A Comparative Analysis of the Categorization of Multidimensional Stimuli: I. Unidimensional Classification Does Not Necessarily Imply Analytic Processing; Evidence From Pigeons (*Columba livia*), Squirrels (*Sciurus carolinensis*), and Humans (*Homo sapiens*)

A. J. Wills, Stephen E. G. Lea, Lisa A. Leaver, Britta Osthaus, Catriona M. E. Ryan, Mark B. Suret, Catherine M. L. Bryant, Sue J. A. Chapman, and Louise Millar
University of Exeter

Pigeons (*Columba livia*), gray squirrels (*Sciurus carolinensis*), and undergraduates (*Homo sapiens*) learned discrimination tasks involving multiple mutually redundant dimensions. First, pigeons and undergraduates learned conditional discriminations between stimuli composed of three spatially separated dimensions, after first learning to discriminate the individual elements of the stimuli. When subsequently tested with stimuli in which one of the dimensions took an anomalous value, the majority of both species categorized test stimuli by their overall similarity to training stimuli. However some individuals of both species categorized them according to a single dimension. In a second set of experiments, squirrels, pigeons, and undergraduates learned go/no-go discriminations using multiple simultaneous presentations of stimuli composed of three spatially integrated, highly salient dimensions. The tendency to categorize test stimuli including anomalous dimension values unidimensionally was higher than in the first set of experiments and did not differ significantly between species. The authors conclude that unidimensional categorization of multidimensional stimuli is not diagnostic for analytic cognitive processing, and that any differences between human's and pigeons' behavior in such tasks are not due to special features of avian visual cognition.

Keywords: pigeon, undergraduate, eastern gray squirrel, categorization, rule-governed behavior

In a recent review, Lea and Wills (2008) showed that very different inferences have been drawn from the human and animal experimental literatures about the effect of including multiple cues in discrimination or categorization tasks. Under some conditions, humans tend to cat-

egorize complex stimuli on the basis of only one of the available dimensions (a "unidimensional" sort), whereas under others, they categorize according to some kind of average of the information from all available dimensions (an "overall similarity" or "family resemblance" sort). Early data, especially from the simple categorization task referred to as the "triad task," indicated that unidimensional sorting was more likely to occur with more separable stimuli (e.g., Garner, 1974), with less time pressure (Smith & Kemler Nelson, 1984), and in adults rather than children (Smith & Kemler, 1977). On the basis of these and other results, it has been widely inferred that these two sorting strategies correspond to the use of two different cognitive mechanisms, with unidimensional classification being the product of a more explicit, analytic, verbal, or verbalizable process on the one hand, and overall similarity classification being the product of a more implicit, intuitive, nonverbalizable process on the other (e.g., Kemler Nelson, 1984; Regehr & Brooks, 1995). A distinction between verbal (rule-based) and implicit (information-integration) mechanisms of categorization forms the basis of Ashby et al.'s (1998) influential COVIS theory. Similar distinctions are elsewhere referred to as lying between rule-based and associatively based processes (e.g., Pinker, 2001; Spiegel & McLaren, 2006), or between rule-based and similarity-based processes (e.g., Pothos, 2005). In the current paper, the terms analytic and nonanalytic processes are used.

Lea and Wills (2008) argued that the identification of unidimensional classification with analytic processing, and overall similarity classification with nonanalytic processing, is mistaken. There

were two strands to their argument. The first focused on experiments with humans: if unidimensional sorting indicates analytic processing, procedural manipulations that favor analytic processing should always result in, if anything, more unidimensional sorting. In practice they do not: there are tasks and procedures where giving people more opportunity to deploy advanced cognitive strategies results in more overall similarity sorting, not less (Longmore, Milton, & Wills, 2009; Milton, Longmore, & Wills, 2008).

The second strand of Lea and Wills' (2008) argument against the identification of unidimensional classification with analytic processes and overall similarity classification with nonanalytic processes was comparative, and it is the focus of the present paper. It is beyond reasonable doubt that humans can process stimuli analytically and that our behavior is at least sometimes governed by rules that we could verbalize (e.g., Nosofsky, Clark, & Shin, 1989; Skinner, 1969), although associative processing does also occur (e.g., Shanks & Darby, 1998; Spiegel & McLaren, 2006). In animals of other species, on the other hand, the use of associative, nonanalytic processes is taken for granted, but any claim of analytic rule use in nonhuman animals is controversial (Lea & Kiley-Worthington, 1996) and calls for elaborate demonstration (e.g., Beckers, Miller, De Houwer, & Urushihara, 2006). If overall similarity classification is the product of nonanalytic processing, one would therefore predict that nonhuman animals would tend to show overall similarity classification to a greater extent than humans. Lea and Wills argued that, on the basis of current literature, there is no evidence for a consistent species difference, and that nonhumans do not reliably show family resemblance sorting. It is difficult to carry out categorization experiments with animals that follow the same procedures that are used with humans. In the absence of such experiments, the widespread success in training animals in category discrimination experiments, using polymorphous stimulus sets that can only be accurately discriminated by attention to overall similarity, has been taken as establishing that animals do indeed sort on the basis of overall similarity (see the pioneering experiment by Herrnstein & Loveland, 1964, and the review of this literature by Huber, 2001). However, there is abundant evidence from experiments using artificial polymorphous categories that birds (the subjects in most nonhuman experiments involving the discrimination of multidimensional stimuli) rarely if ever use all the stimulus dimensions available in such discriminations spontaneously, and in some cases it is impossible to train them to do so at all (e.g., von Fersen & Lea, 1990; Lea, Lohmann, & Ryan, 1993; Lea, Wills & Ryan, 2006).

The comparative argument put forward by Lea and Wills (2008) is unsatisfactory in one important respect. The human and animal research literatures that Lea and Wills (2008) reviewed have developed essentially independently, and in consequence, the comparison of their results is vexed by numerous procedural differences. In humans at least, the tendency to unidimensional classification can be greatly affected by procedural variations (e.g., Milton & Wills, 2004; Regehr & Brooks, 1995), and we have found in previous studies of comparative cognition that making procedures for human and nonhuman experiments as similar as possible can reduce apparent species differences in the results (Goto, Wills, & Lea, 2004). In the present paper, we describe two sets of experiments on multiple-feature discriminations in which the procedures were kept as similar as possible for the humans and

other species. In the first set of experiments, pigeons and humans were tested; in the second set of experiments, we added a third species, the gray squirrel (*Sciurus carolinensis*). The purpose of adding a third species was partly to check the generality of our conclusions, and partly to examine whether any pigeon-human differences we might find reflected differences between humans and nonhumans in general, or whether they might reflect differences between mammals in general and birds in general, perhaps derived from the very different visual systems and brain architectures characteristic of the two classes (see Zeigler & Bischof, 1993; Cook, 2000; Emery, 2006).

In previous experiments where we have wanted to compare human and pigeon performance under closely similar conditions (e.g., Goto et al., 2004), we found that the pragmatic constraints of providing a workable procedure were greater for the pigeon experiments than for the human experiments. In those studies, we therefore first developed procedures that were effective with pigeons, and then developed analogues of them for the human version of the experiments. We used the same general strategy in the present experiments. However in Experiment 2, where squirrels as well as pigeons were used, we started by developing a procedure that would work with the squirrels, since there is comparatively little literature available on training this species.

In summary, therefore, in the current paper we aimed to contrast two hypotheses:

(a) Unidimensional classification is the product of analytic processing of information. From the theoretical positions that unidimensional classification is an indication of analytic processing, and that humans are more likely than other species to show analytic processing, one would predict that when an experimental classification procedure is applied in a similar way to different species, humans will show a greater tendency toward unidimensional sorting than other species, regardless of the experimental situation.

(b) Unidimensional classification as the product of procedural factors. The alternative position put forward by Lea and Wills (2008) was that the tendency to unidimensional classification is mainly determined by procedural factors, at least some of which will affect humans and other species in the same way. From this perspective, we predict that there will be no substantial or consistent species difference in the tendency to unidimensional classification. On the other hand, there may well be differences in the prevalence of unidimensional classification between experimental procedures, and these will be similar for different species, including humans.

Experiment 1a

The first two experiments were designed to allow a comparison of pigeon and human performance in classification of multidimensional stimuli. The task was based on existing human sorting tasks (e.g., Regehr & Brooks, 1995), but in accordance with our general methodological approach, we first developed a procedure that could be given to pigeons, then used a close analogue of it with humans. The task chosen was the match-to-standards procedure, which was introduced by Regehr and Brooks and has been widely used since (e.g., Milton & Wills, 2004). To emulate it with pigeons, we used a conditional discrimination task, in which upon

each trial a single stimulus was presented, and there were two possible responses.

