Multiple Feature Use in Pigeons' Category Discrimination: The Influence of Stimulus Set Structure and the Salience of Stimulus Differences

Stephen E. G. Lea
University of Exeter
Emmanuel M. Pothos
Swansea University
Andy J. Wills, Lisa A. Leaver,
Catriona M. E. Ryan, and Christina Meier
University of Exeter

Accepted for publication, subject to completion of administrative procedures, in the *Journal of Experimental Psychology: Animal Learning and Cognition.* Date of editorial acceptance 27th January 2018.

Author Notes

Stephen E. G. Lea, School of Psychology, University of Exeter, UK; Emmanuel M. Pothos, School of Psychology, Swansea University, UK; Andy J. Wills, Lisa A. Leaver, Catriona M. E. Ryan, and Christina Meier, School of Psychology, University of Exeter, UK.

Emmanuel Pothos is now at the Department of Psychology, City University of London; Andy Wills is now at the School of Psychology, University of Plymouth, UK; Christina Meier is now at the Department of Psychology, University of Manitoba, Canada

Experiment 1 was supported by the New and Emerging Science and Technologies activity of the European Community Framework Programme 6, under grant no. 516542, "From Associations to Rules in the Development of Concepts" (FAR). We are grateful to Eloise Nicklin, Dominique Pender and Kelly Warr for conducting a pilot experiment for Experiment 2; to Ruth Dakin and Guido de Filippo for helping run the pigeons in Experiment 2; and to Catherine Bryant for the initial development of the client programs used in the experiments. All animal housing and testing procedures were in accordance with the guidelines of the Association for the Study of Animal Behaviour and applicable UK and EU law. The authors declare that they have no conflict of interest.

Correspondence concerning this article should be addressed to S. E. G. Lea, University of Exeter, Department of Psychology, Washington Singer Laboratories, Exeter EX4 4QG, United Kingdom. E-mail : s.e.g.lea@exeter.ac.uk

Abstract

Two experiments investigated what makes it more likely that pigeons' behavior will come under the control of multiple relevant visual stimulus dimensions. Experiment 1 investigated the effect of stimulus set structure, using a conditional discrimination between circles that differed in both hue and diameter. Two training conditions differed in whether hue and diameter were correlated in the same way within positive and negative stimulus sets as between sets. Transfer tests showed that all pigeons came under the control of both dimensions, regardless of stimulus set structure. Experiment 2 investigated the effect of the relative salience of the stimulus differences on three visual dimensions. Pigeons learned a multiple simultaneous discrimination between circular patches of sinusoidal gratings that differed in hue, orientation and spatial frequency. In initial training, each stimulus only included one positive or negative feature, and the stimulus differences on the three dimensions were adjusted so that the rates of learning about the three dimensions were kept approximately equal. Transfer tests showed that all three dimensions acquired control over behavior, with no single dimension dominating consistently across pigeons. Subsequently the pigeons were trained in a polymorphous category discrimination using all three dimensions, and the level of control by the three dimensions tended to become more equal as polymorphous training continued. We conclude that the salience of the stimulus differences on different dimensions is an important factor in whether pigeons will come under the control of multiple dimensions of visual stimuli.

Keywords: Category learning, attention, conditional discrimination, polymorphous categories, pigeons

Multiple Feature Use in Pigeons' Category Discrimination

The experiments described in the present paper investigate the circumstances under which multiple stimulus dimensions acquire equal, or unequal, control over pigeons' discrimination behavior. This question has both an empirical and a theoretical motivation. Empirically, it has been found that in some experiments pigeons come under the control of all stimulus dimensions whose variation is correlated with reinforcement, whereas in other experiments they do not; in the extreme case, a single stimulus dimension controls behavior. As will be explained further below, two quite different theoretical interpretations have been placed upon the finding of control by a limited number of stimulus dimensions. On the one hand it has been seen as a sign of cognitive limitation (the inability to attend to multiple dimensions), on the other it is taken as a sign of an advanced cognitive process (categorization by rules). Our aim in this paper is to identify some conditions that favour control by multiple or single dimensions, and thereby to clarify the appropriate theoretical interpretation of such limited-dimension control.

Three experimental approaches have been adopted to address this question so far:

(1) Provision of multiple dimensions, none of which is individually perfectly predictive of reinforcement, but which together allow optimal discrimination. This is the approach of setting pigeons to learn polymorphous discriminations (e.g. Jitsumori, 1993, 1996; Lea & Harrison, 1978; Lea, Lohmann, & Ryan, 1993; von Fersen & Lea, 1990), conjunctive discriminations (e.g. Teng, Vyazovska & Wasserman, 2015; Vyazovska, Teng & Wasserman, 2014), or other multidimensional discriminations (e.g. Herbranson, Fremouw, & Shimp, 1999, 2002; Smith, Ashby, Berg, et al., 2011)..

(2) Provision of multiple dimensions, all of which are equally and perfectly predictive of reinforcement. This was the approach taken in much of the older literature (e.g. Butter, 1963; Chase, 1968; Chase & Heinemann, 1972; Farthing & Hearst, 1970; Newman & Baron,

1965; Reynolds, 1961) and also in some more recent papers (Lea, Wills, Leaver, Ryan, Bryant, & Millar, 2009, Experiments 1a, 1b & 2; Wills, Lea, Leaver. et al., 2009).

(3) Provision of multiple dimensions that differ in how well they predict reinforcement, so that the impact of validity can be assessed in relation to other relevant variables such as salience and attention. This approach was taken in the experiments of Lea et al. (2009, Experiments 3a & 3b) and Nicholls, Ryan, Bryant, & Lea (2011, Experiment 2).

Taken together, these three strands of experimental work show that control by multiple stimulus dimensions is much more likely to be seen when it is necessary for perfect discrimination. In polymorphous or conjunctive tasks, control by all or at least several of the available dimensions is commonly seen (e.g. Jitsumori, 1993, 1996; Lea & Harrison, 1978; Teng et al., 2015; Vyazovska et al., 2014). Where control by a single dimension is sufficient for perfect discrimination, however, experiments from the classic studies of Reynolds (1961) to more recent ones such as those of Wills et al. (2009) and Nicholls et al. (2011) concur in finding that some birds seem to come under the control of only a single dimension. This suggests that, for pigeons, coming under the control of multiple dimensions is in some sense difficult or effortful, so it will only happen when it is necessary to maximise reward rate.

But this is not the whole story. Making multiple dimensions necessary for perfect discrimination is not always sufficient to ensure that they will all be used: sometimes one or more dimensions never acquire control over behavior (e.g. Lea et al., 1993), and sometimes special training is required to ensure that all of them do (e.g. von Fersen & Lea, 1990). Furthermore, when multiple dimensions have different levels of correlation with reward, pigeons do not always direct their behavior towards the best correlated cues (Lea, et al., 2009; Nicholls et al., 2011).

On the other hand, there is at least one experiment in which pigeons might be using both stimulus dimensions even though optimal performance can be achieved using just one dimension. Smith et al. (2011) used a large set of stimuli drawn from a two-dimensional stimulus space (specifically, Gabor patches differing in bar width and orientation). They reported that pigeons learned a single-dimension classification and a multi-dimensional classification at rates that were not significantly different from each other (in contrast, adult humans would be expected to learn the single-dimension classification much more quickly). Smith et al. (2011) interpret their results as showing that pigeons treated these stimuli as unanalyzed wholes, and hence (in a holistic sense) were using both dimensions even when only one was diagnostic.

The question of when multiple dimensions will be used in categorization has theoretical significance because an influential account of human categorization (Ashby, Alfonso-Reese, Turken, & Waldron, 1998) claims that when humans are discriminating categories by means of a verbal rule, a single-dimensional rule is the most likely choice, at least initially. However, when we solve the task by associative learning, all available dimensions will typically influence categorization responses; this is referred to as overall similarity or family resemblance categorization. This account has been challenged: for example, Milton, Longmore, and Wills (2008) showed that some conditions that putatively favor more advanced cognitive processing in human categorization nonetheless lead to less unidimensional control. Lea and Wills (2008) therefore argued that unidimensional control could not be taken as an unambiguous indicator of rule use. However, the idea that associative acquisition of discrimination will automatically use all available information remains highly influential (e.g. Ashby & Maddox, 2011); and since it seems much more likely that pigeons will solve a task by associative learning than by anything analogous to a verbal rule, we would expect pigeons to rely consistently on all available dimensions.

So at the theoretical level, we have a paradox: on the one hand, control by a single stimulus dimension is being taken as the sign of an advanced cognitive process (verbal rule learning), and on the other, it seems to be a sign of a limited cognitive capacity (inability to attend to multiple dimensions). At the empirical level, we have some confusion, with control by multiple dimensions occurring in some experiments but not others. The present paper aims to reduce this confusion, by investigating two manipulations that might make control by multiple dimensions more or less likely. We introduce two relatively novel procedures, both designed to bring factors not considered in previous research under experimental control, so that clearer predictions can be made. In both procedures, we ensured that control by a single dimension could not result from the pigeons simply not seeing some aspects of the stimulus, by using stimuli in which all dimensions were properties of the entire stimulus, rather than parts of it (a possible criticism of some of our earlier experiments, e.g. Lea et al., 2009, Wills et al., 2009).

In Experiment 1, we used two completely valid dimensions, and sought to control the impact of stimulus set structure. Several investigators have argued that the same kind of stimuli can bias towards either unidimensional or multidimensional classification, depending on the particular similarity relations between the stimuli employed (Ashby & Gott, 1988; Pothos & Close, 2008; cf. Herbranson et al., 1999, 2002, for arguments that Ashby and Gott's procedures can be adapted for use with pigeons). Potentially, this might explain why some authors obtain unidimensional and others multidimensional classification. We sought to test this possibility by training two groups of pigeons with stimulus sets of quite different structure (illustrated in Fig. 1). To ensure that any differences were not due to gross differences in average similarity, we controlled the similarity relations between the stimuli of Experiment 1 so that no preference for unidimensional rather than multidimensional classification was predicted on the basis of two formal models (those of Pothos & Chater, 2002, and Rosch & Mervis, 1975).

