been mounted.

Second, the test also controlled for another possible way in which the pigeons may make category-based judgements. The defining characteristics of category exemplars may be correlated with and thus confounded with category-irrelevant local features in the instances to be discriminated (e.g., man-made objects), which could enhance or even overshadow the acquisition of the discrimination of interest (Edwards & Honig, 1987). The present experiment investigated whether any such feature provided a spurious basis for discrimination.

Method

Subjects. The experimental subjects were the same as those in Experiment 1.

Stimuli. The training stimuli were the same as those in Experiment 1. The test stimuli consisted of people from Class P pictures presented against backgrounds of Class NP pictures. The resulting P stimuli were called P_1. Figure 1D shows a P_1 stimulus, as well as the original pictures (P and NP) from which it has been derived. The P_1 stimuli were created from training stimuli (*familiar*, f) and test stimuli from Test 1 of Experiment 1 (*semifamiliar*, sf), as well as from completely novel pictures (*novel*, n). This allowed for several interesting combinations between people and backgrounds of different degrees of familiarity, although only five were actually tested.[1] An overview of all investigated conditions is given in Table 1. Stimuli of the types Pf + Bf (familiar people on familiar backgrounds), Psf + Bf (semifamiliar people on familiar backgrounds), and Pf + Bsf (familiar people on semifamiliar backgrounds) provided contradictory information about class membership with respect to what the pigeons had learned about the contingencies of the constituting elements (backgrounds and people). The stimulus type Pn + Bf (novel people on familiar backgrounds) provided a contradiction between the class rule (people present) and the experienced background contingency (nonpeople). Stimuli of Type Pf + Bn (familiar people on novel backgrounds) did not contain any conflicting information and served mainly as a control. A total of 25 stimuli of each of the five test stimulus types was created.

Apparatus. The apparatus was the same as that in Experiment 1.

Procedure. The test stimuli (P_1) were presented with training contingencies and were interspersed into sequences of training stimuli at a rate of 10 per session. Since there were five testing conditions, each involving the presentation of 25 test stimuli, the whole experiment consisted of 25 sessions.

Results

Figure 5 shows the results of Experiment 3 as mean standardized response rates separately for the two groups (within-group variation is indicated by standard deviations) and for the five test conditions, as well as the first presentation of the corresponding training component to provide a reference value. We carried out two-sample t tests to compare performance on the five types of P_1 test stimuli (Pf + Bf, Psf + Bf, Pf + Bsf, Pf + Bn, and Pn + Bf) with the first presentation of the training stimuli of both classes (TrP and TrNP).

The results indicate that the birds of Group P classified the test stimuli (P_1) in a way similar to that for regular training stimuli of Class P (TrP). Performance of Bird T71-P did not deviate significantly from TrP on any of the five types of P_1 stimuli. By contrast, there were strong deviations of all types from TrNP (all $ps \leq .0001$). In case of Bird T72-P, performance on Pf + Bn, Pf + Bsf, and Pn + Bf did not differ from TrP, but there was a significant difference in the case of Psf + Bf [$t(25) = 2.81$, $p \leq .01$]. Regarding Pf + Bn, test performance even exceeded training level significantly [$t(40) = -2.28$, $p \leq .05$]. Differences between all types of P_1 and TrNP were highly significant (all $ps \leq .001$). It is noteworthy that the birds of Group P performed well not only on stimuli that did not contain any conflicting information (Pf + Bn), but also on stimuli of the types Pf + Bf, Psf + Bf, and Pf + Bsf, in which conflicting item-specific information interacted with the class rule. Correct classification was also found for stimuli of the type Pn + Bf, in which the class rule was pitted against item-specific information about former background contingency.

By contrast, the birds of Group NP responded to all types of P_1 stimuli in a rather intermediate way of classification, instead of assigning them to either the positive or the negative class. Performance of both birds deviated significantly from TrP on all types of P_1 (Bird T73-NP, all $ps \leq .0001$; Bird T77-NP, all $ps \leq .05$), as well as from TrNP (Bird T73-NP, all $ps \leq .05$; Bird T77-NP, all $ps \leq .05$).