Creating stimuli whose dimensions are all relatively salient to the participants is particularly challenging in studies of nonhuman animals. With human experiments, experimenters at least have the advantage of sharing the same visual and cognitive system with their subjects, and therefore having reasonable confidence about how the stimuli will look to them; we cannot have the same certainty with subjects of other species. In Experiment 1a, therefore, we used a procedure that ensured that all the stimulus dimensions used were discriminable by our pigeon subjects, and had already acquired control over their behavior. In the first phase of training pigeons were trained to discriminate between a number of pairs of simple geometrical elements (the stimuli used are shown in Figure 1). Once the pigeons had learned this first phase discrimination to an adequate level, they were trained in a second phase, using stimuli composed of three of the elements ("Prototype" stimuli). Each prototype stimulus consisted of three elements; examples are shown in Figure 2. The discrimination could be solved equally well by using any of the dimensions, or any combination of them. After the pigeons had reached a criterion of correct performance on this second phase discrimination task, we proceeded to the cue-conflict test trials that were the main purpose of the experiment. These test trials used stimuli in which one of the dimensions took the opposite value from the others ("One-away" stimuli, as shown in Figure 2). Using such One-away stimuli is a generalization of the procedure introduced by Reynolds (1961), who used cue-conflict tests after training pigeons in a discrimination in which there were two perfectly correlated relevant dimensions. Many subsequent experiments have used a similar design (e.g., Yokoyama, Dailey, & Chase, 2006). However, with more than two dimensions, the One-away test becomes more powerful; in particular, by comparing behavior toward different One-away stimuli, it is possible to assess whether behavior is under the control of the overall similarity of the stimulus to the training stimuli, or of a single dimension (e.g., Medin, Wattenmaker, & Hampson, 1987; Regehr & Brooks, 1995; Milton & Wills, 2004). In experiments like these, the similarity between two stimuli is assessed in terms of the number of features that they have in common; this definition can be difficult to use with naturally occurring stimuli, but with simple geometrical stimuli of the sort used in the present experiments it is satisfactory.

Because the dimensions of the stimuli were spatially separated, it was possible to determine which elements of them the pigeons pecked at. It has long been supposed that peck location within a
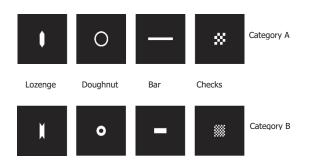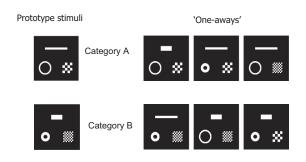

Figure 1. Elements of the stimuli used in Experiments 1a and 1b.

Figure 2. Examples of the multi-element training stimuli used in Phase 2 of Experiments 1a and 1b (Prototypes), and the corresponding One-away test stimuli used in Phase 3.

complex stimulus will be linked to the aspects of a stimulus that acquire control over behavior. Processes like this provide a plausible explanation of the feature positive effect, for example (Jenkins, 1973). In the present experiment, such an association could be tested both at the overall level and on a trial-by-trial basis, allowing an independent investigation of what features of the stimulus might be controlling behavior.

Method

Subjects

The subjects were eight pigeons (*Columba livia*). Five were obtained from a local fancier, and three were bred in the laboratory. They were housed in an indoor aviary, and were transferred to individual cages on days when they were to be tested; after testing they were weighed and given any supplementary feeding needed to maintain their weight at around 90% of free feeding levels. On nontesting days the pigeons remained in the aviary and were given a limited food supply there. The pigeons had all taken part in a previous discrimination learning experiment, but the procedure, stimuli, and equipment set-up were all different from those of the present experiment.

Apparatus

The experiment used four identical operant conditioning chambers, measuring $710 \times 505 \times 435$ mm. One long wall of each box included a 38-cm (15-in.) touch monitor, which consisted of a resistive touchscreen mounted in front of a TFT computer display screen (ELO Accutouch, Model 1547L; ELO Touchsystems Inc.) The bottom edge of the screen was 120 mm above the grid floor of the chamber. Two 2.8-W white houselights were mounted in the top corners of the operant panel above and to either side of the screen. Two apertures, each measuring 60×50 mm and giving access to grain hoppers when the hopper solenoids were activated, were located directly below the houselights and 40 mm above the grid floor of the chamber. The hoppers were illuminated by a 2.8-W white light when activated, and contained a 2:1 mixture of hemp seed and conditioner. White noise was played into the box from a 50-ohm loudspeaker located centrally below the touchscreen. Pecks to target areas of sufficient force to trigger the touchscreen were followed by an immediate 6-dB bleep, the interior of the box could be observed by a video camera mounted on the side of the chamber 250 mm above the floor. The chambers

were housed in a darkened room together with other similar apparatus. Stimulus presentation and reinforcement contingencies for all chambers were controlled, and responses and other data recorded, by a computer (supplied by Quadvision Ltd., Dorset, U.K.) located in an adjacent laboratory area, with software written in Visual Basic using the Whisker control system (Cardinal & Aitken, 2001). Each box had a webcam fitted into the side wall, 250 mm above the floor. The pigeons' behavior could thus be observed from outside the test room using the imaging software ViewCommander (Internet Video and Imaging, Ltd.). Each pigeon was assigned its own test chamber for all stages of the experiment.

Stimuli

In an effort to ensure that all the stimulus dimensions were of roughly equal salience, highly salient dimensions as hue and brightness were avoided. All the stimuli were filled white geometrical forms shown on a black background, and all consisted of the same number of white pixels, within a tolerance of 2%. The stimuli used are illustrated in Figure 1. There were four two-valued stimulus dimensions, designated as Bar, Doughnut, Checks, and Lozenge, so there were eight possible stimulus elements. One value of each dimension was designated as belonging to stimulus set A and the other as belonging to stimulus set B, as shown in the figure. At different stages of the experiment these elements were presented either individually or in various combinations.

Procedure

Standard procedures were used to train the pigeons to take food from either food hopper when it was operated, and then to peck a 30-mm diameter white disk located centrally on the touchscreen. They were then exposed to the following element-training schedule (similar to the "Features in parallel" procedure of Lea et al., 2006). At the beginning of each trial, a single stimulus element appeared on the touchscreen, centered 205 mm above the grid floor in the vertical dimension, and centrally in the horizontal. Pecks anywhere in a 400-pixel square region centered around this element were recorded, and reinforced on a *tand* FT 10s VI 5s schedule by the presentation of two choice response areas on the computer screen (i.e., there was a period of 10 s during which pecks had no scheduled consequences, and then the first peck after a further period of from 0 to 10 s caused the choice keys to appear). The choice keys were white disks 30 mm in diameter, centered 185 mm above the grid floor of the chamber and 50 mm from each edge of the touchscreen, so that they were near the hopper apertures. One of the choice keys was designated as correct, depending on what stimulus element was displayed: for four of the pigeons (Ba, Io, Le, and Ly) the right key was correct if a set A element was presented, and the left key was correct if a set B element was presented; for the remaining pigeons (As, At, PA, and Sf) these contingencies were reversed. The first peck on the correct choice key was reinforced with 3-s access to a mixture of hemp-seed and pigeon conditioner from the hopper nearer to it; pecks to the incorrect key were recorded but had no scheduled consequences. The trial was recorded as having a correct response if the first choice key peck was made to the correct key. Following reinforcement, there was an intertrial interval of between 10 and 15 s. Sessions consisted of 64 trials, made up of eight cycles through the

eight possible stimulus elements. The order of presentation of elements was randomized afresh for each cycle, subject to the constraint that no more than three consecutive trials involved stimuli from the same set. One session was given per day, normally 5 days per week.

Training continued for each pigeon until it reached a criterion of 75% correct on both elements within each of the four dimensions, or for 40 sessions, whichever happened first. If a pigeon did not reach 75% on all four dimensions within 40 sessions, training continued until it had reached a less severe criterion of 75% correct on both elements of at least three dimensions. Once criterion had been reached, the pigeon was transferred to prototype training. The three dimensions on which the pigeon had reached the criterion were selected (or, if it had reach criterion on all four, the three dimensions on which its performance was best), and prototype stimuli consisting of the three set-A or the three set-B elements from these dimensions were constructed. The elements were arranged in an upright triangle, centered 40 mm apart. Examples are shown in Figure 2. Training on prototype stimuli continued, with the same contingencies of reinforcement as in element training, until each pigeon reached a criterion of 87.5% correct.

Once they had reached criterion on prototype training, the pigeons were transferred to one-away testing. One-away stimuli were constructed by substituting one of the elements in the prototype stimuli by the corresponding element from the other stimulus set, with the same triangular arrangement as before. Examples are shown in Figure 2. Across test sessions, all six possible one-away stimuli were used, in all possible arrangements within the triangle. Test sessions consisted of 60 trials; the first 20 used only prototype stimuli, with training contingencies in force. The remaining 40 included 24 with prototype stimuli and training contingencies in force, four with prototype stimuli and test contingencies in force, and 12 with test stimuli and test contingencies. Within each test session, we presented all six possible test stimuli using a single geometrical arrangement of the stimulus dimensions (e.g., Bar at the top, Doughnut to lower left, Checks to lower right) twice each, but the locations of the dimensions were varied between test sessions. On test trials, the first peck to either choice key was reinforced, regardless of the stimulus presented. Following each test session, further prototype training sessions were given with a different pair of prototypes until the criterion of 87.5% correct performance was again achieved. This procedure was repeated until all pairs of prototypes had been used and therefore six test sessions had been given, for a total of 72 trials with test stimuli for each pigeon.

