Interpreting the effects of image manipulation on picture perception in pigeons (*Columba livia*) and humans (*Homo sapiens*)

Kazuhiro Goto, Stephen E. G. Lea, Andy J. Wills and Fraser Milton

University of Exeter

Running head: Spatial frequencies and picture perception

Please address correspondence to:
K. Goto
Kyoto University
Kokoro Research Center
46 Shimoadachi-cho Yoshida, Sakyo-ku
Kyoto, 606-8501
Japan
Phone: +81-(0)75-753-9684
Fax: +81-(0)75-753-9684
Email kgoto@psy.flet.keio.ac.jp

Manuscript prepared for publication in Journal of Comparative Psychology.

This version last saved 14 January 2011

Abstract

The effects of picture manipulations on humans' and pigeons' performance were examined in a go/no-go discrimination of two perceptually similar categories, cat and dog faces. Four types of manipulation were used to modify the images. Mosaicization and scrambling were used to produce degraded versions of the training stimuli, whilst morphing and cell exchange were used to manipulate the relative contribution of positive and negative training stimuli to test stimuli. Mosaicization mainly removes information at high spatial frequencies, whereas scrambling removes information at low spatial frequencies to a greater degree. Morphing leads to complex transformations of the stimuli that are not concentrated at any particular spatial frequency band. Cell exchange preserves high spatial frequency details, but sometimes moves them into the "wrong" stimulus. The four manipulations also introduce high-frequency noise to differing degrees. Responses to test stimuli indicated that high and low spatial frequency information were both sufficient but not necessary to maintain discrimination performance in both species, but there were also species differences in relative sensitivity to higher and lower spatial frequency information.

Keywords: picture perception, spatial frequency, pigeons, humans

Interpreting the effects of image manipulation on picture perception in pigeons (*Columba livia*) and humans (*Homo sapiens*)

One of the cognitive capacities that pigeons and other birds share with humans and other primates is the ability to respond to photographs of natural scenes in a categorical manner. For example, Herrnstein and Loveland (1964) demonstrated that pigeons could discriminate between color slide images that contained at least one human being, and those that contained no human beings. However, it is not yet known whether the perceptual and cognitive mechanisms underlying this capacity are the same in different species.

Evidence from the study of simpler, more artificial visual stimuli suggests that there may be important differences in visual processing between pigeons and humans. Different results are obtained in pigeon and human experiments on, for example, ‘pop-out’ of features in visual search (Allan & Blough, 1989), perception of occluded figures (Sekuler, Lee, & Shettleworth, 1996), geometrical illusions (Nakamura, Fujita, Ushitani, & Miyata, 2006), identification of rotated patterns (Delius & Hollard, 1995), the effect of a configural context on orientation discrimination (Donis & Heinemann, 1993), and global versus local advantage in the processing of hierarchical stimuli (Cavoto & Cook, 2001). Several of the results cited above can be accounted for by a generalization that pigeons and humans differ in terms of the relative importance of configural and elemental information in their pattern recognition of artificial stimuli.

If it is generally true that elemental processing is relatively more important in pigeon than human visual cognition, the same principle should account for any species differences in the perception of complex, natural images as well as artificial stimuli. However, despite considerable experimental effort since the work of Herrnstein and Loveland (1964), the information underlying pigeons' discrimination of photographs has proved difficult to specify (Fetterman, 1996; Huber 2001), partly because of the polymorphous nature of such stimuli (von Fersen & Lea, 1990; Lea, Wills & Ryan, 2006). In general, natural categories such as "human" cannot be described or discriminated by any single simple property, such as wavelength, intensity, or spatial frequency (Honig & Urcuioli, 1981; Lubow, 1974). Nevertheless, if images can be systematically manipulated in ways that eliminate or reduce some aspects of the information in the stimuli, it is possible to examine how different aspects of the stimuli contributed to the discrimination. Early experiments used simple forms of this strategy; for example Herrnstein and Loveland (1964) used monochromatic test stimuli after training with full color images, and found that pigeons' discrimination of humans from non-humans was maintained, showing that the hue components were not essential. More recent experiments have used more sophisticated methods of image manipulation. For example, Huber, Troje and their colleagues (Troje, Huber, Loidolt, Aust, & Fieder, 1999) examined the types of information employed by pigeons in discrimination between male and female human faces. Troje et al. (1999) examined whether pigeons' discrimination was predominately based on "texture", that is the spatial distribution of color across the stimuli, or "shape", that is the components of the stimuli (e.g., the nose). Three groups of pigeons were trained in a successive go/no-go discrimination procedure. The first group was presented with intact faces of males and females from a face database. The second group was presented with pictures whose facial parts were averaged across the entire face database but with the texture of the intact face superimposed (texture-only stimuli). The third group was presented with pictures whose shape was intact but with texture averaged across the entire face database (shape-only stimuli). The pigeons in the intact and texture-only group acquired the discrimination faster and reached higher asymptotic levels of discrimination than those in the shape-only group, suggesting that the shape features were less important than texture cues for this discrimination.