In Experiment 2, we combined the multiple valid dimensions approach with the polymorphous training approach, and sought to control the impact of the relative salience of the stimulus differences on the different dimensions. Pigeons were first trained to solve three separate unidimensional discriminations, as in the "features in parallel" condition of Lea, Wills and Ryan (2006). The training procedure used allowed the stimulus control achieved by each dimension to be measured as training progressed, and the conditions were adjusted to ensure that control was equated across all three. We then introduced two further training conditions, using stimuli in which all three dimensions were varied. In the first of these new conditions, all dimensions were valid. In the second, the dimensions were combined so as to form a polymorphous training set. This allowed us first to examine whether control by all three dimensions continued despite the preceding training with all dimensions valid, and then to examine whether control by multiple dimensions would be further strengthened by the polymorphous contingencies, which require control by all dimensions if complete accuracy is to be achieved.

Across the two Experiments, therefore, we consider two factors that both potentially affect unidimensional vs. multidimensional classification, and that complement each other, in that one focusses on stimulus relations and the other on individual stimulus properties.

Experiment 1

In the first experiment, stimuli were defined in a two-dimensional stimulus space, as in the experiments of Smith et al. (2011) and many previous experiments with humans. As shown in Figure 1, the stimulus categories were defined on two dimensions, in such a way that the category boundary could be placed parallel to either stimulus dimension, or diagonally, and still allow perfect discrimination. To allow a more sensitive test of the models, two different stimulus sets were used. The application of formal models can help provide specific predictions with respect to two critical design issues: first, for each stimulus set, is there a prediction that individuals will learn at different rates, depending on whether they come under control of a single dimension or both ? Second, is the learning rate for one stimulus set predicted to be equivalent or not to the learning rate for the other set?

Figure 1 about here

Formal modelling may challenge intuition in the following way. Both sets have the property that each dimension on its own is sufficient for perfect discrimination. However, in Figure 1(a), the two stimulus dimensions are monotonically related across the entire set of stimuli, but in Figure 1(b) this is not the case. When only one of the stimulus dimensions is considered, the distributions of intra-category and inter-category similarities between stimuli are identical. But when both dimensions are considered, those distributions differ between the two kinds of stimulus set. At an intuitive level, it might be expected that, if both dimensions are involved in the discrimination, rates of learning to discriminate the two kinds of stimulus set might differ. For example, with the non-monotonically related dimensions, we have a visual impression of two distinct clusters of stimuli when they are viewed in two dimensions, but not when they are viewed in a single dimension, suggesting that an unsupervised learning mechanism should favor two dimensions. Alternatively, monotonically related dimensions may facilitate the emergence of synthetic dimensions, which may aid learning.

More detailed theoretical analysis, however, suggests a different prediction. Two distinct computational frameworks predict that, regardless of whether one or two dimensions are used to solve the problem, it should be equally easy to learn to discriminate the two kinds of stimulus set and, moreover, that there should be no difference in category learning difficulty between the two stimulus sets. Both models are based on the same idea, comparing within-category similarity to between-category similarity, with 'better' (easier to learn and more intuitive, cf. Pothos, Edwards & Perlman, 2011a) categorizations reflecting greater within-category similarity. Regarding the prediction of unidimensional classification vs. classification using both dimensions, the approach in both models is to compare 'intuitiveness' of a categorization when the stimuli are represented with each dimension and with both dimensions (Pothos & Close, 2008). For example, if a required categorization is predicted to be most intuitive using only the y dimension, instead of x or x and y, then this is the prediction for the way the stimuli would be represented.

The first computational framework is Pothos and Chater's (2002) simplicity model. The model assumes that grouping behavior is driven by a prerogative to *compress* the information contained in the similarity relations between a set of objects (cf. Atick, 1992). An information code is constructed for describing all similarity relations in a stimulus set without categories vs. with categories, assuming that categories imply all within category similarities to be greater than between category ones (if there are erroneous constraints, then a further information code is required for correction and a code is required to specify the classification as well).

The second computational framework is based on Rosch and Mervis's (1975) proposal that more intuitive categories are those with greater within-category similarity and lower between-category similarity. This approach is similar to Pothos and Chater's (2002), except for the fact that instead of employing information theory to balance a description with and without categories, a simple index of category intuitiveness is computed as an average of within vs. between category similarities (details in Pothos, Perlman, Bailey et al., 2011b).

These two models for categorization do not always converge (Pothos et al., 2011b), but they do so when applied to the stimulus sets shown in Figure 1. The models make the surprising prediction that the two kinds of discrimination should be equally easy, and that it should be equally easy to make either kind of discrimination based on one dimension or on two.

While it is fruitful to motivate categorization predictions from formal models, several complications arise. First, Figure 1 is based on the physical properties of the stimuli, and we do not know the precise relationship between distance on a physical stimulus dimension and psychological similarity for pigeons. Second, we do not know how the overall similarity between two stimuli might depend on their similarity in separate dimensions – for example, whether similarity would be determined by e.g. a Euclidean, city-block or Tchebyshev metric. Soto and Wasserman (2010) suggest that different metrics will apply with different combinations of stimulus dimensions. Third, the predictions from such categorization models tend to be more robust with greater numbers of stimuli. Notwithstanding the above points, to a reasonable approximation, the basic prediction that these stimulus structures create no bias between single- and multiple-dimensional learning, and no bias between each stimulus structure, should not depend much on these issues.

Our formal models thus make two predictions:

(i) There should be no difference in the ease of learning the two kinds of discrimination shown in Figure 1(a) and 1(b). This prediction was tested by comparing the rate of learning by separate groups of pigeons, trained with the two categorizations.

(ii) There should be no bias towards using either a single dimension, or both dimensions, in making the discrimination: unidimensional and overall similarity categorization should be equally likely, and roughly equal numbers of individuals should therefore show the two types of control. This prediction stands in contrast to what we would expect from previous research: according to the conclusions of Smith et al., (2011), there should be a strong bias towards control by both dimensions; and according to the conclusions of Lea and Wills (2008), there should be a strong bias towards unidimensional control. These contrasting predictions were tested by carrying out transfer tests once training was complete, using stimuli located in parts of the stimulus space unused during training.

In addition to examining the effect of stimulus structure, in Experiment 1 we took steps to avoid either of the dimensions we were studying dominating behavior because of differences of perceptual salience. In some previous experiments with animals, color cues have proved to be highly salient (e.g. Wills et al., 2009). Although we used color in Experiment 1, we used relatively small color differences, to avoid any bias towards unidimensional control due to simple stimulus factors.

Method

Subjects. Eight pigeons were used. They were obtained as discards from local fanciers, and had no previous experimental experience of discriminating the stimulus dimensions used in the present investigation. The pigeons were housed in an indoor aviary, measuring 2.2 m by 3.4 m by 2.4 m, and given constant access to water and grit. Prior to testing each day, the pigeons were held in individual cages in which they had access to water and were released back into the aviary when testing had finished for the day. All pigeons were maintained on a 12:12 hr light/dark cycle at 95% of their free-feeding weight.

Apparatus. Four identical operant chambers (internal size; 640 mm x 430 mm x 470 mm) were used. Each consisted of a plywood box, with a 15-inch (39-cm) resistive touch-monitor (Elo Touchsystems Accutouch model 1547L) mounted in the front wall; its base was 18.5cm above the grid floor of the box. The screen resolution was set to 1024 x 768 pixels, so that 1 cm on the screen corresponded to 30 pixels. The monitor was controlled by software written in Visual Basic using the Whisker control system (Cardinal and Aitken 2001) running on a computer supplied by Quadvision (Quadvision Ltd., Dorset, UK). Two food hoppers, positioned one on each side of the main screen, could be used to deliver a 2:1 mixture of hemp seed and conditioner to the pigeons, for 2.5 seconds. Each box had a webcam fitted into the side wall, 250 mm above the floor, allowing the pigeons' behavior to 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.

Stimulus Materials. The stimuli were drawn from a population of 49 colored circular patches, which differed in diameter and hue, with 7 levels of each dimension. The diameters used ranged from 100 to 400 pixels (3 to 12 cm), in a logarithmic series in which each successive diameter was 26% greater than the previous one. The hues were generated using only the green and blue guns in the computer display, in the following series: (168, 70), (166, 126), (164, 147), (163, 164), (159, 180), (153, 205), (143, 252). These were selected by visual matching to Munsell chips of constant brightness to the normal human eye, across a range from a definite green to a definite blue. There are many uncertainties in mapping between an RGB coding for a monitor and its perceptual effect for an eye whose fundamental receptors are different from those of a human, but by using only two of the monitor's three guns, and restricting the ranges at which they were used, we aimed to generate stimuli that for both pigeons and humans showed a smooth gradation between two distinct hues. Wright and Cumming (1971) argue that the pigeon has a transition point between hues at around 545nm, somewhat lower in wavelength than the transition between yellow and green to the normal human eye, but higher than the wavelength that would be matched, by a human, to the hue obtained by using the green gun alone on a typical monitor. In the rest of this paper, the stimuli are designated by number pairs (Hue-*x*, Diameter-*y*) where *x* designates a level of hue, from the greenest stimulus (168,70) as 1, to the bluest (143, 252) as 7 and *y* designates a level of diameter, from 100 pixels as 1, to 400 pixels as 7.