Results and Discussion

As in the corresponding condition of Lea et al. (2006), learning to discriminate the individual stimulus elements was slow: only one pigeon reached criterion on all four dimensions within 40 sessions, and all remaining pigeons required at least 10 further sessions to reach it on three dimensions. For all pigeons but one, the dimension learned most slowly was the Lozenge, and apart from the pigeon that reached criterion on all four dimensions, only one of the subjects reached criterion on that dimension.

Transfer from element training to prototype training was not instantaneous, but was quick. The pigeons required from two to five sessions to reach the 75% criterion.

The response made in each test trial involving a one-away stimulus was classified as consistent or inconsistent with each of four possible simple hypotheses about control over behavior in the case of cue conflict: unidimensional control by each of the three dimensions, and control by the preponderance of feature values (overall similarity). Table 1 shows the percentages of the 72 test trials for each pigeon whose results were consistent with each of the four simple hypotheses.

In interpreting the results in Table 1, it is important to bear in mind that the figures for an individual pigeon in the four columns are neither independent nor complementary. For example, in a situation where a subject is perfectly following an overall similarity classification, so that the overall similarity proportion is 100%, the proportions for the other hypotheses will be determined by the design of the experiment. However they would not be 0%, they would be 67%. Similarly, a subject perfectly following a unidimensional strategy produces an overall similarity proportion not of 0%, but of 67%.

Because of this interdependence, it is not possible to test whether the proportion of trials predicted by one hypothesis is significantly different from the proportions predicted by the others. For several birds, the difference of the best from the next best hypothesis is small. However, in each case one of the four hypotheses accounts for more of the test trial outcomes than any of the others, so for each pigeon, one of the four hypotheses can be described as the numerically best-fitting hypothesis. Clearly, given the small differences in numbers of trials correctly predicted between the four hypotheses, any particular classification might be in error—that is to say, the data are inherently noisy (and much more noisy than would be expected on the basis of results with

humans in the studies from which our procedure was derived, e.g., Milton et al., 2008). However there is no reason to expect the errors to be other than random, so that the data can be used to make comparisons between species and procedures of the proportions of subjects that have a unidimensional or an overall similarity best-fitting hypothesis. It is these comparisons that are the central concern of the current paper. The noise in the data reduces the power of such comparisons, and it is therefore important to demonstrate that differences can be found.

On the basis outlined above, four of the pigeons are categorized as overall similarity classifiers, and four of them as unidimensional classifiers. One pigeon (Ba) showed a very strong position bias during test trials, which attenuated any cue effects, but even this pigeon could be assigned to one of the groups (unidimensional Bar classification). The amount of overall similarity responding is higher than might be expected from previous results on birds' discrimination of artificial polymorphous stimuli (e.g., Lea et al., 1993; Lea et al., 2006), but it must be recalled that in the present experiment, the pigeons were explicitly trained to discriminate all the dimensions in the first phase of the experiment.

One possible explanation for the relatively small differences observed between the proportion of trials predicted by the four hypotheses is that control over behavior might be restricted to one location within the stimulus. In the extreme case where a subject only came under the control of the element at one of the three locations in the stimulus, and was only under the control of one of the stimulus dimensions (responding at random when other dimensions were presented in its preferred location), it is easily shown that only 67% of test trial responses would be consistent with the unidimensional hypothesis for that element, while 44% of re-

Table 1

Experiments 1a and 1b Response on Test Trials: Proportions of Trials Out of 72 (Pigeons) or 16 (Humans) in Which Response Was Consistent With Each of Four Simple Hypotheses About the Subject/Participant's Categorization Strategy, and the Best-Fitting Hypothesis

	Consistent with control by					
	Bar	Checks	Lozenge	Doughnut	Overall similarity	Apparent strategy
Pigeons						
As	39%	61%		58%	58%	Checks
At	39%	72%		50%	61%	Checks
Ba	54%	51%		46%	51%	Bar (position bias)
Io	67%	67%		58%	72%	Overall similarity
Le		53%	53%	72%	64%	Doughnut
Ly	49%	49%		51%	64%	Overall similarity
Pn	49%	51%		49%	63%	Overall similarity
Sf	63%	54%		54%	65%	Overall similarity
Students						
1	67%	67%		67%	100%	Overall similarity
2	50%	83%		50%	83%	Overall similarity/checks
3	67%	67%		67%	100%	Overall similarity
4	50%	67%		67%	83%	Overall similarity
5	75%	58%	58%		92%	Overall similarity
6		67%	33%	67%	67%	Overall similarity/checks/Doughnut
7	58%	75%		58%	92%	Overall similarity
8	33%		33%	100%	67%	Doughnut
9		42%	42%	92%	75%	Doughnut
10	33%	33%		100%	67%	Doughnut
11	67%	67%		67%	100%	Overall similarity
12	67%	67%		67%	100%	Overall similarity

sponses would be consistent with the other unidimensional hypotheses, and 55% of responses would be consistent with an overall similarity strategy. In contrast, the figures for a subject under control of one element, wherever it appeared in the stimulus, are 100%, 33%, and 67% respectively. It would be expected that a pigeon that was under preferential control of stimuli appearing in one area of the response key would peck more at that area. Table 2a shows that several of the pigeons did show marked disparities in the mean numbers of pecks directed to the three different areas within the stimulus; the differences were significantly consistent across the one-away test trials (Kendall's W test) for five of the eight birds.

In further analyses, we investigated whether peck rates to stimulus dimensions (as opposed to location within a stimulus) were a valid indicator of selective stimulus control. Tables 2b and 3 show the results of this analysis. Table 2b shows mean peck rates to each of the dimensions in the first 10 s of all test trials (during which the choice keys were never available). Several birds had markedly different mean response rates toward different dimensions, and for six of the eight birds these differences were significant (Kendall's concordance test). However the differences in mean response rate were not consistently related to the dimensions that controlled

Table 2

Experiment 1a. Behavior on Test Trials: Mean Pecks at Each Element Location (Table 2a) and Stimulus Dimension (Table 2b) in the First 10 s of a Trial, and Concordance (Kendall's W) of Rankings of Peck Rates Across One-Away Test Trials. Figures in Parentheses Are Upper Bounds of 95% Confidence Limits of Mean Peck Rates. In All Cases the Lower Bound Was Zero

(a): Response rates as a function of the location being pecked

| Pigeon | Location | | | W |
	Top	Bottom left	Bottom right	
As	1.47 (6.16)	1.58 (6.61)	1.31 (5.55)	.003
At	.17 (0.81)	1.75 (5.97)	1.92 (5.91)	.37***
Ba	1.30 (5.16)	1.58 (4.58)	3.00 (7.91)	.16***
Io	2.21 (9.19)	1.40 (7.57)	2.08 (8.14)	.03
Le	.33 (6.00)	4.97 (14.74)	4.97 (15.48)	.23***
Ly	.78 (2.92)	3.96 (9.51)	1.00 (4.91)	.42***
Pn	1.03 (3.80)	2.06 (8.57)	2.67 (9.53)	.01
Sf	.00 (0.00)	1.10 (4.89)	4.65 (9.97)	.74***

(b): Response rates as a function of the stimulus dimension being pecked

| Pigeon | Stimulus dimension | | | | W |
	Bar	Checks	Lozenge	Doughnut	
As	0.07 (0.57)	3.71 (8.55)		0.58 (3.98)	.68***
At	1.33 (4.37)	1.54 (4.88)		2.11 (6.33)	.21***
Ba	1.79 (5.87)	2.69 (6.88)		1.79 (5.58)	.10**
Io	3.65 (11.95)	0.31 (2.53)		1.74 (7.13)	.17***
Le		1.00 (6.23)	4.94 (15.84)	3.85 (14.62)	.16***
Ly	1.78 (6.76)	1.83 (7.01)		2.13 (6.95)	.02
Pn	2.15 (8.39)	0.88 (3.91)		2.72 (9.61)	.09***
Sf	1.38 (5.40)	2.19 (9.17)		1.64 (6.15)	.03

** $p < .01.$ *** $p < .001.$

choice behavior, probably because of high variability in response rates, reflected in the wide confidence intervals on mean rates and the low values of the W coefficients shown in Table 2b. It is notable that pigeon As, which has a much higher W value than any other bird, did have a perfect correlation between mean peck rates to a dimension and the proportions of test trials that were consistent with its value (Checks > Doughnut > Bar in both cases).