There were no consistent differences between performance on the various types of P_1 tests, suggesting that responding was independent of the degree of familiarity with either the presented persons or the backgrounds.

Table 1
Combinations of People and Backgrounds
Used as P$_1$ Stimuli in Experiment 3

People	Background	Test Condition
f	f	Pf + Bf
sf	f	Psf + Bf
f	sf	Pf + Bsf
f	n	Pf + Bn
n	f	Pn + Bf

Note—The backgrounds of the P$_1$ stimuli were former NP stimuli. The test condition column specifies the different types of resulting test stimuli: f, familiar (i.e., stimuli from the training set used in Experiment 1); sf, semifamiliar (i.e., stimuli from the test set used in Experiment 1, Test 1); n, novel; P, people; B, background.

Discussion

The most interesting finding to emerge from the reversed contingencies test was the difference between the two groups on the P$_1$ stimuli. In Group P, performance was not at all disrupted by the modified contingencies conditions. Even stimuli that contained conflicting information with respect to previous contingencies (Pf + Bf, Psf + Bf, Pf + Bsf) or in which the class rule was in conflict with former background contingencies (Pn + Bf) were classified correctly. Group P, therefore, did not rely on any background cues confounded with the presence of a person. Instead, the results suggest that the birds of Group P made use of a category-based response rule, with classification being coupled to category-relevant features—that is, features of the target *human figure*. Their good performance with novel people on a familiar background (Pn + Bf stimuli) lends particularly powerful support to this assumption. The familiar backgrounds had previously been paired with nonreinforcement, and they covered a large area in the pictures. By contrast, the people in the pictures were unfamiliar, had thus not been previously paired with reinforcement, and occupied a smaller area in the pictures. The fact that the pigeons classified these stimuli as positive strongly supports a category-based response rule.

In Group NP, performance was dominantly controlled neither by the category rule, as in Group P, nor by the background, as reported by Greene (1983). Instead, the pigeons responded to the stimuli of Class P$_1$ in an intermediate way. A possible explanation for this result may lie in a parallel use of item- and category-specific information or, according to Greene's formulation, insufficient distinction between relevant and irrelevant features. If familiar stimuli are recognized by retrieving item-specific information and novel stimuli are classified by searching for a target (i.e., a person), combinations of familiar and novel elements or of familiar elements belonging to converse classes should lead to confusion and decrements in performance.

Taken together, the results of Experiment 3 are in sharp contrast with what Greene (1983) has reported from a similar test with pigeons. She found that when new positive slides were introduced, consisting of the target *SE* (a particular person) added to former negatives, the birds treated these slides as if they still were negatives. Complementary, all new negatives, formed by removing *SE* from former positives, were treated as if they still were positives. She concluded that "behavior was controlled much more by the backround of slides, that is, by features that were conceptually irrelevant, than by the concept *SE*" (Greene, 1983, p. 223). It seems, thus, that our pigeons came closer to the claim of Herrnstein and Loveland (1964) that pigeons, properly trained, learn to respond to the presence or absence of human beings in photographs.

GENERAL DISCUSSION

The training (Experiment 1) and the generalization test (Experiment 1, Test 1) confirmed the finding of Herrnstein and Loveland (1964) that pigeons learn to classify complex photorealistic pictures according to the presence or absence of people and that this ability can be generalized to novel instances of the two classes. The presentation of novel gray-scale stimuli (Experiment 1, Test 2) provided evidence that color was used as an important source of information for classifying novel stimuli in Group NP but that its loss could be compensated for by other information in Group P. The scramble test (Experiment 2) suggested that classification was based, at least in part, on the recognition of tiny idiosyncratic elements of previously shown training exemplars, rather than on global or configural properties of the class-defining human figures. In particular, idiosyncratic pixel properties, like the color or the brightness values of the stored training stimuli, turned out to serve as discrimination cues. Finally, the reversed contingencies test (Experiment 3) showed that both experimental groups had also learned about aspects that determined category membership but that only the performance of Group P was dominantly controlled by that type of information.