Experimental Design. The pigeons were trained on a conditional discrimination in which discriminative stimuli were displayed in a central area of the touchscreen; in the presence of some stimuli, reinforcement was available for pecks on a white target area to the left of the stimulus, and in the presence of others, reinforcement was available for pecks on a similar area to the right of the stimulus. Each pigeon was trained with a different set of stimuli as right-positive. For four of the pigeons, larger diameter circles were right-positive and smaller diameter circles were left-positive, and for the other four pigeons these contingencies were reversed. Similarly, for four of the pigeons bluer circles were right-positive and greener circles were left-positive, and for the other four pigeons these contingencies were reversed. These two assignments were counterbalanced, generating four groups of two pigeons. Within each of these groups, one of the pigeons was trained with monotonically related stimulus dimensions (as in Panel a of Figure 1), and the other with non-monotonically related dimensions (Panel b). For example, for Pigeons Le and Sf, greener, larger circles signalled that food was available on the left, while bluer, smaller circles signalled that food was available on the right. Pigeon Le had monotonically related stimulus dimensions, so stimuli (Hue-1, Diameter-5), (Hue-2, Diameter-6) and (Hue-3, Diameter-7) signalled food on the left and stimuli (Hue-5, Diameter-1), (Hue-6, Diameter-2) and (Hue-7, Diameter-3) signalled food on the right. Pigeon Sf had the corresponding non-monotonically related stimulus dimensions sets, so stimuli (Hue-1, Diameter-7), (Hue-2, Diameter-6) and (Hue-3, Diameter-5) signalled food on the left, and (Hue-5, Diameter-3), (Hue-6, Diameter-2) and (Hue-7, Diameter-1) signalled food on the right.

Procedure. The pigeons were trained by conventional means first to find food in both the food hoppers, then to peck a white circular target area next to each food hopper to obtain food from that hopper, and finally to peck a white circular target area to turn on one or other of the side targets, which they could then peck to obtain food.

Discrimination training sessions each consisted of 96 trials, separated by a variable inter-trial interval of between 10 and 30s during which the touch screen was blank. At the beginning of each trial, one of the training stimuli appeared in the central target area. The 6 training stimuli were used in a pseudo-random sequence, constrained so that no more than three consecutive stimuli would be drawn from one set (left-positive or right-positive). Pecks at the stimulus were reinforced, on a *tand* FT 10s VI 3s schedule, by the presentation of both the white target areas. The first peck to the correct target area was reinforced by the presentation of the food hopper, with illumination, with the touch screen going blank. Pecks to the incorrect target area were recorded but had no scheduled consequences.

Discrimination training continued for each pigeon until it reached a criterion of 80% correct in 3 successive sessions. Following this, three transfer sessions were given. Four different types of stimuli were included in the transfer tests, and their relationships are illustrated in Figure 2.

Figure 2 about here

(a) The 6 training stimuli, used under test contingencies of reward (stimuli marked "A" and "B" in Figure 2)

(b) Four non-conflict stimuli that combined the stimulus dimensions in the same general way as the training stimuli, but used combinations not seen during training (stimuli marked "TN" in Figure 2). In practice the combinations used were those that had been used in training for the other stimulus dimension relationship group using the same stimulus dimension assignments. So, for example, the pigeons that had been trained with the monotonically related stimulus dimensions (Hue-1, Diameter-1), (Hue-2, Diameter-2) and (Hue-3, Diameter-3) versus (Hue-5, Diameter-5), (Hue-6, Diameter-6) and (Hue-7, Diameter-7) were tested with the stimuli (Hue-3, Diameter-1), (Hue-1, Diameter-3), (Hue-5, Diameter-7) and (Hue-7, Diameter-5), which had been used in training with the corresponding non-monotonically related stimulus dimensions.

(c) The ambiguous stimulus (Hue-4, Diameter-4), using the middle values of both stimulus dimensions, neither of which was used in any training stimulus (stimulus marked "?" in Figure 2).

(d) Ten stimuli in which the dimensions that had been used in training were put into conflict (stimuli marked as "T1", "T2", "T3", "T4", and "TC" in Figure 2). The stimuli chosen were those that had been used in training other pigeons with different assignments of the stimulus dimension, with both monotonic and nonmonotonic relationships between the stimulus dimensions. So, the pigeons that had been trained with the monotonically related stimulus dimensions (Hue-1, Diameter-1), (Hue-2, Diameter-2) and (Hue-3, Diameter 3) versus (Hue-5, Diameter-5), (Hue-6, Diameter-6) and (Hue-7, Diameter-7), and those that had been trained with the nonmonotonically related stimulus dimensions (Hue-1, Diameter-3), (Hue-2, Diameter-2) and (Hue-3, Diameter-1) versus (Hue-5, Diameter-7), (Hue-6, Diameter-6) and (Hue-7, Diameter-5), were all tested with the ten stimuli (Hue-1, Diameter-7), (Hue-3, Diameter-6), (Hue-2, Diameter-5), (Hue-1, Diameter-5), (Hue-3, Diameter-5), (Hue-7, Diameter-1), (Hue-7, Diameter-3), (Hue-5, Diameter-2), (Hue-5, Diameter-1) and (Hue-5, Diameter-3).

For each pigeon there were thus 15 transfer stimuli, plus the 6 training stimuli used under test conditions. Each of these stimuli was used in three trials, making 63 test trials in all. These were distributed between six test sessions, with the tests of the non-conflict and ambiguous stimuli occurring in the first and the conflict tests in the second, and so on. Test sessions were separated by at least two sessions of training conditions to ensure that performance remained at criterion. Each test session included an opening and closing block of 24 trials using training stimuli and contingencies only. Test trials were separated by at least two trials in which normal training stimuli and contingencies were presented. On all test trials, the first choice peck made was reinforced regardless of which target was pecked. The sessions involving non-conflict and ambiguous stimulus tests consisted of a total of 88 trials, and those involving conflict tests of 112 trials.

Results

Training. The number of sessions required to meet criterion ranged from 6 to 13 in the group with monotonically related stimulus dimensions, and from 7 to 13 in the group with non-monotonically related stimulus dimensions (including the three sessions at about 80% performance that were required to meet the criterion) Figure 3 shows the mean performance of the pigeons in the first 18 training sessions (the number that the fastest-learning pigeon received); these included some sessions given after transfer had started. It can be seen that in the first few sessions, the pigeons trained with monotonically related stimulus dimensions performed slightly less well than those trained with the nonmonotonically related stimulus dimensions, but performance converged by Session 8. By Session 12 mean performance for both groups was at 90% accuracy. The number of sessions required to reach criterion did not differ significantly between the groups (2-tailed Mann-Whitney test, $U = 6$, $N_1 = N_2 = 4$, $P > .05$), and a Bayesian t-test indicated that the data provided little evidence either against or in favour of the null hypothesis (Bayes $\text{factor}_{10} = 0.541$ against a non-informative prior).

Figure 3 about here

Transfer. Mean accuracy of responding to the training stimuli under test conditions was 91% (range 85% to 97%). This did not differ from the mean accuracy of responding to the same stimuli under training conditions during the test sessions (mean 89%, range 83% to 95%; 2-tailed Wilcoxon test, $T = 11$, $N = 8$, $P > 0.05$). Accuracy of response to the non-conflict transfer stimuli was high. Its mean over all pigeons was 88% (range 56% to 100%). The low value for one pigeon was traced to a procedural error, in which the pigeon was given some training sessions with the wrong reinforcement contingencies for Hue; it has been retained as a conservative assumption). This mean did not differ significantly from the accuracy of response to training stimuli under test conditions (2-tailed Wilcoxon test, $T = 16$, $N = 8$, $P > 0.05$).

The ambiguous stimulus (Hue-4, Diameter-4), involving the middle value of both stimulus dimensions, was responded to in the way that was appropriate for a greener stimulus on 54% of trials (range across pigeons 20% to 83%), and in the way that was appropriate for a smaller stimulus on 46% of trials (range 17% to 80%). Neither of the means differed significantly from 50% (2-tailed 1-sample Wilcoxon tests, $T = 10$ and $T = 12$ respectively, $N = 8$, $P > 0.05$ in each case).

Responses to the conflict stimuli were classified as hue-appropriate or diameter-appropriate. For example, if a pigeon that had been trained with small green circles as right-positive pecked to the right when given a large green circle, that response was classified as hue-appropriate. The mean proportion of hue-appropriate responses was 38%. Proportions for individual pigeons were 68%, 25%, 15%, 40%, 35%, 58%, 32% and 35%. The trend towards a majority of diameter-appropriate responses was not significant across pigeons (one-sample Wilcoxon test, $T = 5$, $N = 8$, $P > 0.05$). However the proportion of responses that were hue-appropriate differed significantly from 50% for four of the individual pigeons ($\chi^2(1)$ values > 3.84); by virtue of the additive property of the chi-square distribution (Weatherburn, 1957, p. 177), the sum of the $\chi^2(1)$ values for the eight pigeons can be tested as a $\chi^2(8)$ value. Pooling in this way gives a $\chi^2(8)$ value of 43.21 ($P < 0.001$), confirming that the individually significant results can be accepted despite the lack of significance in the other subjects. In three of the pigeons the significant tendency was for diameter-appropriate responses, in the fourth it was for hue-appropriate responses. The proportions of conflict stimuli responded to in a hue-appropriate way did not differ significantly between the training conditions (2-tailed Mann-Whitney test, $U = 7.5$, $N_1 = N_2 = 4$, $P > .05$).

The conflict tests were subdivided into those where, in terms of the physical properties of the stimulus, the hue dimension gave a relatively strong cue while the diameter dimension gave a weaker one, those where the diameter dimension gave a stronger cue and the hue dimension a weaker one, and those where the two dimensions gave equally strong cues. For example, for pigeons trained with the monotonically related stimulus sets (Hue-1, Diameter-1), (Hue-2, Diameter-2) and (Hue-3, Diameter-3) versus (Hue-5, Diameter-5), (Hue-6, Diameter-6) and (Hue-7, Diameter-7), the conflict stimulus (Hue-1, Diameter-5) was one where the hue dimension gives the stronger cue; the stimulus (Hue-3, Diameter-7) was one where the diameter dimension gives the stronger cue; and in the stimuli (Hue-1, Diameter-7), (Hue-2, Diameter-6) and (Hue-3, Diameter-5) the two dimensions are equally strong. In all, 73% of the responses to test stimuli where hue gave the stronger cue were hue-appropriate (range across pigeons 38% to 100%), and only 14% (range 0% to 38%) of the responses to test stimuli where diameter gave the stronger cue were hue-appropriate. All eight pigeons made more hue-appropriate choices when hue was the stronger cue than when diameter was the stronger cue, so the trend was significant (2-tailed binomial test, $P < 0.01$). For the stimuli where the two dimensions gave equally strong cues, 34% (range 8% to 67%) of responses were hue-appropriate.