At a more detailed level, the relation between peck rates and eventual choice was consistent across birds. Table 3 shows that with the sole exception of pigeon As on the Checks dimension, each bird's mean peck rates to each dimension were higher on trials when the bird eventually made a choice that was consistent with the information given by that dimension (e.g., if left choice pecks had been reinforced in the presence of the long Bar, peck rates to the Bar dimension were higher on trials when the long Bar was shown and the bird did make a left choice, or on trials when the short Bar was shown and the bird made a right choice). Because peck rates were quite variable, and the number of test trials relatively low, confidence intervals for these peck rates are wide. However, the effect is large (the geometric mean of the ratio between the rates under consistent and inconsistent trials was 2.66), and since all eight birds showed such an effect for at least two out of three dimensions, this difference is significant (2-tailed binomial test, $N = 8$, $p < .01$). It appears that although overall peck rates varied greatly between trials, on individual trials the pigeons pecked more at a particular dimension when they were going to go on to make a choice that was consistent with the value of that dimension.

Because the first response on each choice trial was reinforced regardless of its location, it is possible that the pigeons could have learned to recognize the test stimuli and respond to them in accordance with reinforcement received on early presentations. However it is not obvious how any such strategy could have produced any systematic behavior of the kinds that were in fact observed.

Experiment 1b

Experiment 1b was as closely as possible a human analogue of Experiment 1a. Because human subjects learn this kind of discrimination much more quickly than pigeons, the procedure was compressed into a single session. However the same stimuli, contingencies and learning criteria were used, and results were analyzed in the same way.

Method

Participants

Fourteen undergraduates (*Homo sapiens*) were recruited through the School of Psychology's participant panel. Participation was incentivized by a small money payment or course credit.

Apparatus, Stimuli, and Procedure

Experimental sessions took place in a small, quiet room. Participants sat at a table on which there was a keyboard and computer screen, with the same kind of touchscreen as used in Experiment 1a. The stimuli were identical to those in Experiment 1a, and the same program was used to control the experimental procedure.

Table 3
Experiment 1a. Behavior on Test Trials: Mean Pecks at Each Stimulus Dimension in the First 10 s of a Trial, Broken Down by Whether the Choice Response Was Inconsistent or Consistent With the Information Given by That Dimension. Upper Figure: Mean. Lower Figure: Upper Bound of 95% Confidence Interval of Mean (All Lower Bounds Were Zero)

Final response consistent with dimension value	Stimulus dimension							
	Bar		Checks		Lozenge		Doughnut	
	No	Yes	No	Yes	No	Yes	No	Yes
Pigeon								
As	0.05	0.11	4.07	3.48		.	0.03	0.98
	0.46	0.72	9.66	7.78			0.39	5.27
At	1.20	1.54	1.25	1.65		.	2.06	2.17
	4.44	4.23	5.11	4.78			6.06	6.65
Ba	1.64	1.92	1.80	3.54		.	1.49	2.15
	6.10	5.70	4.80	8.03			4.60	6.56
Io	0.04	5.46	0.18	0.44		.	0.27	2.79
	0.44	13.55	1.61	3.30			2.04	8.93
Le			1.03	1.63	3.45	6.62	2.75	4.27
			4.80	7.72	12.61	18.46	10.77	15.89
Ly	1.54	2.03	1.15	2.24		.	1.69	2.54
	6.22	7.33	6.00	7.50			5.42	8.13
Pn	1.08	3.29	0.49	1.24		.	0.78	3.89
	4.87	10.78	2.63	4.80			3.46	11.51
Sf	0.59	1.84	1.76	2.66			0.64	2.49
	3.10	6.32	7.91	10.39			3.22	7.61

Participants responded by touching the screen. The timing of events differed slightly from Experiment 1a, in that the intertrial interval was reduced to 1 s and was constant, and correct responses were followed by the presentation of a large smiley face in the center of the screen for 1 s. Other timings, including the 10-s period during which responses to the stimuli were recorded but had no scheduled consequences, were the same as in Experiment 1a. However the session structure was different from Experiment 1a. There was a single session, divided into three Phases.

Phase 1 involved single element training, organized into 64-trial blocks: the two versions of the four elements shown in Figure 1 were presented individually, with each stimulus occurring on eight trials within each block. Up to three blocks were given. If the number of errors to each of the eight stimuli was below three within a block (i.e., if performance on each stimulus was 75% correct or better), Phase 2 began. If the participant did not reach this criterion by the third block of Phase 1 training, the session terminated. Half the participants were trained under conditions where the left response key was correct in the presence of category A stimuli, and half in the opposite conditions.

Phase 2 involved training with three-element stimuli as shown in Figure 2. As for the pigeons, the three dimensions for which accuracy was greatest in Phase 1 were included in the Phase 2 stimuli. Only the two prototype stimuli were presented. Trials were organized into 64-trial blocks with each stimulus being presented on 32 trials within each block. If the number of errors to each of the stimuli was below nine within a block (i.e., if performance on each stimulus was 75% correct or better), Phase 3 began.

Phase 3 involved testing with one-away three-element stimuli as shown in Figure 2. To maintain comparability with the pigeon procedure, test trials were interspersed among training trials of the kind used in Phase 2. There was one 40-trial block. The test block included 12 trials with each of the two prototype stimuli, with the same contingencies as used in training. The remaining 16 trials were test trials. The 16 test trials within the block included two with each of the prototype stimuli as used in training trials, and two with each of the six one-away stimuli derived from that participant's prototype stimuli. Whatever response the participant made on a test trial was treated as correct.

Results and Discussion

Learning was relatively quick. Two participants failed to reach criterion in the first two Phases and therefore did not progress to testing. As with the pigeons, the Lozenge was the stimulus dimension that tended to be learned most slowly (8 out of 12 students who completed training). Table 1 includes the results of the test trials with the 12 participants who reached criterion. As the table shows, seven participants could be unambiguously categorized as responding in accordance with overall similarity, three could be unambiguously categorized as responding according to a single dimension (the Doughnut in all cases) and the other two gave unclear patterns. The individual participants could be assigned to sorting strategies with more confidence than the pigeon subjects in Experiment 1a, though there was more ambiguity than in typical free sort experiments (e.g., Milton et al., 2008), presumably as a result of the procedural changes we had to introduce to ensure that the human and pigeon procedures were as similar as possible. Analysis of relative response rates to different areas of the stimulus, as carried out for the data of Experiment 1a, was not performed for the data from the present experiment since the participants tended to show idiosyncratic location-based response patterns with little variation between trials.

In test trials, the proportions of the two species showing overall similarity sorting did not differ significantly (Fisher's exact test, $p = .36$); the humans were if anything more likely than the pigeons

to respond in accordance with overall similarity rather than a single dimension. The level of overall similarity sorting observed in humans is comparable to that found in other human match-to-standards procedures with spatially separate stimuli (e.g., Milton & Wills, 2004).

The results of these first two experiments provide no evidence that humans, when tested with multidimensional stimuli under comparable conditions, show a higher prevalence of unidimensional classification than pigeons. Under the assumption that analytic processes are more available to adult humans than to pigeons, these results therefore provide no evidence in support of the hypothesis that unidimensional classification is the product of an analytic process. If anything, these two experiments give more support to the hypothesis that overall similarity sorting can be the product of an analytic process as the ratio of overall similarity classifications to unidimensional classifications is numerically higher in humans than in pigeons.

On the other hand, the present results do not support the generalization that pigeons cannot classify stimuli in terms of overall similarity. The results of Experiment 1a differ from those of past experiments that have found that animals show unidimensional control over behavior when it would be advantageous to use information from multiple dimensions (e.g., Lea et al., 1993, 2006). The present results suggest that such behavior may, at least partially, result from a failure to detect the contingencies relating to some dimensions. With the present procedure, where the contingencies on each dimension were established by preliminary training with stimulus elements, a reasonable amount of overall similarity responding was seen. Nonetheless there was some tendency for unidimensional responding, and in the pigeons at least this may well be attributable to limited ability to process multiple dimensions, as suggested by Lea et al. (2006).

Experiment 2a

Experiments 1a and 1b showed that humans were, if anything, more likely than pigeons to categorize ambiguous stimuli on the basis of overall similarity (as opposed to categorizing them on the basis of a single stimulus dimension). This result does not seem entirely consistent with a hypothesis that overall similarity sorting indicates a nonanalytic, "associative" process, while single-dimension sorting indicates an analytic, "rule-based" process.