The object of the present experiments was to see whether a single principle could be found that would account for differences between pigeon and human discrimination of a set of complex photographic stimuli. One candidate principle is the relative roles of configural and elemental information, which as noted above seems to differ in pigeon and human discrimination of simple artificial visual stimuli. Cerella (1980) first postulated that pigeons' pattern recognition is largely based on elemental information and only to a small extent on spatial relationship among elements. In an experiment using line drawings, he found that pigeons' discrimination of cartoon characters was unimpaired by transformation such as occlusion, deletion of parts and scrambling of parts. Also using line drawings, Kirkpatrick and her colleagues conducted a more extensive series of experiments that examined the role of spatial organization of elements in complex visual pattern recognition; these experiments demonstrated that the spatial organization of elements does play some part in pigeons' pattern recognition (Kirkpatrick-Steger, Wasserman & Biederman, 1996, 1998). However even if pigeons do use configural information, it remains possible that it is generally less important for them than it is for humans

In addition to this work on line drawings, the relative importance of these two types of cues for pigeons has been studied in relation to full-color picture perception (e.g., Aust & Huber, 2001, 2003). Aust and Huber analyzed the features supporting discrimination of presence and absence of humans in three studies using various picture scrambling methods. Scrambling did not completely disrupt the discrimination though it decreased discrimination accuracy. Aust and Huber's results thus indicate that both elemental and configural features play a part in pigeons' picture perception. However, few studies have directly compared pigeons' and humans' picture perception, and it is unknown whether the relative importance of configural and elemental information in picture discrimination differs between pigeons and humans, as it appears to in discriminations of simple, artificial stimuli.

The distinction made by Aust and Huber (2003) between configural and elemental information in picture stimuli can be conceptualized in three distinct but overlapping ways. The most general is to refer to global versus local properties of the stimuli – those pertaining to the stimulus as a whole and those pertaining to particular regions within it. Overall brightness, for example, is a global cue, whereas the shape of a nose is a local cue. Secondly, one can talk about configural versus elemental features. This terminology is probably most appropriately applied to shape information; for example, the hierarchical stimuli used by Cavoto and Cook (2001) were composed of different-shaped elements that could be arranged in different configurations, which might or might not be the same shape as the elements. Finally, we can give the configural versus elemental distinction a more psychophysical interpretation by framing it in terms of high and low spatial frequencies. Spatial frequency analysis of natural scenes generally shows that they contain roughly equal amounts of energy per octave across the range of detectable spatial frequencies (Field, 1987), but different frequency ranges may have different roles in picture perception for different species. Configural cues, and most global cues, must be specified in terms of the lower spatial frequencies present in a scene, whereas local, elemental cues must be specified in terms of higher spatial frequencies. Ghosh, Lea and Noury (2004) showed that an unexpected difference between human infants and pigeons in behavior towards artificially modified pictures could be explained by identifying the areas of the stimulus that had higher concentrations of high spatial frequencies, and assuming that the pigeons' categorizations were more dependent on these areas than the human infants' categorizations.

In the present study, we examined the roles of elemental and configural information in pigeons' and humans' recognition of perceptually similar complex images, by comparing the effects of four distinct image manipulations. All but cell interchange manipulations have been used in previous experiments on pigeons concept discrimination, but they have not been systematically compared. The objectives were to examine 1) how these manipulations affect the discrimination of natural photographic images, 2) whether their relative impacts differ between pigeons and humans, and 3) whether such differences could be interpreted in terms of the impacts of the manipulations in the spatial frequency domain. Subjects were first trained on a concept discrimination between the unmodified stimuli, and the effects of the image manipulations were assessed in test sessions.

The first two manipulations, mosaicization and scrambling, degraded both positive and negative stimuli but did not introduce information from the positive stimuli into negative stimuli or vice versa. They filter the spatial frequency information in each stimulus in opposite ways, though neither is a pure filtering operation. Both start by dividing the image into tiles. Mosaicization (Harmon, 1973) averages the color within each tile, and thus removes all information whose spatial frequency is higher than the reciprocal of the tile size, while leaving lower spatial frequency information largely intact. On the other hand, scrambling randomizes the location of the tiles, and thus removes all information whose spatial frequency is lower than the reciprocal of the tile size, while conserving most of the higher spatial frequency information. Both manipulations also introduce some high spatial frequency noise at tile boundaries, and it is known that for humans looking at mosaic stimuli, such added information can disrupt recognition of a stimulus (Gordon & Field, 1978). For our stimuli and manipulations, the added frequency content is greater for scrambling than for mosaicization and, in terms of pixel-based image similarity (as measured by cross-correlation analyses), the scrambling manipulation distorted the images more than the mosaicization manipulation. These factors have to be taken into account when interpreting the data.