In summary, we have found evidence that *both* types of information—item-specific as well as category-specific—were stored in parallel during the people/nonpeople category learning in both groups. This is in keeping with the idea that identification and categorization need not occur to the exclusion of the other but that item-specific and category-specific information are just "two sides of the same coin" (Wasserman, Kiedinger, & Bhatt, 1988). However, as was shown in Group P in the reversed contingencies test, novel items have not been endorsed as belonging to the people category as a function of the similarity between the new item and the exemplars of the people category already stored in memory. Instead, we found evidence that the members of Group P used only category-level information to determine category membership of novel items, whereas idiosyncratic information was irrelevant for that decision. Thus, different conditions ob-
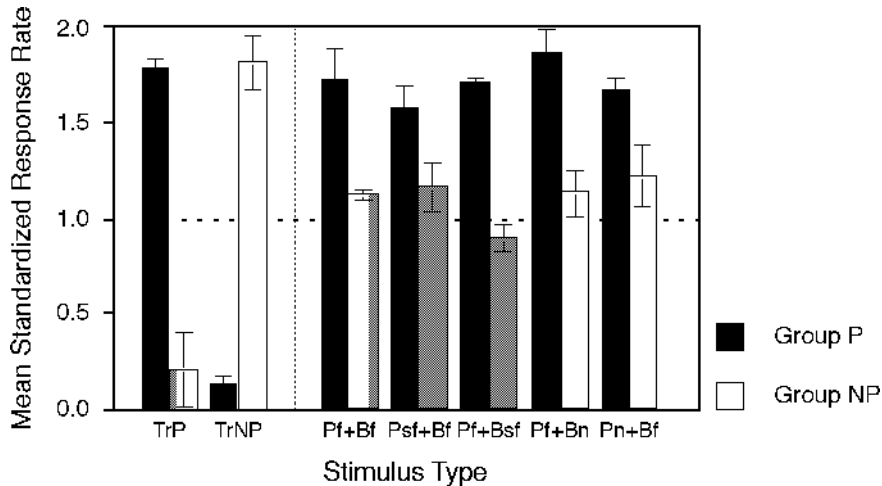
Figure 5. Performance in Experiment 3 on the five types of test stimuli and performance on the first presentation of the training stimuli of Class P (TrP). The results are shown separately for the two groups as mean standardized response rates ($\pm SD$). Means were taken across the 2 birds of each group. P, people; B, background; tr, training component; f, familiar (training set); sf, semifamiliar (test set); n, novel.

viously encourage the use of different mixtures of these two types of information, and more of one may imply less of the other (Medin et al., 1983; Wasserman et al., 1988).

Although differences between the groups must be interpreted with caution because of the small sample size, it appears that Group P actually performed somewhat better than Group NP. Compared across experiments, that difference is reflected by a consistent trend. Figure 2 seems to indicate that in Experiment 1, the training stimuli were acquired more rapidly and accurately by Group P than by Group NP. In Experiment 1 (Test 2) and in Experiment 2 (Set 2), performance of Group P was found to be superior to that of Group NP. In Experiment 3, Group P, but not Group NP, was able to classify the test stimuli of Class P_1 correctly. Most likely, that intergroup difference was rooted in the ways in which the two sources of information—item-specific and category-specific—were weighed. Whereas they seemed to be rather equal in significance for Group NP, the birds of Group P appeared to give precedence to category-specific over item-specific information. Consequently, Group P was able to deal more effectively than Group NP with many of the tasks.