The second series of experiments had three aims. First, we wanted to establish whether our method of classifying subjects as overall similarity or unidimensional sorters had sufficient power to be useful. As noted in the discussion of Experiment 1a, the classification results involved considerable noise. The lack of observed species differences could therefore possibly be due to a lack of statistical power. We therefore ensured that the procedure in the second series of experiments was substantially different from that of Experiments 1a and 1b. If we observed a change in the frequency of unidimensional sorting compared with Experiments 1a and 1b, we could be sure that if there had been large species differences, the method had enough sensitivity that we would have detected them in Experiments 1a and 1b, although obviously a small species difference could still have escaped detection. Our second aim was related: we wanted to see whether the similarity of behavior across species found in Experiments 1a and 1b would still be found under quite different conditions. Accordingly, in this

second set of experiments, we sought conditions where unidimensional sorting might be expected to dominate in humans, as opposed to the dominance of overall similarity sorting in Experiment 1b. Milton and Wills (2004) found that unidimensional classification is more likely to be dominant in humans in situations where the stimulus dimensions are spatially integrated, and hence Experiment 2 employed spatially integrated stimuli. Additionally, we wanted to examine conditions that were somewhat more typical of those in which animals have been trained in the past (e.g., Reynolds, 1961), so in these experiments we used spatially integrated stimuli, with more salient stimulus dimensions, and without initial training on each stimulus dimension separately.

Our third aim was to investigate whether any differences that might be found between pigeons and humans under these different conditions reflected a general animal/human difference, or whether there might be, for example, a bird/mammal difference, with nonhuman mammals behaving more like humans than pigeons. Accordingly, in Experiment 2a, the subjects were eastern gray squirrels (*Sciurus carolinensis*).

Gray squirrels were chosen because they are diurnal, nonprimate mammals and they are capable of learning hue discriminations in food rewarded tasks both in the laboratory and the field (Jacobs, 1976; Macdonald, 1992). On both psychophysical and molecular grounds they appear to be, like most mammals, cone dichromats (Carvalho, Cowing, Wilkie, Bowmaker, & Hunt, 2006; Jacobs, 1976), but it is not known how important color is in their visual cognition. It was therefore of interest to see whether color would emerge as the dominant dimension if they did show unidimensional sorting, as would be expected for birds in similar experiments (cf. Lea et al., 2006).

There is little preexisting literature on laboratory learning in squirrels and, as a nondomestic species, they are easily stressed by laboratory procedures. Extensive pilot testing was therefore needed to find an appropriate procedure. In particular it was difficult to find a training procedure that would be effective for the squirrels while controlling for the influence of olfactory cues, to which we have found them to be highly sensitive. The procedure finally adopted was based on the "pokeboard" apparatus described by Ferster & Skinner (1957, p. 32) for rapid training of pigeon key-pecking. We have developed a version of this task for use in discrimination learning with nonhuman primates under zoo conditions (Warner & Lea, 2007), and this was adapted for use with the squirrels. The pokeboard was manually operated; Experiments 2b and 2c, with pigeons and humans respectively, were designed to simulate the squirrel procedure as closely as possible using automated apparatus. The final version closely resembles the multiple-stimulus procedure described by Huber, Apfalter, Steurer, and Prossinger (2005), who showed that pigeons could learn a category discrimination very rapidly in a procedure where multiple instances of two categories (achromatic human faces of male or female gender) appeared on a touch screen, and the pigeons' task was to peck at instances of one of the categories. Pigeons that had failed to show any improvement at the gender discrimination task in over 60 hours of training under a conventional go/no-go discrimination procedure learned it within a few hours under the new procedure.

As in Experiments 1a and 1b, the subjects' task in the second series of experiments was to discriminate between two prototypical stimuli that differed on three dimensions; and successful learn-

ing was followed by testing with one-away stimuli. However, in comparison with the first series of experiments, the stimuli were more conventional. They involved more salient dimensions (hue, gross shape, and surface pattern), the dimensions were not spatially separate, and there was no initial training with individual dimensions. These changes were made partly in the expectation that they would induce a dominance of unidimensional sorting in humans, partly to make the procedures more similar to those used in standard experiments such as that of Reynolds (1961), and partly to simplify procedures as far as possible for the squirrel experiments (and to use the most salient cues possible) to maximize the chance of training them successfully.
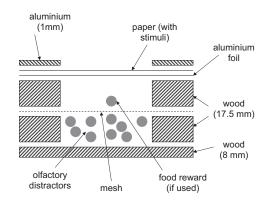
Method

Subjects

The subjects were two male and two female hand-reared eastern gray squirrels (*Sciurus carolinensis*), housed in mixed-sex pairs. They were part of a colony maintained in large cages; one pair was in a cage 2 m × 1.8 m × 2.4 m and the other in a cage 3m × 1.8m × 2.4m. The colony occupied two air-conditioned, temperature controlled rooms under a 12:12 lighting regime. The cages were connected by tunnels and wire mesh bridges, 2.1 m above the floor, to a similar room containing two large cages of similar size that were used for testing. Normally access to the bridge and tunnel system from the living cages was closed off by sliding doors. As part of their regular maintenance, the squirrels were trained to come one at a time to the test cages to feed when the sliding doors were opened, and to return to the living cages afterward. The entire procedure of the experiment could be carried through without any need to handle the squirrels.

Apparatus

The apparatus consisted of a wooden pokeboard, 210 × 210 × 45 mm in size, with an array of 12 45-mm diameter food wells arranged in a hollow square around the edge. It was made up of seven layers, as shown in Figure 3. The lower parts of the food wells were always filled with the food used as rewards in the experiment, pieces of "Honey Nut Cheerios" breakfast cereal
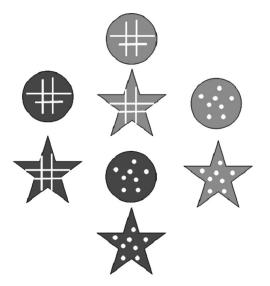

Figure 3. Cross-section through the pokeboard apparatus used in Experiment 2a, taken at a foodwell. Not to scale. Foodwells were 45 mm in diameter.

(refreshed daily), to ensure that all wells gave the same olfactory cues. The next layer consisted of a fine metal mesh to prevent the squirrels reaching this food source. The food rewards, where used, were placed in the upper layers of the foodwells. The foodwells were closed with a sheet of aluminum foil, to mask olfactory cues, and a sheet of paper on which the stimuli were printed; these were held in place by the top layer of sheet aluminum. The components of the pokeboard were held together with four bolts and wing-nuts, positioned in a square inside that formed by the food-wells.

Procedure

The experiment began with pretraining, with one trial being given per day. When a squirrel entered the experimental cage at the beginning of the trial, the exit door was closed, and the squirrel found the pokeboard in position on the floor. It was allowed to explore the pokeboard, and open any number of foodwells. Once the squirrel had opened all the foodwells, the exit door was opened and the squirrel was allowed to leave the experimental room. This was also done immediately if the squirrel appeared distressed, if 60 minutes had elapsed from the squirrel's entry to the test room without any interaction with the pokeboard, or if 60 minutes had elapsed since the last contact with the pokeboard. On all trials, the order in which the squirrels opened and explored the 12 foodwells was recorded. Initially, the squirrels were trained to find food in the foodwells, with white paper covering them, but with the paper torn to expose the food beneath. No aluminum foil was present. Once they were finding food quickly, in successive phases the paper was cut but not opened, was not cut at all, and the aluminum foil was added. Next they were exposed to a condition where intact paper and foil covered all foodwells, and only half the foodwells had rewards in them; this procedure allowed a check that the squirrels were not detecting the food rewards by olfaction. Once the squirrels were opening all foodwells rapidly in this condition, and showing no sign of olfactory discrimination, a preliminary discrimination was introduced in which the foodwells containing food were marked with a regular, black, filled hexagon (side length 15 mm). On the first trial, all foodwells were marked with hexagons and contained food; on subsequent trials, a randomly chosen half of them were. Four trials were given per day. Discrimination performance was measured by the ρ statistic of Herrnstein, Loveland, and Cable (1976), calculated from the order in which foodwells were visited. Foodwells were scored as visited if the paper covering them was perforated, regardless of whether the squirrel extracted any food from it. Thus if all positively marked foodwells were perforated before all negatively marked ones, ρ would take the value 1.0; if they were opened in random order, ρ would be approximately 0.5. For each squirrel, training with the hexagon stimuli was given for a minimum of four trials, and then continued until ρ exceeded 0.8 on three successive trials, which could include the first four.

Once the squirrels had reached criterion with the hexagon stimuli, the experimental discrimination was introduced. The stimuli are shown in Figure 4. They had three dimensions, all of which gave perfectly valid information. Six of the foodwells were marked with each of the stimuli. For two of the squirrels, food reward was available in the foodwells marked by the orange circles with a hash pattern, and not in the foodwells marked by the blue stars with a dotted pattern; for the remaining two squirrels, these contingencies

Figure 4. The training and testing stimuli used in Experiments 2a, 2b, and 2c. Light gray represents orange; darker gray represents blue. The two colors were of roughly equal brightness to the human eye. The top and bottom stimuli were used in training; the remaining six stimuli are the "one-away" test stimuli.

were reversed. Training was continued until the same ρ-based criterion was reached.