Discussion

The training results did not detect any difference in speed of acquisition between the monotonically related stimulus dimensions condition and the non-monotonically related stimulus dimensions condition. This makes it unlikely that the pigeons were basing their discrimination on any kind of perceptual synthesis of the two dimensions, since that would have been easier to achieve with the monotonically related stimulus dimensions. This null result is consistent with the predictions of our formal models. Given our small sample size, it is of course possible that there might be a small difference in ease of learning between the two groups, which we did not detect

The transfer results imply that the pigeons consistently came under the control of both dimensions, and to a roughly equal extent. This result is in line with the conclusions of Smith et al. (2011) about pigeons, even though our procedure was quite different, with a much smaller total population of stimuli involved in training, and it confirms that multidimensional control can occur under conditions where it is not necessary for perfect discrimination. Like the results of Smith et al., our data contrast with results from non-human primates (e.g. (Smith, Beran, Crossley, Bloomer & Ashby, 2010; Smith, Crossley, Boomer, Church, Beran and Ashby, 2012), They also contrast with the predictions of our formal models, which predicted that categorization should have been equally likely to be single-dimensional or two-dimensional. So, the conclusion from the present results is that it is not stimulus set structure that predicts whether pigeons come under unidimensional or multidimensional control.

The results of our transfer tests contrast also with many previous results from birds learning multidimensional discriminations. For example, Lea et al. (2006) exposed chickens to a 5-dimensional polymorphous discrimination, and found that behavior when all dimensions were available was largely controlled by a color cue, despite the fact that the chickens could be shown to be able to discriminate all five dimensions. Wills et al. (2009, Experiment 2a) found that when pigeons had three perfectly valid stimulus dimensions available, five out of six pigeons classified conflict stimuli in terms of a single dimension, and for all five it was color. The results of the present experiment show that color will not always be the most salient dimension in a multidimensional discrimination task; the high levels of salience for color in some previous experiments may have resulted from using very large differences of hue, whereas in the present experiment the differences were deliberately made moderate.

Moreover, it was not just the case that either dimension could acquire control over behavior; the behavior of individual pigeons came under the control of both stimulus dimensions. For every pigeon, behavior in cue-conflict transfer trials was more likely to be controlled by the hue dimension than the diameter dimension when the hue cue was strong and the diameter value was weak, and more likely to be controlled by the diameter dimension than the hue dimension when the reverse was the case. Thus, the pigeons had learned about both dimensions. Unlike the preponderance of the pigeons (and squirrels and undergraduates) studied by Wills et al. (2009), but like the rhesus monkeys studied by Couchman, Coutinho and Smith (2010) and the pigeons studied by Smith et al. (2011), the pigeons in the present experiment responded to the cue conflict stimuli according to "Family Resemblance" or "Overall Similarity" rather than unidimensionally. This result contrasts with the prediction of no bias for unidimensional vs. multidimensional behavior, from the categorization models we employed: the pigeons showed a clear bias towards using multiple dimensions.

As discussed in our General Introduction, pigeons sometimes come under the control of multiple dimensions when they are available, and sometimes do not. The present experiment suggests that in at least some cases, unidimensional control may result from uncontrolled differences in the salience of the stimulus differences on the different dimensions involved. With the saliences of the differences along two dimensions carefully balanced, and with stimulus sets that should not be biasing towards either unidimensional or multidimensional control, we found reliable and consistent multidimensional control. In our next experiment we pursued the question of balancing salience between dimensions experimentally.

Experiment 2

Experiment 1 showed that, if the differences between stimulus values on dimensions were moderate, each of two dimensions would acquire roughly equal control over pigeons' behavior in a category discrimination task, even though a single dimension would be sufficient for perfect discrimination. Experiment 2 continued the approach of controlling the salience of stimulus differences, and added manipulations of the contingencies of reinforcement for control by multiple dimensions. To allow more sensitive tests of the use of different stimulus dimensions, the number of relevant dimensions was increased from two to three. During initial training, the pigeons could discriminate individual stimuli by using a single stimulus dimension, but different dimensions were relevant to different stimuli. This procedure allowed an independent assessment of the salience of different stimulus dimensions. An adaptive procedure was used so that if one dimension seemed to be acquiring more (or less) control over behavior than the others, the difference between its positive and negative stimulus values could be decreased (or increased). Then in subsequent training we used stimulus sets in which all stimuli could be discriminated using only a single dimension, to see whether control by multiple dimensions would be stable when it was no longer enforced. Finally we introduced a training procedure using polymorphous stimulus sets, to examine whether multidimensional control could be further increased by making it necessary for perfect discrimination

In all phases of the experiment, we presented multiple positive and negative stimuli simultaneously, with different cues available in each. The procedure was similar to that used by Wills et al (2009, Experiment 2b), Nicholls et al. (2011), Lea, De Filippo, Dakin and Meier (2013) and Lea, Poser-Richet and Meier (2015); see also Huber, Apfalter, Steurer, & Prossinger, (2005). In this procedure, discrimination is demonstrated if the subject responds by pecking at the positives rather than the negatives. In the initial training phase of the present experiment, each positive stimulus contained the positive value of only a single stimulus dimension, and each negative stimulus contained the negative value of only a single dimension. The salience of the stimulus differences on the dimensions could therefore be assessed by recording (a) which kinds of positive stimuli were pecked first and (b) which stimulus dimension was discriminated most accurately. If the stimulus difference on one of the dimensions appeared to be more salient than the others according to either of these criteria, the difference between its positive and negative values was reduced, and conversely if the difference between the stimuli on one dimension appeared to be less salient than the others, the difference was increased.

After successful training with only a single relevant dimension in each stimulus, the pigeons were transferred to stimulus sets in which all three dimensions were relevant, and all provided valid information, as in Wills et al. (2009). As in that experiment, following successful discrimination training, tests were carried out with "one-away" stimuli, in which one of the three stimulus dimensions was put into conflict with the other two (so that, for example, it took its negative value in a stimulus where the other two dimensions took positive values). This allows an assessment of the extent to which the pigeons' behavior is under the control of a particular one of the three stimulus dimensions, or whether behavior is controlled roughly equally by all dimensions, which is referred to as "Overall Similarity" control.

Finally, the pigeons were transferred to a 2-out-of-3 artificial polymorphous category task, formally similar to that used by Lea and Harrison (1978). In this task, all the stimuli are "one-aways", in which one of the stimulus dimensions is in conflict with the other two. Perfect performance in such a task requires discrimination according to Overall Similarity. While birds can be trained in such tasks (e.g. Lea and Harrison, 1978; Lea et al., 1993; Jitsumori, 1993), learning in these experiments has typically been slow, and frequently not all the stimulus dimensions involved acquire control over behavior, so that performance is never perfect (e.g. Lea et al, 1993; Lea, et al., 2006). In the present experiment, we were interested to see whether previous training on each of the stimulus dimensions separately, and careful balancing of the saliences of the stimulus differences on them, would allow more rapid and complete acquisition of a polymorphous category discrimination than is usually found. We also investigated whether the polymorphous condition would increase control by multiple dimensions, which is required for perfect discrimination under that condition.

Method

Subjects. Eight pigeons were acquired as discards from local fanciers. They had no previous experimental experience. They were maintained in the same way as the subjects in Experiment 1.

Apparatus. Four test chambers were used, identical to those used in Experiment 1 except that they were fitted with infra-red touchscreens. As before, 30 pixels on the screen corresponded to 1 cm. The control apparatus was the same as in Experiment 1.

Stimulus materials. The stimuli were circular patches, 80 pixels (2.4 cm) in diameter, consisting of sinusoidal gratings that varied in three dimensions. In each dimension, two extreme values and a neutral value were defined, as follows; additional intermediate values were also used.

Hue. Hue was varied from Red to Orange. The blue pixel value was set to zero in all stimuli. The strongest Red hue was produced by setting the pixel values used to red 255 (the maximum value) and green 0. The strongest Orange hue was produced by setting them to red 180 and green 75. A neutral hue was produced by setting red to 218 and green to 37. Intermediate hues were produced by increasing the green value and reducing the red value, while maintaining their total at 255. Because the stimuli were gratings, intensity varied across them according to a sinusoidal function; at minimum intensity, the pixel values were zero for both red and green, and at intermediate intensities, the proportions of the red and green values were always the same as for the maximum intensity for the stimulus concerned.

Orientation. The orientation of the grating was varied from 0° (Horizontal) to 90° (Vertical). The neutral orientation was 45°. Only orientations in the positive quadrant were used, i.e. between 0° and 90° anti-clockwise from horizontal.

Spatial frequency. The spatial frequency of the grating was varied from 7 (Low) to 17 (High) cycles across the diameter of the stimulus, i.e. from .083 cycles/pixel to .213

cycles/pixel. The neutral value of spatial frequency was defined as 11 cycles across the diameter of the stimulus (.138 cycles/pixel), the approximate geometric mean of the extreme values.

Different combinations of these stimulus values were used for different pigeons at different stages of the experiment, as detailed below.

Procedure.