A feature-positive effect may be responsible for that asymmetry, which means that discrimination was facilitated when instances containing the target were presented on reinforced trials (see, e.g., Bottjer & Hearst, 1979; Edwards & Honig, 1987; Hearst, 1978; Jenkins & Sainsbury, 1969, 1970; Manabe & Kawashima, 1982; Pace, McCoy, & Nallan, 1980; Sainsbury, 1971; Sainsbury & Jenkins, 1967). The fact that the target appeared on positive trials may have promoted the pigeons' ability to distinguish between relevant and irrelevant features in Group P. Vice versa, the lack of reinforcement regarding

the category-relevant features may have impeded correct weighing in Group NP. However, several alternatives—for instance, effects of stimulus homogeneity—are also possible and cannot be ruled out with the present data.

The research reported here addressed the relative roles of item-specific and category-specific information for classification in a people-present/people-absent discrimination task and also took an initial step toward identifying the feature domains that were exploited to accomplish that discrimination (e.g., color). But of course, the present experiments did not exhaust the conditions under which a stimulus element is recognized as a target (i.e., a person). They may, however, serve as a starting point for a more detailed investigation of the target properties that are actually used for classification.

REFERENCES

BLOUGH, D. S. (1982). Pigeon perception of letters of the alphabet. *Science*, **218**, 397-398.

BOTTJER, S. W., & HEARST, E. (1979). Food delivery as a conditional stimulus: Feature-learning and memory in pigeons. *Journal of the Experimental Analysis of Behavior*, **31**, 189-207.

CERELLA, J. (1980). The pigeon's analysis of pictures. *Pattern Recognition*, **12**, 1-6.

COOK, R. G., WRIGHT, A. A., & KENDRICK, D. F. (1990). Visual categorization by pigeons. In M. L. Commons, R. J. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), *Quantitative analyses of behavior* (Vol. 8, pp. 187-214). Hillsdale, NJ: Erlbaum.

D'AMATO, M. R., & VAN SANT, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 43-55.

EDWARDS, A., & HONIG, W. K. (1987). Memorization and "feature selection" in the acquisition of natural concepts in pigeons. *Learning & Motivation*, **18**, 235-260.

FETTERMAN, J. G. (1996). Dimensions of stimulus complexity. *Journal of Experimental Psychology: Animal Behavior Processes*, **22**, 3-18.

FLANNAGAN, M. J., FRIED, L. S., & HOLYOAK, K. J. (1986). Distribu-

tional expectations and the induction of category structure. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 241-256.

GREENE, S. (1983). Feature memorization in pigeon concept formation. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analysis of behavior* (Vol. 4, pp. 209-229). Cambridge, MA: Ballinger.

HEARST, E. (1978). Stimulus selection and feature selection in learning and behavior. In S. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior*. Hillsdale, NJ: Erlbaum.

HERRNSTEIN, R. J., & LOVELAND, D. H. (1964). Complex visual concept in the pigeon. *Science*, **146**, 549-551.

HERRNSTEIN, R. J., LOVELAND, D. H., & CABLE, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **2**, 285-311.

HONIG, W. K., & URCUIOLI, P. J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior*, **26**, 405-445.

HUBER, L. (1994). Amelioration of laboratory conditions for pigeons (*Columba livia*). *Animal Welfare*, **3**, 321-324.

HUBER, L., AUST, U., MICHELBACH, G., ÖLZANT, S., & NOWOTNY, R. (1999). Limits of symmetry conceptualization in pigeons. *Quarterly Journal of Experimental Psychology*, **52B**, 351-379.

HUBER, L., & LENZ, R. (1993). A test of the linear feature model of polymorphous concept discrimination with pigeons. *Quarterly Journal of Experimental Psychology*, **46B**, 1-18.

JACOBS, G. H. (1993). The distribution and nature of color vision among the mammals. *Biological Review*, **68**, 413-471.

JENKINS, H. M., & SAINSBURY, R. S. (1969). The development of stimulus control through differential reinforcement. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning*. Halifax, NS: Dalhousie University Press.