When criterion had been reached, one-away trials were introduced. In these trials, two of the training stimuli (one positive and one negative) were replaced by 'one-away' stimuli as shown in Figure 4, where one dimension of the training stimulus was replaced by its equivalent from the other training stimulus. There were six possible one-away stimuli. On test days, six trials were given, and within each such session, each one-away stimulus was used twice. On one of the trials with each stimulus the foodwell it covered was baited, and on the other it was not. Three of the test stimuli were associated with baited foodwells on first presentation: the other three on second presentation. Each squirrel was given three test sessions. Each test stimulus was scored as being treated as a positive if it was among the first six stimuli perforated on a trial, and scored as treated as negative otherwise. After each test day, at least 2 days of training were given; if necessary, this was continued until the ρ value again exceeded to 0.8.

Results and Discussion

The squirrels required three to five sessions of four arrays to reach criterion on the preliminary discrimination of hexagons from blanks. The number of sessions to criterion on the experimental discrimination ranged from 2 to 10.

Table 4 includes the data from the test trials with one-away stimuli. It reports the proportions of the 36 test stimuli for which the squirrels' behavior toward them (perforating them within the first six stimuli of a trial, or not) was consistent with each of the four simple hypotheses about the control of behavior—the three unidimensional hypotheses (color, shape, or pattern) and the overall similarity hypothesis. It can be seen that for all four squirrels a unidimensional hypothesis is the numerically best-fitting; for one

of the squirrels the stimulus shape was the dominant dimension and the other three it was stimulus color.

Previous experiments (e.g., Jacobs, 1976; Macdonald, 1992) have shown that gray squirrels can make discriminations on the basis of hue. It should be noted that although the two stimulus colors used were approximately balanced for brightness to the human eye, no attempt was made to correct for the different characteristics of squirrel color vision, so we are not able to say whether the squirrels in the present experiment were influenced by hue rather than saturation or brightness. However this experiment is the first demonstration that any dimension of color can dominate shape and pattern cues in this species. The results suggest that color discriminations may be important in squirrels' sensory ecology.

Experiment 2b

Experiment 2b was designed to test pigeons in a situation that would be as close an analogue as possible to that used with squirrels in Experiment 2a. Given the efficient learning described by Huber et al. (2005) in a similar situation, an automated version of the pokeboard was used, in which pecking illuminated areas of a touchscreen replaced perforating paper as an instrumental response. To ensure that the pigeons could reach all the cells in the display, the touchscreen array was smaller than the pokeboard. Subject to this, the stimuli and procedures were the same.

Method

Subjects

The subjects were six domestic pigeons, obtained as loft discards from local fanciers. They had previously been trained to peck a white disk on a touchscreen for food reward, but were otherwise experimentally naïve. They were maintained under the same conditions as the pigeons in Experiment 1a.

Apparatus and Stimuli

The same operant chambers were used as in Experiment 1b, controlled in the same way, except that the resistive touchscreens were replaced with infrared equivalents (CarrollTouch). The stimuli were arranged within a 105 × 105 mm area at the bottom of the touchscreen.

Procedure

In the first pretraining session, the pigeons were exposed to a centrally placed white disk of diameter 30 mm (the observing key). When the pigeon pecked this once, it disappeared, to be replaced by an array of 12 white 25 mm diameter outline circles, arranged in a hollow square with 1 mm separation. Centered within each circle was a filled white regular hexagon, measuring 8 mm across. If the pigeon gave two pecks in succession within any circle, the hexagon within that circle disappeared, and a 30-mm diameter filled white circle appeared to one side of the stimulus array; if the hexagon pecked was on the left side, this circle appeared to the left of the stimulus array, otherwise it was to the right. A single peck to the filled white circle was followed by a 2.5-s operation of the corresponding feeder, which contained hemp seed and pigeon

Table 4

Experiments 2a, 2b, and 2c: Proportions of Trials Out of 36 (Squirrels and Pigeons) or 12 (Humans) in Which Response Was Consistent With Each of Four Simple Hypotheses About the Subject/Participant's Categorization Strategy, and the Best-Fitting Hypothesis

	Consistent with control by				
	Color	Shape	Pattern	Overall similarity	Apparent strategy
Squirrels					
Mo	53%	75%	36%	64%	Shape
It	83%	39%	44%	67%	Color
De	75%	36%	31%	42%	Color
Sy	72%	44%	22%	39%	Color
Pigeons					
Zb	81%	53%	53%	86%	Overall similarity
Rv	75%	47%	36%	58%	Color
Ot	75%	58%	31%	64%	Color
Rr	100%	33%	33%	67%	Color
Ra	92%	47%	42%	75%	Color
Rp	97%	36%	31%	64%	Color
Students					
1	67%	50%	67%	83%	Overall similarity
3, 5, 6	67%	33%	67%	67%	
4	50%	67%	33%	50%	Shape
7	75%	42%	25%	42%	Color
9	67%	67%	67%	100%	Overall similarity
10, 24	83%	50%	50%	83%	
11, 21	67%	67%	33%	67%	
12	50%	50%	67%	67%	
13	100%	33%	33%	67%	Color
14	50%	50%	83%	83%	
15	50%	83%	50%	83%	
16	33%	67%	33%	33%	Shape
18, 19	83%	50%	17%	50%	Color
20	25%	75%	58%	58%	Shape
22, 23	50%	83%	50%	83%	

conditioning mixture. Once the hexagon within a circle had disappeared, it did not reappear within that trial, and further pecks to that circle had no scheduled consequences. The trial ended when all cells containing hexagons had been opened: at this point the array disappeared. After an intertrial interval that varied between 1 s and 5 s, the observing key reappeared, signaling the start of the next trial.

In subsequent pretraining sessions, only six randomly selected circles in the array contained hexagons. Making two pecks in succession at any of these circles led to reinforcement as described above: pecks to empty circles were followed by a 2.5-s period during which pecks had no scheduled consequences. When all the filled circles had been opened the trial terminated. Four trials were given per session, separated by 1 s to 5 s intertrial intervals, and one session was given per day, normally 5 days per week. Discrimination was assessed in terms of the ρ statistic of Herrnstein et al. (1976), calculated from the order in which the pigeons pecked at the circles.

Since all the pigeons reached a criterion of ρ exceeding 0.8 on three successive trials within a single session, training of the experimental discrimination began at the start of the next session, using the same stimuli as in Experiment 2a (see Figure 4), and the same procedure as in pretraining. At the beginning of each trial, each of the 12 circles included either a positive or a negative stimulus (six of each). The stimuli measured 8 mm across. For three of the pigeons, the positive stimuli were orange circles with a hash pattern; for the remaining

three pigeons, blue stars with a dotted pattern were positive. The same contingencies on pecks to positive and negative stimuli were in force as above. Once any stimulus had been pecked twice, it disappeared for the remainder of the trial, and further pecks to that circle had no scheduled consequences.

The same criterion of ρ exceeding 0.8 on three successive trials was set as for the squirrels, but in addition, each pigeon was given a minimum of three training sessions. All pigeons met the criterion within three sessions. Test sessions were then run in the same way as for the squirrels in Experiment 2a. Test stimuli were scored as being treated as positive or negative in the same way as in Experiment 2a.

Results and Discussion

All the pigeons learned the preliminary discrimination, of hexagons from blanks, to criterion level in a single four-trial session. The number of trials required to learn the experimental discrimination to the criterion of $\rho > 0.8$ in three successive trials ranged from 3 to 6.

Table 4 includes the results of the test trials, showing the proportions of the 36 trials for each pigeon in which behavior was consistent with each of the four simple hypotheses about what cues were controlling it. Like the squirrels in Experiment 2a, most of the pigeons behaved in a way that was most consistent with unidimensional control by the color cue. One pigeon's behavior was most consistent

with overall similarity, although even for that pigeon color appeared to be a stronger cue than the others. The proportion of pigeons whose behavior was most consistent with unidimensional control was higher than in Experiment 1a, though the difference is not statistically significant (Fisher's exact test, $p = .30$).

Experiment 2c

Experiment 2c was designed to provide a human analogue of Experiments 2a and 2b. To allow automated testing, the procedure was based on that used with pigeons in Experiment 2b rather than that used with squirrels in Experiment 2a where the two diverged. As in Experiment 1b, the time allowed for training the experimental discrimination was reduced to take into account human participants' more rapid learning.

Method

Participants

Twenty-four undergraduate students were recruited through the School of Psychology's participant panel. Participation was incentivized by a small money payment or course credit.