Pretraining. The pigeons were trained by conventional means to peck an Observing Key, a circular white patch of diameter 80 pixels (2.4cm) located centrally on the touch screen and with its centre 238 pixels (8 cm) above the base of the screen; and then to peck either of two Side Keys, each consisting of a circular white patch of diameter 80 pixels centred 60 pixels (3 cm) from one side of the touch screen and 100 pixels (3 cm) above its base. Pecking a Side Key always led to a 3-s presentation of the food hopper at the corresponding side of the touch screen; the food hopper recess was illuminated while the hopper was presented. Following completion of pretraining on the Observing and Side Keys, the pigeons were exposed to a procedure in which a single peck on the Observing key led to the presentation of a circular array of 12 white outline circles. The individual circles were 100 pixels in diameter; the ring of circles was 225 pixels (13.4 cm) in diameter. Within each circle, a white hexagon, 80 pixels in diameter, was displayed. If the pigeon made two pecks at the same hexagon, without any intervening pecks at any other member of the array, that hexagon disappeared (but its surrounding circle remained), and the nearer Side Key was illuminated; a single peck on that key led to hopper presentation. The array remained in place until all the hexagons had been removed in this way, after which there was an inter-trial interval of between 1 and 5 s, followed by the presentation of the Observing Key allowing the pigeon to obtain the next array. Sessions consisted of the presentation of four such arrays. For some pigeons, the brightness of the hexagons was reduced in initial sessions. Most of the pigeons required 5 sessions before they were removing all the hexagons from each array rapidly and smoothly, but one bird required 6 and another 18. Once the pigeons were performing smoothly with 12 hexagons, they were given a single session in which only 6 of the circles in each array contained hexagons. Some pigeons did not transfer smoothly from pretraining to the discrimination training that followed: these pigeons were given a small amount of additional pretraining using only the positive stimuli from the discrimination procedure, in some cases at reduced saturation. Data from these sessions are not included in the analyses that follow.

Discrimination training: general. The experiment involved discrimination training and test sessions using several different sets of stimuli, but the general procedure was the same for all of them. As in the pretraining with the hexagon stimuli, it involved the presentation of a series of circular arrays of 12 stimuli, with each stimulus contained within a white circle. However, in discrimination training, some of the stimuli were positive and others were negative, so that the pigeon's task was to identify and peck the positive stimuli while ignoring the negative stimuli. Positive and negative stimuli were arranged in pseudo-random order around the array, with the constraint that no more than three stimuli of one category could occur without an intervening stimulus of the other category.

Procedure within trials was the same as in Wills et al. (2009, Experiment 2b), Nicholls et al. (2011) and Lea et al. (2013, 2015). As in pretraining, two successive pecks to a positive stimulus led to its disappearance and the presentation of a side key and access to food reward. Two successive pecks to a negative stimulus led to its disappearance, but no side key was presented; instead, the touchscreen became inactive for a time equal to that of the hopper presentation. Once all six positive stimuli had been removed, the array disappeared, and the inter-trial interval began. Discrimination training sessions consisted of the presentation of 8 arrays. Performance on each array was measured using the rho statistic of Herrnstein, Loveland, and Cable (1976), which compares ranks assigned to positive and negative stimuli on the basis of the subject's tendency to respond to them. The first stimulus pecked was assigned the highest rank, the second stimulus peck the next highest, and so on. Negative stimuli that were removed unpecked when all the positive stimuli in an array had been pecked were all assigned the mean of the remaining, unassigned ranks. For each session, overall discrimination performance was assessed by the mean value of rho across the eight arrays.

Discrimination training: dimensions in parallel. The first stage of discrimination training involved stimulus sets similar to those used in the "Features in parallel" procedure of Lea et al. (2006). An example of the training arrays used is shown in Figure 4. In each array there were six positive and six negative stimuli. In all stimuli, a single stimulus dimension took either its positive or its negative value, and the remaining two dimensions took their neutral value. Two of the positive stimuli contained a positive value of Hue, two contained a positive value of Orientation, and two contained a positive value of Spatial Frequency; and similarly two of the negative stimuli contained the negative value of Hue, two the negative value of Orientation, and two the negative value of Spatial Frequency. In the example shown in Figure 4, the positive values of Hue, Orientation and Spatial Frequency are Red, Horizontal and Low respectively; the positive values of the three dimensions were counterbalanced across the eight pigeons.

Figure 4 about here

To assess performance on each stimulus dimension, a rho value was calculated using only the four stimuli containing positive or negative values of that dimension. To assess whether the pigeons tended to peck the stimuli with positive values on specific dimensions more readily than those with positive values on the other dimensions, the mean ranks assigned to the positive stimuli from the three stimulus dimensions were calculated for each array, and the consistency of differences between these means across the arrays in a session was assessed using Kendall's W coefficient of concordance.

Stimulus difference adjustment. Initially, the extreme values of each stimulus dimension were used in all stimuli. However, if a pigeon showed a sustained or substantial tendency to discriminate one dimension better or worse than the other two, or to peck the positive stimuli in which one dimension was relevant consistently earlier or later in an array than those in which the other two dimensions were relevant, the stimulus values were adjusted for the next training session, using criteria defined below. Hue adjustments were made by changing the pixel levels of red and green by 9, maintaining their total at 255. Angle adjustments were made by changing the angles by 10°. Frequency adjustments were made by changing the spatial frequency by 1 cycle across the diameter of the stimulus (2 cycles if the frequency was above 12 cycles across the stimulus diameter). Initial adjustments were necessarily in the direction of reducing the stimulus difference on one or two of the dimensions; subsequent adjustments were made in the direction of increasing a stimulus difference if possible. The criteria for making an adjustment were developed in the light of experience but remained subjective, since the two factors of equal discrimination and equal salience of the three dimensions had to be balanced, while maintaining high levels of discrimination. The aim was to continue training and adjustment until rho values for discrimination performance on all three dimensions exceeded 0.8 in a session, the W value for consistency of the order of pecking the three different kinds of positive stimulus fell to 0.20 or below, and there was no dimension for which the stimuli in which it was positive tended to be pecked either before, or after, both of the others. For one pigeon (Mo) the attempt to achieve all these criteria seemed to be leading to increasing instability, so adjustment and training were terminated when the first criterion (rho values on all dimensions of at least 0.80) was met. In all subsequent training and testing, the positive and negative values of each stimulus dimension used were those in use at the end of the stimulus adjustment procedure.

Prototype training. The pigeons were then exposed to stimuli in which all three dimensions took their positive values, or all three took their negative values. Sessions consisted of 8 arrays each containing six identical positive stimuli and six identical negative stimuli. Two such sessions were given.

One-away tests. Test sessions followed. Each test session consisted of 7 arrays. Odd numbered arrays included only prototype stimuli, as in the immediately previous training, six positive and six negative. Even-numbered arrays (test arrays) included four positive and four negative prototype stimuli, plus four test stimuli. In two of the test stimuli, two of the stimulus dimensions took their positive values and the remaining one (the anomalous dimension) took its negative value; in the other two test stimuli, two dimensions took their negative values and one took its positive value. To ensure that there were equal numbers of stimuli associated with positive and negative contingencies in each array, one of each kind of test stimulus was associated with the contingencies appropriate to a positive stimulus, and one with the contingencies appropriate to a negative stimulus. The anomalous dimension was the same for all the stimuli in a given array, and was different in each of the three test arrays in a session. Three test sessions were given to each pigeon, separated by a single prototype training session; an additional prototype training session was given after the final test session. The order in which the dimensions were used as anomalous was varied across pigeons and counterbalanced between the three test sessions for a given pigeon.

Polymorphous training. Following completion of the one-away tests, the pigeons were given further training sessions, with 8 arrays per session, in which all the stimuli included one anomalous dimension. Stimuli were associated with positive reinforcement contingencies if two dimensions took their positive values, and with negative contingencies if two dimensions took their negative values. There were 6 positive and 6 negative stimuli in each array, with each stimulus dimension taking its anomalous value (negative in a positive stimulus, or positive in a negative stimulus) in two of the stimuli of each category. 10 sessions of this type were given.

Results

Training with dimensions in parallel and stimulus difference adjustment. Table 1 includes a summary of performance in the initial training and stimulus adjustment procedures The rho values for each dimension show that all pigeons came under control of all three dimensions in this procedure. However, as the final stimulus values show, five of the eight pigeons required some adjustment of the stimulus values on one of the stimulus dimensions to ensure roughly equal control of behavior by all three of them. At the end of training, no pigeon was requiring adjustment on two dimensions, though some experienced adjustment on more than one dimension during the course of training.

Prototype training. Table 1 also summarises performance in the two sessions of prototype training that followed initial discrimination training. It can be seen that all pigeons showed effectively perfect performance under this condition.

Table 1 about here

One-away tests. Table 2 summarises performance in the one-away test sessions, using the same form of analysis as Wills et al (2009, Tables 1 & 4). That is to say, a stimulus is counted as being treated as positive if it was one of the first 6 that the pigeon removed, out of the 12 available in an array, and the table records the number of test stimuli whose positive or negative categorization is correctly predicted by each of the three dimensions on its own, or by the Overall Similarity rule (treat a stimulus as positive if it has a majority of positive features). It can be seen that for seven of the eight pigeons, the overall similarity rule made the most successful predictions, and for the remaining pigeon it was jointly most successful.

However for at least three of the pigeons (Bw, Hy and Ax), one of the three dimensions individually did notably better than the others.

Table 2 about here

Polymorphous training. Figure 5 summarises performance in the final polymorphous concept training. As would be expected from their performance in one-away tests, all pigeons showed above-chance performance from the beginning of training (a significant proportion: binomial test, $P<.01$ two-tailed). Mean performance improved across the sessions (Spearman's rho between mean rho and session number = 0.86, $N=10$, $P<.01$ two-tailed). Table 3 shows the level of discrimination of each stimulus dimension during the first and second five sessions of polymorphous concept training. In a polymorphous categorization, the positive value of each stimulus dimension occurs in some negative stimuli, and the negative value of each dimension occurs in some positive stimuli. As a result, fully accurate performance on the overall discrimination can only occur if the mean rho value with reference to each individual dimension is 0.67; higher single-dimension rho values than this reflect excessive control by one dimension, which must be at the expense of one or both of the other dimensions. Values of rho for a single dimension of 0.5 or below reflect failure to come under control of a dimension. The table also shows the standard deviation of the mean rho values for the three dimensions in each session block. A lower standard deviation indicates more equal control by the three dimensions. It can be seen from the table that, by the second session block, no mean rho values lower than 0.5 were recorded for any dimension. The standard deviation of mean rho values was more than halved between the first and second blocks of training, a significant reduction (Wilcoxon test, $T=1$, $N=8$, $P<.02$ 2-tailed).