JENKINS, H. M., & SAINSBURY, R. S. (1970). Discrimination learning with the distinctive feature on positive or negative trials. In D. Mostovsky (Ed.), *Attention: Contemporary theory and analysis* (pp. 239-274). New York: Appleton-Century-Crofts.

JITSUMORI, M. (1996). A prototype effect and categorization of artificial polymorphous stimuli in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **22**, 405-419.

LEA, S. E. G. (1984). In what sense do pigeons learn concepts? In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 263-276). Hillsdale, NJ: Erlbaum.

LEA, S. E. G., LOHMANN, A., & RYAN, C. M. E. (1993). Discrimination of five-dimensional stimuli by pigeons: Limitations of feature analysis. *Quarterly Journal of Experimental Psychology*, **46**, 19-42.

LUBOW, R. E. (1974). Higher-order concept formation in the pigeon. *Journal of the Experimental Analysis of Behavior*, **21**, 475-483.

MALOTT, R. W., & SIDDALL, J. W. (1972). Acquisition of the people concept in pigeons. *Psychological Reports*, **31**, 3-13.

MANABE, K., & KAWASHIMA, T. (1982). The feature-positive effect based on positional discrimination in pigeons. *Annual of Animal Psychology*, **31**, 89-102.

MEDIN, D. L., DEWEY, G. I., & MURPHY, T. D. (1983). Relationships between item and category learning: Evidence that abstraction is not automatic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 607-625.

MOLLON, J. D. (1989). "Tho' she kneel'd in that place where they grew . . . ." *Journal of Experimental Biology*, **146**, 21-38.

MONEN, J., BRENNER, E., & REYNAERTS, J. (1998). What does a pigeon see in a Picasso? *Journal of the Experimental Analysis of Behavior*, **69**, 223-226.

MOSTOFSKY, D. (1965). *Stimulus generalization*. Stanford, CA: Stanford University Press.

PACE, G. M., McCOY, D. F., & NALLAN, G. B. (1980). Feature-positive and feature-negative learning in the Rhesus monkey and pigeon. *American Journal of Psychology*, **93**, 409-427.

PALMER, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

POOLE, J., & LANDER, D. G. (1971). The pigeon's concept of pigeon. *Psychonomic Science*, **25**, 157-158.

RESTLE, F. (1957). Theory of selective learning, with probable reinforcement. *Psychological Review*, **64**, 182-191.

SAINSBURY, R. S. (1971). Effect of proximity of elements on the feature-positive effect. *Journal of the Experimental Analysis of Behavior*, **16**, 315-325.

SAINSBURY, R. S., & JENKINS, H. M. (1967). Feature-positive effect in discrimination learning. *Proceedings of the Annual Convention of the American Psychological Association*, **2**, 17-18.

SIEGEL, R. K., & HONIG, W. K. (1970). Pigeon concept formation: Successive and simultaneous acquisition. *Journal of the Experimental Analysis of Behavior*, **13**, 385-390.

THEEUWES, J. (1996). Parallel search for a conjunction of color and orientation: The effect of spatial proximity. *Acta Psychologica*, **94**, 291-307.

TREISMAN, A. M., & GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

TROJE, N. F., HUBER, L., LOIDOLT, M., AUST, U., & FIEDER, M. (1999). Categorical learning in pigeons: The role of texture and shape in complex static stimuli. *Vision Research*, **39**, 353-366.

VAUGHAN, W. J., & GREENE, S. L. (1984). Pigeon visual memory capacity. *Journal of Experimental Psychology: Animal Behavior Processes*, **10**, 256-271.

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**, 235-246.

WURM, L. H., LEGGE, G. E., ISENBERG, L. M., & LUEBKER, A. (1993). Color improves object recognition in normal and low vision. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 899-911.

NOTE

1. Correspondingly, people depicted in Class P pictures were removed or obscured in order to create stimuli of Class NP. But since the respective tests provided inconclusive results—probably owing to remnant people cues in the pictures—they were excluded from the present report.