Apparatus, Stimuli, and Procedure

The apparatus was the same as in Experiment 1b, and the stimuli were the same as in Experiment 2b. The procedure was closely based on that used in Experiment 2b, and the same program was used to control the experiment and collect data. Participants responded by touching the screen. As in Experiment 1b, correct responses were followed by the presentation of a large smiley face in the center of the screen for 1 s; incorrect responses were followed by a corresponding interval during which the screen was inactive. The array remained on the screen during these events. The session structure was different from Experiment 2b. There was a single session, divided into two phases, a training phase and a test phase. The training phase was planned to consist of up to three blocks of four trials, each trial involving the display of a single array of stimuli. For half the participants, the reward was given following responses to circles marked by the orange circles with a hash pattern, and not for responses to circles marked by the blue stars with a dotted pattern; for the other half of the participants, these contingencies were reversed. Once a participant had reached a criterion of a ρ of 0.8 on three successive trials within a block, a test phase began; if a participant did not reach the criterion within three training blocks the experiment terminated at that point. The test phase consisted of six trials, during which all six one-away test stimuli were presented, two per trial, with the same constraints as in Experiment 2b, so that each test stimulus was associated with reward once and with nonreward once. In both training and test phases, stimuli were scored as being treated as positive or negative in the same way as in Experiments 2a and 2b.

Results and Discussion

Twenty-one of the participants reached the criterion of a ρ of 0.8 on three successive trials within a four-trial block in the first block of trials. The remaining three did not reach the criterion within three four-trial blocks, so their participation was ended before the

test phase began. Results of test trials are included in Table 4. It can be seen that two of the students showed performance best described by an overall similarity classification, seven showed performance best described by a unidimensional classification; and the remaining 12 could not be assigned to any simple strategy. It is possible that these participants were under the control of two dimensions but not all three, or that they were using some more complex strategy.

As predicted, unidimensional classification was more prevalent than overall similarity sorting for humans in Experiment 2c, in contrast to the dominance of overall similarity sorting in humans in Experiment 1b. This difference in relative prevalence of unidimensional classification between Experiment 2c and Experiment 1b is statistically significant, (Fisher's exact test, $p = .03$). As noted above, the prevalence of unidimensional sorting was also higher for pigeons in Experiment 2b than pigeons in Experiment 1a, although this difference is not statistically significant. Collapsing across pigeons and humans, the procedural differences between Experiments 1 and 2 had a significant effect on the prevalence of unidimensional sorting (Fisher's exact test, $p = .02$).

No clear effects of species on the relative prevalence of overall similarity and unidimensional classification were found. As noted in the discussion of Experiment 1b, the effect of species was not significant in Experiment 1, and it was also not significant in Experiment 2 (Fisher's exact test, $p = 1.0$).

General Discussion

An accurate understanding of the cognitive differences between humans and other animals is an essential preliminary to any discussion of the features of the proto-human ecological niche that may have resulted in the unique enlargement of the human brain.

The spur to the experiments described here was an apparent anomaly among such cognitive differences. Theorists of categorization have often assumed that the use of a single dimension to sort multidimensional stimuli is diagnostic for the use of analytic processing, as against a nonanalytic process which would result in classification on the basis of overall similarity (e.g., Kemler Nelson, 1984). It is also widely assumed that analytic processes are more characteristic of adult human cognition than nonhuman cognition. It follows from these assumptions that unidimensional classification should be more prevalent in humans than nonhumans, and that overall similarity classification should be more prevalent in nonhumans than in humans. One reason to doubt this inference is that, in numerous experiments on animal concept discrimination (e.g., Lea & Ryan, 1983; Lea et al., 1993; Lea et al., 2006), we and others have found that nonhuman animals tend to make such discriminations on the basis either of single dimensions or of a small subset of the available dimensions. Reviewing this literature, Lea and Wills (2008) accepted that it is reasonable to assume that humans can use rules and that animals will typically learn in a less analytic way, but argued on the basis of experimental data from both humans and other animals that if this is accepted, the use of unidimensional sorting as a criterion for analytic processing will have to be dropped.

Lea and Wills (2008, p.124) acknowledged a difficulty with this conclusion, which is that the experiments on which it rests have used very different procedures with humans and nonhumans. At least in humans, procedural details can make a great difference to sorting

strategies (e.g., Milton & Wills, 2004; Regehr & Brooks, 1995). Furthermore, in previous experiments we have found that when procedures are made more similar in human and nonhuman experiments, apparent species differences sometimes diminish or disappear (e.g., Goto et al., 2004). The purpose of the present experiments, therefore, was to compare human and nonhuman sorting strategies in experiments whose procedures would be as similar across species as we could make them, to see whether species differences would persist under those conditions. With these precautions taken, we found that humans were no more likely than other species to show unidimensional sorting. In Experiments 1a and 1b, there was no great difference between pigeons and students in the tendency to unidimensional sorting, but if anything more of the pigeons than the humans showed it. In Experiments 2a, 2b, and 2c, among those subjects where a clear determination could be made, again there were no clear species differences in prevalence of unidimensional sorting although, if anything it was more prevalent in pigeons and squirrels than in humans. Furthermore, the two sets of experiments used distinctly different procedures and stimuli, and there was a greater tendency to unidimensional sorting in Experiments 2b and 2c than in Experiments 1a and 1b, showing that the pattern of sorting was more sensitive to the differences in procedure and stimuli than to the difference of species. The methodology depends on the classification of subjects as unidimensional or overall similarity sorters, and given the pattern of results we found, this classification is subject to error. However we did observe differences in the proportions of the two types of sorters between the two sets of experiments. This means that the method clearly has enough power to detect at least large differences. There may still be small differences between species in their sorting strategies, which we were unable to detect, but they are clearly not as large as the differences between procedures.

Between them, the experiments control for some of the more obvious explanations for an enhanced tendency toward unidimensional sorting in the previous literature on nonhuman animals, as summarized by Lea & Wills (2008). First, there is nothing to suggest that the current results flow from any difference between bird and mammal visual systems, since the squirrels in Experiment 2a behaved in essentially the same way as the pigeons in Experiment 2b. Second, it was not the case that unidimensional sorting in nonhumans happens only because subjects could not discriminate or had not noticed all the stimulus dimensions, since in Experiment 1a the pigeons reached a demanding criterion of discrimination of all the dimensions used in the test stimuli. Third, within each series of experiments, procedures and stimuli were kept as similar as possible for the different species involved.

Of course, procedures can never be identical for species as different as pigeons, squirrels and humans; indeed, even if they could, we could not be sure that their effects were the same (Bitterman, 1960). One obvious way in which they differed was in the total amount of training given, and as Cook and Smith (2006) have shown, total exposure can make a big difference to the way stimuli are coded and subsequent generalization performance, in both pigeons and humans. However Cook and Smith's results give reasonable grounds for believing that training to the same criteria in the two species, as was done in our experiments, should produce similar codings. Methodologically, the results of the present experiments demonstrate the importance of making procedures parallel between species so far as that can be done: a pigeon-human comparison based on the results of Experiments 1a and 2c, for

example, would have given completely misleading results. Clearly the squirrel procedure of Experiment 2a differed more from the touchscreen procedures of Experiments 2b and 2c than they differed between themselves, and it would be desirable in future to establish touchscreen responding in squirrels so that they could be tested in the same way as pigeons. However the general similarity of the results of these three experiments suggests that the formal similarities among them were more important than the differences of instrumentation.

While procedures within each series of experiments were kept as similar across species as possible, the procedures used in the two series of experiments quite deliberately differed substantially. In part, these procedural differences flowed from the different primary purposes of each series: in the first series, our goal was to ensure that the pigeons were successfully discriminating all the dimensions of the stimuli; in the second, it was to include squirrels as a model nonprimate, diurnal mammal species. However the differences between them had the useful effect of demonstrating the malleability of the tendency to unidimensional sorting. Because of the multiple differences between the two experimental series, the origins of the differences in the results obtained remain to be explored in further experiments. In particular, it would be interesting to investigate the effects of the pretraining on individual elements in Experiment 1, the effects of exposure to multiple copies of the stimuli in Experiment 2, and the effects of using highly salient stimulus dimensions such as color. Lea et al. (2006) explored the first of these factors systematically, but not in relation to the tendency to unidimensional sorting. Some information on the second point is available from Huber et al. (2005), but again not in relation to the tendency to unidimensional sorting; and in any case a more detailed investigation of this powerful training technique would clearly be valuable. One obvious difference between the two experimental series is that Experiments 2a–2c used spatially integrated stimuli, whereas in Experiments 1a and 1b the dimensions were spatially separated; spatially integrated stimuli tend to produce more unidimensional sorting (Milton & Wills, 2004), which is indeed what we observed, but given the other differences too much should not be made of this.

Given our present experimental results, two conclusions are possible concerning the suggestion that unidimensional sorting can be taken as evidence of analytic processing, and overall similarity sorting as evidence of nonanalytic, processing (as proposed, e.g., by Kemler Nelson, 1984). The first possibility is to accept that argument, and conclude that pigeons, squirrels and humans all have similar levels of access to analytic processes. Alternatively, we reject that conclusion as implausible, and therefore reject the argument that unidimensional classification can be taken as evidence of analytic processing. That was the position taken by Lea and Wills (2008), and the present experimental results put their argument onto a firmer basis, by replacing comparisons between experiments with different procedures (which our present experiments confirm should not be trusted) by comparisons between parallel experiments with closely similar procedures.