Figure 5 and Table 3 about here

Discussion

Experiment 2 demonstrated, in agreement with Experiment 1, that when differences in dimensional perceptual salience are made relatively small, pigeons' behavior will come under the control of more than one stimulus dimension simultaneously and to a roughly equal extent. The experiment extended the number of dimensions for which we could demonstrate this from two to three.

The later sessions of polymorphous training demonstrate some tendency for the contingencies to overcome differences in salience, in a way that does not always happen when birds are trained under polymorphous contingencies from the beginning, or even with training on individual dimensions (see, for example, Lea et al., 2006). But even that process was incomplete, with new instances of excessive control by a single dimension opening up even in the later sessions of polymorphous training (e.g. pigeon Ta on orientation).

General Discussion

The current experiments throw further light on the circumstances under which discrimination in terms of multiple stimulus dimensions will be obtained reliably in pigeons. It was already known that this can, but does not always, happen. Evidence for control by all available dimensions has been reported when it is required for accurate performance (e.g. Lea & Harrison, 1978; Teng et al., 2014; von Fersen & Lea, 1990; Vyazovska et al., 2015). It is unlikely when a single dimension is sufficient for discrimination (e.g. Reynolds, 1961; Wills et al., 2009). However, the need to use all dimensions to obtain 100% accuracy is not a sufficient condition to ensure that they will all be used (e.g. Lea et al., 1993), at least unless special training is provided (e.g. von Fersen & Lea, 1990). Nor is it a necessary condition, as the results of Smith et al. (2011) show.

The present Experiment 1 suggests that stimulus set structure is unlikely to be a key factor in biasing pigeons towards the use of multiple stimulus dimensions, contrary to what is known from human categorization and learning (e.g., Milton et al., 2008; Pothos & Close, 2008). However both experiments suggest that a key factor is controlling the saliences of the stimulus differences on the different dimensions. A lack of control over salience may thus explain why inconsistent results have been obtained in the past, with some experiments showing unidimensional control and others showing multidimensional control. In particular, we used relatively small color differences, and the fact that multidimensional control emerged reliably suggests that the large color differences typically used in experiments with pigeons may be so salient that they block control over behavior by other dimensions. For example, Wills et al. (2009) found that 5 out of 6 of their pigeons categorized multidimensional stimuli in terms of color alone. It should be noted, though, that failures of control by all dimensions can occur with purely monochromatic stimuli (e.g. Lea et al., 1993).

Theoretical accounts of human categorization do not fare well in accounting for the results of these experiments. In Experiment 1, if the pigeons had used rules, in the way that Ashby et al. (1998) suggest humans do, they would have consistently used single dimensions; none of them did, in either experiment. If they had categorised according to either Pothos and Chater's (2002) or Rosch and Mervis's (1975) principles, they should have shown no bias between unidimensional and multidimensional classification; they all tended to use multidimensional classification. The work on human categorization processes basically reveals a flexible categorization mechanism, which can accommodate different biases or approaches, depending on the structure of the environment (assuming stimulus dimensions of broadly equivalent salience). Do non-human animals lack this flexibility? The present results suggest that they might, though they are far from being conclusive on this point. Note that formal categorization models (including the ones above) typically make predictions on the basis of the similarity structure of the stimuli (that is, their relative similarities to each other) and are therefore sensitive to issues such as the scaling of the dimensions; despite our controls, it is possible that our assumptions about those were incorrect. In addition, in designing the two category structures, we also considered whether the stimuli could relate to each other in a way that would make it easier to categorize in terms of an overall synthetic dimension or not, but no evidence for this possibility emerged (e.g., Experiment 1).

At first glance, this seems to be what Ashby et al. (1998) would expect of an animal that only has associative learning mechanisms available, and indeed what Smith et al. (2011) observed in a direct test of the differential predictions of the Ashby approach for humans and pigeons. Both Chase (1968) and Blough (1972) have advanced theories about the determination of pigeons' response to multidimensional stimuli, which essentially rely on associative processes alone. But the present results are not entirely comfortable from the point of view of such theories. Although we succeeded in getting balanced control of pigeons' behavior from two or three stimulus dimensions, the process seemed to be easily thrown out of balance: These results suggest that a more sophisticated theory of animal attention may be required. We do not suggest that pigeons formulate rules, and they clearly do learn about multiple stimulus dimensions provided salience differences are controlled. But it does appear that strong control by a single dimension may in some sense shut down control by other available dimensions, so that response is not determined by a static combination of information available across the available dimensions, but by a more dynamic process. Such processes have of course been actively considered in theories of animal attention, from Krechevsky (1932) and Sutherland and Mackintosh (1971) onwards; it may be time to bring them back into play.

The study of categorization processes with non-humans, using models and paradigms developed for human categorization, is not without problems, and our work reveals several cautionary points. First, the biases to categorize in different ways in humans are more clear cut when there are many stimuli. In the case of the computational models considered above, it is easier to create categories which reflect stronger unidimensional vs. multidimensional biases, when there are more stimuli; equally, we would have more confidence in equating unidimensional vs. multidimensional biases, with more stimuli (with few stimuli, changes in the representation of even one stimulus can have a large impact on model predictions). In the case of Ashby et al.'s (1998) work, most of their experiments involve large numbers of stimuli. By contrast, with few stimuli, categorization could be driven by just exemplar memorization. Of course, the use of categorization paradigms with a large number of stimuli with non-humans is not straightforward (e.g., because of the additional controls, which must be in place, or issues with learning competencies). Second, closer scrutiny is perhaps warranted regarding the assumptions for how non-humans represent the relevant stimuli. For humans, it usually suffices to assume a reasonably well-behaved psychological space, but if analogous assumptions do not translate to non-humans, such as pigeons, then this casts doubt on the validity of any predictions from corresponding models. Third, it has to be noted that even in the adult literature there is some disagreement regarding the exact interplay between unidimensional and multidimensional biases in classification (e.g., Murphy, 2004; Pothos & Close, 2008; Milton et al., 2008). So, it could be the case that eventually different perspectives will emerge for this debate in the human literature, with different implications for the non-human learning literature.

Notwithstanding these qualifications, the issue of unidimensional vs. multidimensional biases in non-human learning is clearly a significant one. To apply ideas and theory from the corresponding human categorization literature, as we have attempted in the present work, provides new, exciting perspectives for theoretical understanding. We hope the present work will aid future, corresponding efforts.

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.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning Memory and Cognition*, 14, 33-53. doi:10.1037/0278-7393.14.1.33

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*, 147-161. doi:10.1111/j.1749-6632.2010.05874.x

Atick, J. J. (1992). Could information theory provide an ecological theory of sensory processing? *Network*, *3*, 213-251.

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

Butter, C. M. (1963). Stimulus generalization along one and two dimensions in pigeons. *Journal of Experimental Psychology*, *65*, 339-346.

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

Chase, S. (1968). Selectivity in multidimensional stimulus control. *Journal of Comparative and Physiological Psychology*, *66*, 787-792.

Chase, S., & Heinemann, E. G. (1972). Choices based on redundant information: An analysis of two-dimensional stimulus control. *Journal of Experimental Psychology*, *92*, 161-175.

Couchman, J. J., Coutinho, M. V. C., & Smith, J. D. (2010). Rules and resemblance: Their changing balance in the category learning of humans (*Homo sapiens*) and monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, *36*, 172-183.

Farthing, G. W., & Hearst, E. (1970). Attention in the pigeon: Testing with compounds or elements. *Learning and Motivation, 1,* 65-78.

Herbranson, W. T., Fremouw, T., & Shimp, C. P. (1999). The randomization procedure in the study of categorization of multidimensional stimuli by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 25*, 113-134.

Herbranson, W. T., Fremouw, T., & Shimp, C. P. (2002). Categorizing a moving target in terms of its speed, direction, or both. *Journal of the Experimental Analysis of Behavior, 78,* 249-270.

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., 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.

Jitsumori, M. (1993). Category discrimination of artificial polymorphous stimuli based on feature learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *19*, 244-254.

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., De Filippo, G., Dakin, R., & Meier, C. (2013). Pigeons use low rather than high spatial frequency information to make visual category discriminations. *Journal of Experimental Psychology: Animal Behavior Processes, 39*, 377-382.

Lea, S. E. G., & Harrison, S. N. (1978). Discrimination of polymorphous stimulus sets by pigeons. *Quarterly Journal of Experimental Psychology, 30*, 521-537.

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, 46B,* 19-42.

Lea, S. E. G., Poser-Richet, V., & Meier, C. (2015). Pigeons can learn to make visual category discriminations using either low or high spatial frequency information. *Behavioural Processes, 112,* 81-87.

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: II. Strategic information search in humans (*Homo sapiens*) but not in pigeons (*Columba livia*). *Journal of Comparative Psychology, 123,* 406-420.

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.

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.

Murphy, G. L. (2004). The big book of concepts. MIT Press: Cambridge, USA.

Newman, F. L., & Baron, M. R. (1965). Stimulus generalization along the dimension of angularity: A comparison of training procedures. *Journal of Comparative and Physiological Psychology, 60,* 59-63.

Nicholls, E., Ryan, C. M. E., Bryant, C. M. L., & Lea, S. E. G. (2011). Labelling and family resemblance in the discrimination of polymorphous categories by pigeons. *Animal Cognition, 14,* 21-34.

Pothos, E. M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science, 26*, 303-343.

Pothos, E. M. & Close, J. (2008). One or two dimensions in spontaneous classification: A simplicity approach. Cognition, 107, 581-602.

Pothos, E. M., Edwards, D. J., & Perlman, A. (2011a). Supervised vs. unsupervised categorization: Two sides of the same coin? *Quarterly Journal of Experimental Psychology, 64*, 1692-1713.

Pothos, E. M., Perlman, A., Bailey, T. M., Kurtz, K., Edwards, D. J., Hines, P., & McDonnell, J. V. (2011b). Measuring category intuitiveness in unconstrained categorization tasks. *Cognition, 121*, 83-100.

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

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

Smith, J. D., Beran, M. J., Crossley, M. J., Bloomer, J., & Ashby, F. G (2010). Implicit and explicit category learning by macaques (Macaca mulatta) and humans (Homo sapiens). *Journal of Experimental Psychology: Animal Behavior Processes, 36,* 54-65. doi:10.1037/a0015892

Smith, J. D., Crossley, M. J., Boomer, J., Church, B. A., Beran, M. J., & Ashby, F. G. (2012). Implicit and explicit category learning by capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology, 126,* 294-304. doi:10.1037/a0026031

Soto, F. A., & Wasserman, E. A. (2010). Integrality/separability of stimulus dimensions and multidimensional generalization in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 36*, 194-205.

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

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

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

Weatherburn, C. E. (1957). *A first course in mathematical statistics*, 2nd Edn. Cambridge: Cambridge University Press.

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

Wright, A. A., & Cumming, W. W. (1971). Color-naming functions for the pigeon. *Journal of the Experimental Analysis of Behavior*, *15*, 7-17.

Table 1

Experiment 2: Performance during initial acquisition, stimulus adjustment, and prototype training. Stimulus dimensions are denoted by H (hue), O (orientation) and F (spatial frequency); units are pixel values of green and red color systems for hue, degrees anticlockwise from horizontal for orientation, and cycles per stimulus diameter for spatial frequency. Where the final stimulus values had required adjustment away from the original values, they are shown in bold. Final rho values for individual dimensions are means over the final two sessions, and rho values for prototype training are means over the two sessions given before one-away tests began.

		Final stimulus values (positive; negative)			Final rho values				
Pigeon	Sessions to first criterion performance	H	O	F	H	O	F	Total training and stimulus adjustment sessions	Rho in prototype training
At	6	g000r255; g075r180	**80;** **10**	07; 17	.94	.88	.90	24	1.00
Ax	3	g075r180; g000r255	00; 90	**09;** **13**	.88	1.00	.81	15	.99
Bw	6	g075r180; g000r255	90; 00	17; 07	.97	.91	.84	7	1.00
Fz	4	g075r180; g000r255	00; 90	17; 07	.97	1.00	.97	18	.99

Hy	3	g075r180; g000r255	**80;** **10**	07; 17	.88	1.00	.97	8	.99
Lp	6	g000r255; g075r180	00; 90	07; 17	.89	1.00	.89	21	1.00
Mo	8	g000r255; g075r180	90; 00	**15;** **08**	1.00	.94	.89	26	1.00
Ta	5	**g009r246;** **g066r189**	00; 90	07; 17	.84	.94	.92	10	.96

Note. All pigeons were initially trained using extreme values of all three dimensions (Hue: g000r255 vs g075r180; Orientation 90 vs 0; Frequency 7 vs 17)

Table 2

Experiment 2: Rules predicting responses to test stimuli in "One-away" tests. Entries in the main body of the table are the numbers of test stimuli whose categorization by the pigeons as positive or negative was correctly predicted by each of the three stimulus dimensions on its own, or by the Overall Similarity rule (treat a stimulus as positive if it has a majority of positive features). The right hand column shows the rule that best predicted categorization, as implied by the largest of the figures in the previous four column.

Pigeon	Categorization rule				Inferred strategy
	Hue	Orientation	Spatial frequency	Overall Similarity	
At	23	25	19	31	OS
Ax	19	29	21	31	OS
Bw	27	19	17	27	OS/H
Fz	22	26	22	28	OS
Hy	21	21	27	33	OS
Lp	22	26	22	34	OS
Mo	24	20	24	32	OS
Ta	23	19	25	31	OS

Table 3. Experiment 2: accuracy of discrimination of each stimulus dimension during polymorphous training. Entries are mean values across each session block of the rho statistic of Herrnstein et al. (1976). Because of the polymorphous structure of the stimulus sets, 100% accurate performance would lead to mean rho values of 0.67 for all three dimensions.

	Hue		Orientation		Spatial frequency		Standard deviation of mean rho values	
Pigeon	Sessions 1-5	Sessions 6-10	Sessions 1-5	Sessions 6-10	Sessions 1-5	Sessions 6-10	Sessions 1-5	Sessions 6-10
At	0.60	0.54	0.64	0.66	0.53	0.60	0.056	0.035
Ax	0.57	0.57	0.70	0.60	0.43	0.56	0.135	0.006
Bw	0.62	0.61	0.53	0.63	0.59	0.63	0.046	0.012
Fz	0.51	0.55	0.46	0.66	0.81	0.78	0.189	0.133
Hy	0.44	0.53	0.70	0.67	0.65	0.66	0.138	0.075
Lp	0.65	0.62	0.65	0.63	0.48	0.61	0.098	0.006
Mo	0.66	0.61	0.57	0.68	0.53	0.56	0.067	0.029
Ta	0.60	0.60	0.61	0.70	0.65	0.66	0.026	0.035
Mean	0.58	0.58	0.61	0.58	0.58	0.63	0.094	0.041

Figure captions

Figure 1. Experiment 1: Examples of the training stimulus sets, showing sets with (a) monotonically related and (b) non-monotonically related stimulus dimensions

Figure 2. Experiment 1: Types of stimuli used in transfer tests. Panel (a) shows an example for a bird trained with monotonically related stimulus dimensions, panel (b) for a bird trained with non-monotonically related dimensions. The stimulus types are coded as follows: A and B, training stimuli; TN, novel combinations of stimulus dimension values that preserved the valence combinations used in training; ?, the ambiguous stimulus for which neither dimension gives any information; TC, T1-T4, conflict stimuli, in which the two dimensions gave conflicting information. In stimuli T1 and T4, Dimension 1 would be expected to control behaviour more strongly than Dimension 2, and vice versa for stimuli T2 and T3.

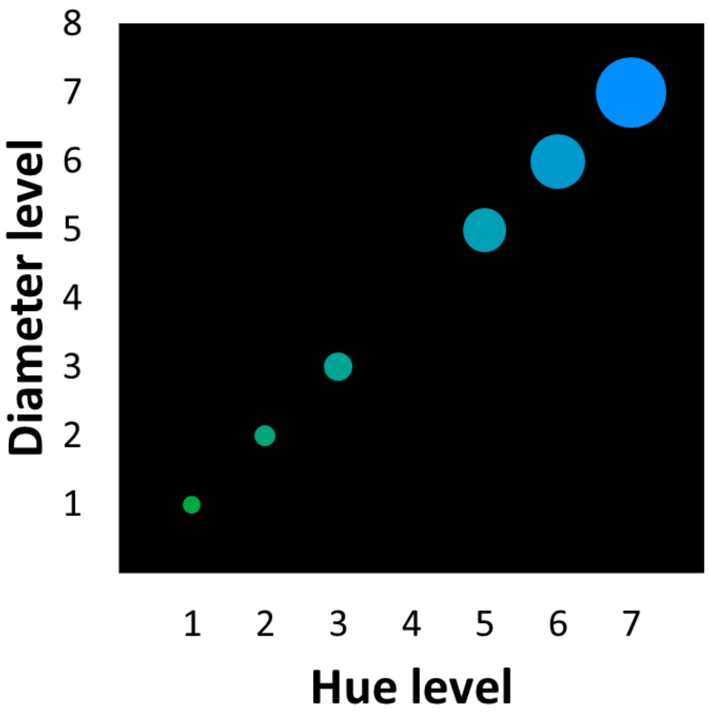
Figure 3. Experiment 1: Mean acquisition performance.

Figure 4. Experiment 2: Example of a stimulus array from the discrimination training phase. The + signs indicate which stimuli were positive, but were not shown to the pigeons. In the example shown, the positive values of the three stimulus dimensions were Red (Hue), Horizontal (Orientation) and Low (Spatial Frequency).

Figure 5. Experiment 2: Performance during polymorphous concept training. All birds received 10 sessions of this training. Data are mean values of the rho statistic across eight birds, with error bars showing the range. Note that chance performance would correspond to a mean rho of 0.5

Figure 1. Experiment 1: Examples of the training stimulus sets, showing sets with (a) monotonically related and (b) non-monotonically related stimulus dimensions. Each cluster of three stimuli constituted a category as used in the experiment.