Of course, the evidence is not yet conclusive. It is not inconceivable for a nonverbal to animal to use a rule; for example, Beckers et al. (2006) argued that rats can use causal inferential reasoning without verbal processes, and in subsequent experiments we have considered more fully what it might mean for a nonverbal animal to be using a rule in categorizing multidimensional stimuli

of the sort used in the present experiments (Lea, Wills, Leaver, Ryan, Bryant, & Millar, in press). But it is hard to believe that nonhumans are as a general matter as inclined to analytic processing as humans. Alternatively, it is conceivable that the present task happened to be one where humans prefer to use a nonanalytic strategy; but that would be an ad hoc assumption and we can see no reason to make it, and some reason to reject it since the procedure was closely related to the match to standards design, in which Milton et al. (2008) found evidence that humans typically use analytic strategies to make overall similarity sorts. We conclude that the present experiments strengthen the conclusion that we should not take unidimensional sorting as reliable evidence of analytic processing.

If we are to understand the uniqueness of human cognition, and therefore understand the unique ecological niche that may have produced it, it is important to explore differences between human and animal information processing. It is quite possible that one such difference is in the mode of processing of multidimensional stimuli, and it may be that future experiments will make such a difference evident. However the present experiments, using two quite distinct procedures, provide no basis for saying that humans deal with such stimuli in a way that marks them off from other warm-blooded vertebrates.

References

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

Beckers, T., Miller, R. R., DeHouwer, J., & Urushihara, K. (2006). Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference. *Journal of Experimental Psychology: General, 135,* 92–102.

Bitterman, M. E. (1960). Toward a comparative psychology of learning. *American Psychologist, 15,* 704–712.

Cardinal, R. N., & Aitken, M. R. F. (2001). *Whisker (version* 2). Cambridge, U.K.: Cambridge University Technical Services Ltd.

Carvalho, L. D., Cowing, J. A., Wilkie, S. E., Bowmaker, J. K., & Hunt, D. M. (2006). Shortwave visual sensitivity in tree and flying squirrels reflects changes in lifestyle. *Current Biology, 16,* R81–R83.

Cook, R. G. (2000). The comparative psychology of avian visual cognition. *Current Directions in Psychological Science, 9,* 83–89.

Cook, R. G., & Smith, J. D. (2006). Stages of abstraction and exemplar memorization in pigeon category learning. *Psychological Science, 17,* 1059–1067.

Emery, N. J. (2006). Cognitive ornithology: The evolution of avian intelligence. *Philosophical Transactions of the Royal Society, Series B, 361,* 23–43.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Garner, W. R. (1974). *The processing of information and structure.* Potomac MD: Erlbaum.

Goto, K., Wills, A. J., & Lea, S. E. G. (2004). Global-feature classification can be acquired more rapidly than local-feature classification in both humans and pigeons. *Animal Cognition, 7,* 109–113.

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,* 285–302.

Huber, L. (2001). Visual categorization in pigeons. In R. G. Cook (Eds.), *Avian visual cognition.* Retrieved March 30, 2008, from http://www.pigeon.psy.tufts.edu/avc/huber/

Huber, L., Apfalter, W., Steurer, M., & Prossinger, H. (2005). A new learning paradigm elicits fast visual discrimination in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 237–246.

Jacobs, G. H. (1976). Wavelength discrimination in gray squirrels. *Vision Research, 16,* 325–327.

Jenkins, H. M. (1973). Noticing and responding in a discrimination based on a distinguishing element. *Learning & Motivation, 4,* 115–137.

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior, 23,* 734–759.

Lea, S. E. G., & Kiley-Worthington, M. (1996). Can animals think? In V. Bruce (Eds.), *Unsolved mysteries of the mind,* pp. 211–244. Hove: Erlbaum (UK) Taylor and Francis.

Lea, S. E. G., Lohmann, A., & Ryan, C. M. E. (1993). Discrimination of 5-dimensional stimuli by pigeons: Limitations of feature analysis. *Quarterly Journal of Experimental Psychology, 46*B, 19–42.

Lea, S. E. G., & Ryan, C. M. E (1983). Feature analysis of pigeons' acquisition of concept discrimination. In M. L. Commons, R. J. Herrnstein, and A. R. Wagner (Eds.), *Quantitative analyses of behavior, Vol 4: Acquisition processes,* pp. 239–253. Cambridge, MA: Ballinger.

Lea, S. E. G., & Wills, A. J. (2008). Use of multiple dimensions in learned discriminations. *Comparative Cognition and Behavior Reviews, 3,* 115–133.

Lea, S. E. G., Wills, A. J., Leaver, L. A., Ryan, C. M. E., Bryant, C. M. L., & Millar, L. (2009). A comparative analysis of the categorization of multidimensional stimuli: I. Strategic information search. Manuscript submitted for publication.

Lea, S. E. G., Wills, A. J., & Ryan, C. M. E. (2006). Why are artificial polymorphous concepts so hard for birds to learn? *Quarterly Journal of Experimental Psychology, 59,* 251–267.

Longmore, C.A., Milton, F., & Wills, A. J. (2009). Analytic processes of overall similarity sorting in free classification. Manuscript submitted for publication.

Macdonald, I. M. V. (1992). Grey squirrels discriminate red from green in a foraging situation. *Animal Behaviour, 43,* 694–695.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19,* 242–279.

Milton, F., Longmore, C. A., & Wills, A. J. (2008). Processes of overall similarity sorting in free classification. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 676–692.

Milton, F., & Wills, A. J. (2004). The influence of stimulus properties on category construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 407–415.

Nosofsky, R. M., Clark, S. E., & Shin, H. J. (1989). Rules and exemplars in categorization, identification and recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 15,* 282–304.

Pinker, S. (2001). Four decades of rules and associations, or whatever happened to the past tense debate? In E. Dupoux (Ed.), *Language, the brain, and cognitive development: Papers in honor of Jacques Mehler* (pp. 157–179). Cambridge, MA: MIT Press.

Pothos, E. M. (2005). The rules versus similarity distinction. *Behavioral and Brain Sciences, 28,* 1–49.

Regehr, G., & Brooks, L. R. (1995). Category organization in free classification - the organizing effect of an array of stimuli. *Journal of Experimental Psycholog: Learning Memory & Cognition, 21,* 347–363.

Reynolds, G. S. (1961). Attention in the pigeon. *Journal of the Experimental Analysis of Behavior, 4,* 203–208.

Shanks, D. R., & Darby, R. J. (1998). Feature- and rule-based generalization in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes, 24,* 405–415.

Skinner, B. F (1969). *Contingencies of reinforcement: A theoretical analysis.* New York: Appleton-Century-Crofts.

Smith, J. D., & Kemler Nelson, D. G. (1984). Overall similarity in adults' classification: The child in all of us. *Journal of Experimental Psychology: General, 113,* 137–159.

Smith, L. B., & Kemler, D. G. (1977). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology, 24,* 279–298.

Spiegel, R., & McLaren, I. P. L. (2006). Associative sequence learning in humans*, Journal of Experimental Psychology: Animal Behavior Processes, 32,* 150–163.

Von Fersen, L., & Lea, S. E. G. (1990). Category discrimination by pigeons using 5 polymorphous features. *Journal of the Experimental Analysis of Behavior, 54,* 69–84.

Warner, F. E., & Lea, S. E. G. (2007, March). Concept-level generalization in monkeys following training on a single exemplar of a familiar category. Poster presented at the Conference of the Comparative Cognition Society. Melbourne, FL.

Yokoyama, K., Dailey, D., & Chase, S. (2006). Processing of conflicting and redundant stimulus information by pigeons. *Learning & Behavior, 34,* 241–247.

Zeigler, H. P., & Bischof, H.-J. (Eds.). (1993). *Bird vision and cognition.* Cambridge, MA: MIT Press.

New Editors Appointed, 2011–2016

The Publications and Communications Board of the American Psychological Association announces the appointment of 3 new editors for 6-year terms beginning in 2011. As of January 1, 2010, manuscripts should be directed as follows:

- *Developmental Psychology* (http://www.apa.org/journals/dev), **Jacquelynne S. Eccles, PhD,** Department of Psychology, University of Michigan, Ann Arbor, MI 48109
- *Journal of Consulting and Clinical Psychology* (http://www.apa.org/journals/ccp), **Arthur M. Nezu, PhD,** Department of Psychology, Drexel University, Philadelphia, PA 19102
- *Psychological Review* (http://www.apa.org/journals/rev), **John R. Anderson, PhD,** Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213

Electronic manuscript submission: As of January 1, 2010, manuscripts should be submitted electronically to the new editors via the journal's Manuscript Submission Portal (see the website listed above with each journal title).

Manuscript submission patterns make the precise date of completion of the 2010 volumes uncertain. Current editors, Cynthia García Coll, PhD, Annette M. La Greca, PhD, and Keith Rayner, PhD, will receive and consider new manuscripts through December 31, 2009. Should 2010 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 2011 volumes.