(a)



(b)

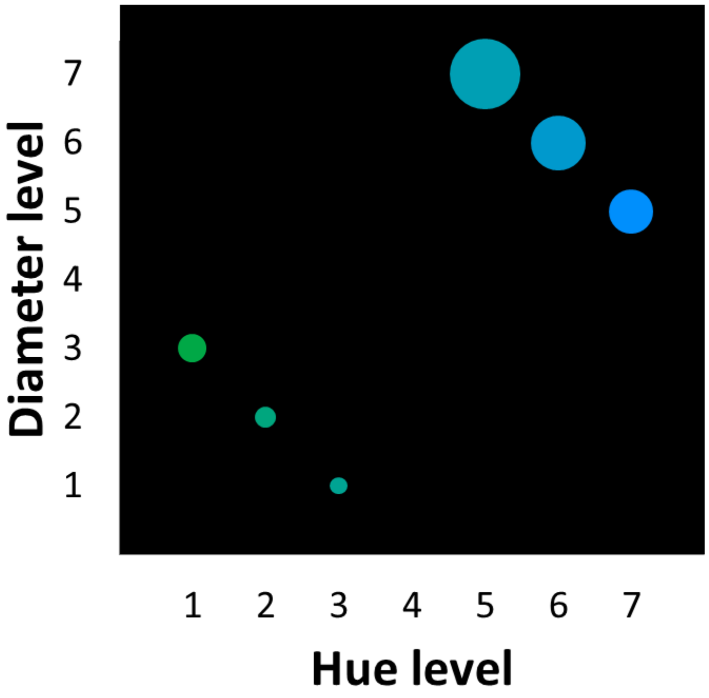
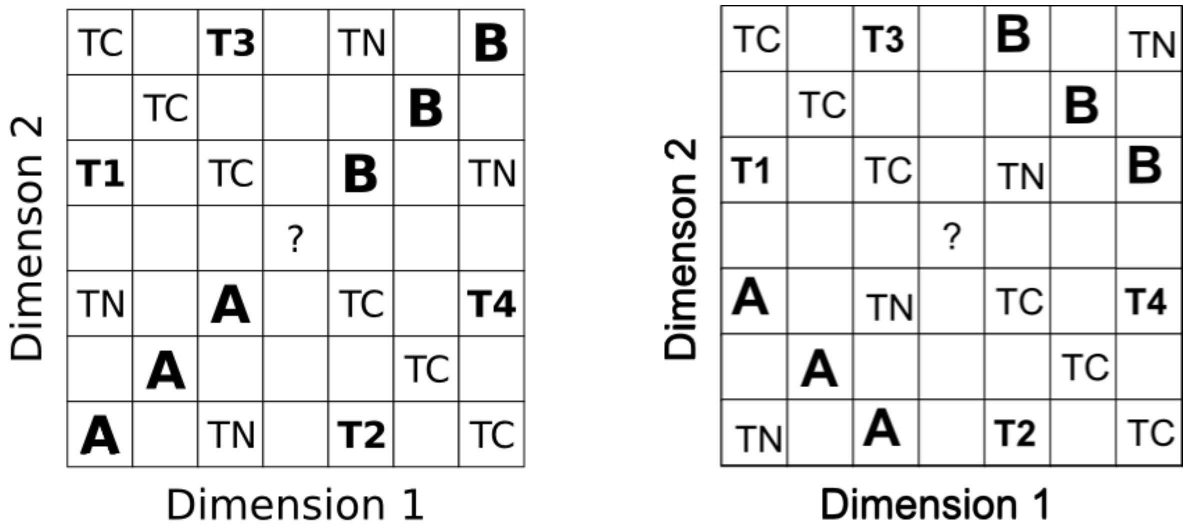


Figure 2. Experiment 1: Types of stimuli used in transfer tests. Panel (a) shows an example for a bird trained with monotonically related stimulus dimensions, panel (b) for a bird trained with non-monotonically related dimensions. The stimulus types are coded as follows: A and B, training stimuli; TN, novel combinations of stimulus dimension values that preserved the valence combinations used in training; ?, the ambiguous stimulus for which neither dimension gives any information; TC, T1-T4, conflict stimuli, in which the two dimensions gave conflicting information. In stimuli T1 and T4, Dimension 1 would be expected to control behaviour more strongly than Dimension 2, and vice versa for stimuli T2 and T3.



(a)

TC		**T3**		TN		**B**
	TC				**B**	
T1		TC		**B**		TN
			?			
TN		**A**		TC		**T4**
	A				TC	
A		TN		**T2**		TC

(b)

TC		**T3**		**B**		TN
	TC				**B**	
T1		TC		TN		**B**
			?			
A		TN		TC		**T4**
	A				TC	
TN		**A**		**T2**		TC

Figure 3. Experiment 1: Mean acquisition performance

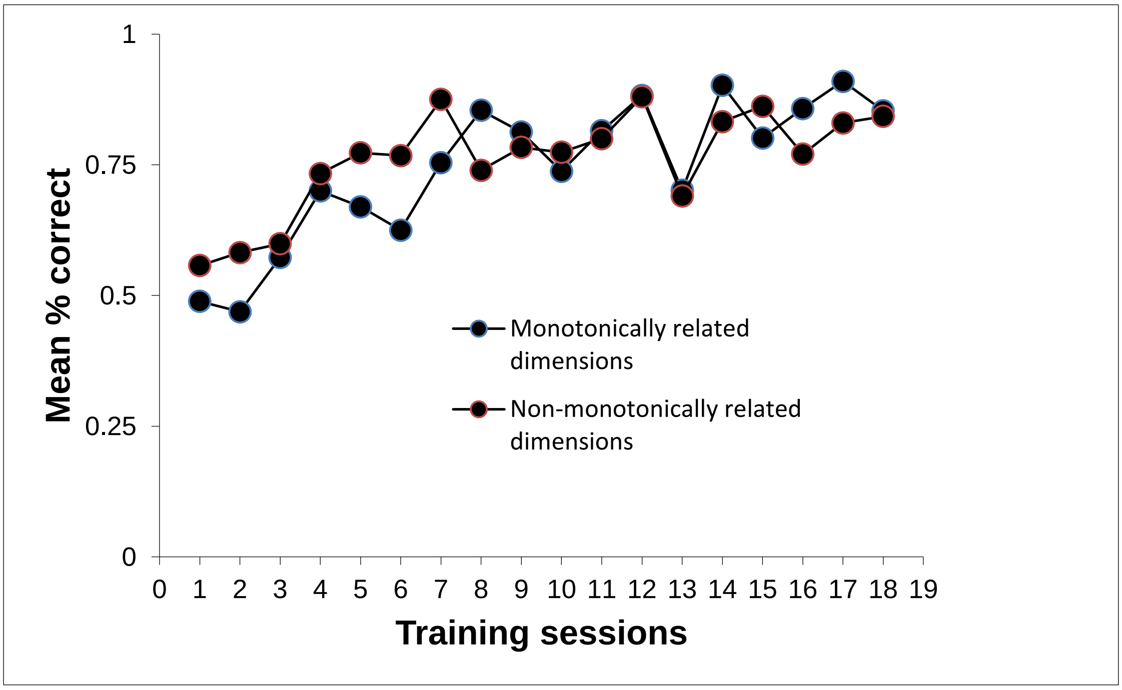


Figure 4. Experiment 2: Example of a stimulus array from the discrimination training phase. In the example shown, the positive values of the three stimulus dimensions were Red (Hue), Horizontal (Orientation) and Low (Spatial Frequency). The labels by each stimulus show which dimension was relevant for it, and whether that dimension took its positive or its negative value (and hence whether the stimulus was positive or negative). The remaining two, irrelevant, dimensions always took their intermediate, neutral values. The labels were not shown to the pigeons.

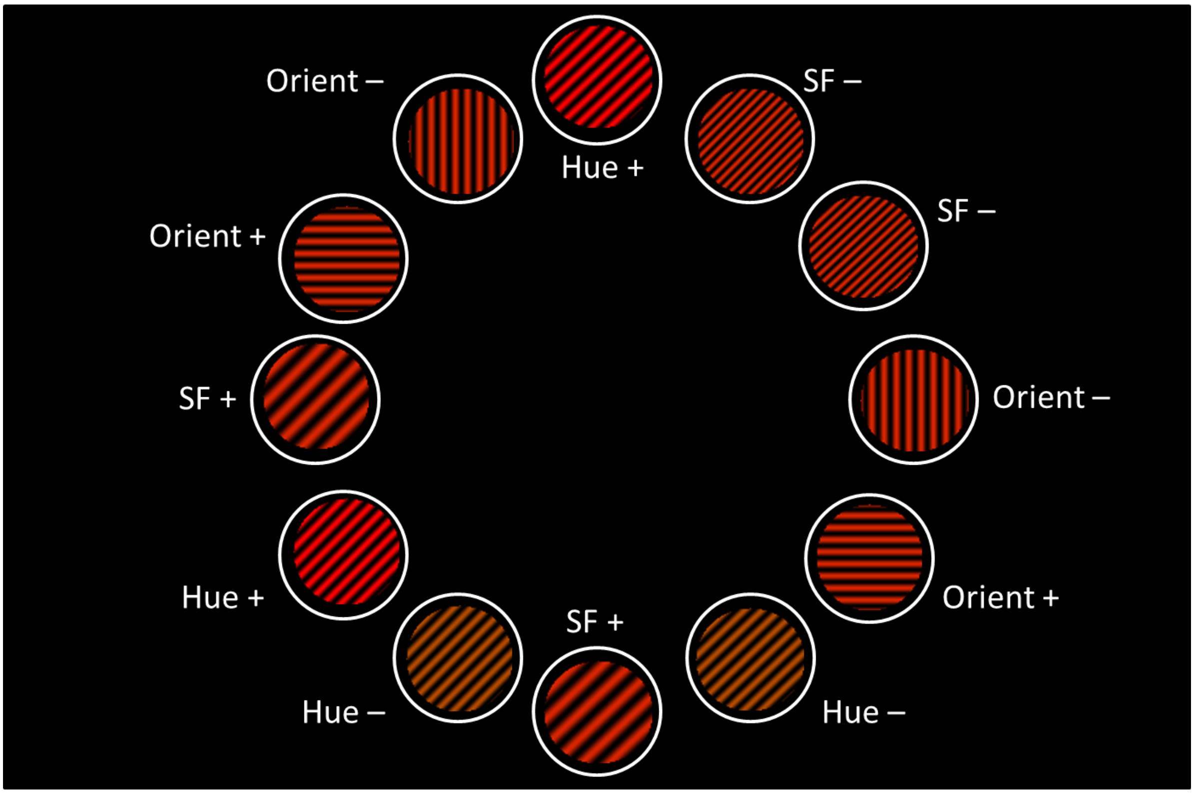


Figure 5. Experiment 2: Performance during polymorphous concept training. All birds received 10 sessions of this training. Data are mean values of the rho statistic across eight birds, with error bars showing the range. Note that chance performance would correspond to a mean rho of 0.5