The other two manipulations, morphing and cell exchange, mix paired positive and negative stimuli in fixed proportions. However, they do so in ways that have different effects in the spatial frequency domain. Morphing involves specifying corresponding anchor points in a pair of stimuli, and deriving intermediate stimuli by taking weighted averages of the properties of the two stimuli (including location) at these points, with interpolation to the remainder of the picture. This leads to complex transformations of the stimuli, but these transformations are not concentrated at any particular spatial frequency band. In cell exchange, we divided the stimuli into tiles, and exchanged a given proportion of these between stimuli. This preserves high spatial frequency details, but sometimes moves them into the “wrong” stimulus. Cell exchange also introduces high spatial frequency noise. In terms of pixel-based image similarity, the cell exchange manipulation distorts the images more than the morphing manipulation. A more detailed analysis of the effects of our four manipulations on the spatial frequency energy spectrum and on the pixel-based similarity of our stimulus sets is provided in the online supplementary materials.

The effects of the first three of these manipulations have been investigated separately in pigeon picture perception studies (e.g., mosaicization: Huber, et al. 2000; scrambling: Aust & Huber, 2001; morphing Ghosh et al., 2004; Makino & Jitsumori, 2001), but cell exchange appears not to have been used before. However, the manipulations have not been compared on the same stimuli. In this study, the four manipulations were applied to the same stimuli and the severity of degradation was systematically varied across several levels to examine how the effect changed according to the manipulation level.

In all the experiments, we used cat and dog faces as stimuli, for two reasons. First, cat and dog faces are perceptually similar categories, and it is difficult to give an instant description of the distinguishing characteristics of these two categories. They are also allospecific in context for both pigeons and humans (though it is unlikely that pigeons would relate the images to real cats or dogs, while humans certainly do). Second, we had already conducted experiments using these stimuli, and confirmed that pigeons show transfer to novel category instances (Ghosh et al., 2004), suggesting that the pictures contained category-diagnostic features.

The first two experiments examined the effects of mosaicization and scrambling in pigeons and humans respectively, and the final two experiments examined the effects of morphing and cell exchange in the two species. As in the experiments of Ghosh et al. (2004), discrimination was established using a go/no-go procedure in pigeons: i.e. the subject was required to respond to one category but not rewarded for making any response to the other. In human experiments, the procedure was made as close as practicable to that used with pigeons.

Experiment 1a: Mosaic and scrambled stimuli in pigeons

Experiment 1a examined the effects of mosaicization and scrambling on pigeons' picture perception. Pigeons first learned to discriminate photographs of cat and dog faces in a successive discrimination task. Generalization tests with mosaicized and scrambled cat and dog faces were then conducted. Generalization tests to novel cat and dog face images were not conducted, because Ghosh et al. (2004), using identical stimuli and similar testing apparatus, had already confirmed that such generalization occurs reliably.

Method

Subjects.

Twelve racing pigeons (*Columba livia*) obtained as discards from local fanciers were used as subjects in this experiment. There were 3 naïve pigeons, 3 pigeons that had been used in an unrelated experiment, and 6 pigeons that had been used in a previous experiment using the same training stimuli but none of the generalization stimuli (Ghosh et al, 2004, Experiment 2). The pigeons were normally housed in two indoor aviaries, measuring 2.2 m by 3.4 m by 2.4 m. Each aviary was equipped with pigeonholes in units of 16, and water and crushed oyster shells were freely available. The pigeons were maintained on a 12:12 hr light/dark cycle, with 30-min simulated dawn and dusk periods. They were moved to individual cages for at least 30 min before and after the test sessions. Each bird was maintained at 85-93% of its free-feeding weight throughout the experiment by the delivery of hemp and conditioner during the experimental sessions and by supplements of mixed grain on non-testing days.

Apparatus.

Four identical three-key operant chambers, 69 cm by 49 cm by 39 cm, were used. Each consisted of a plywood box, with a 3-key intelligence panel (Campden Instruments Ltd, London), 33.5 cm by 35 cm, mounted centrally into the front wall. The three keys had a diameter of 2.5 cm and were centered 10.5 cm apart and 24.0 cm above the grid floor of the chamber. All three keys operated reed switches when struck with a force of 0.035 N. The two side keys were translucent, and could be transilluminated by amber lamps. The center key was transparent, and a shutter operated by a rotary solenoid was situated behind it so that viewing a 15-inch cathode ray tube monitor (HL-5854B, Hyundai), visible 15 cm behind the center key, was prevented during the intertrial intervals. This monitor was controlled by a PC computer (Pentium II 333MHz, Tiny), running a stimulus selection and display program written in Borland Delphi, under the Windows 95 operating system. An aperture in the intelligence panel, 7.0 cm by 7.5 cm, was positioned 15cm below the center key, giving access to a solenoid-operated food hopper attached to the outside of the box; the hopper contained a 1:2 mixture of hemp and conditioner. The availability of this food was signaled by a 1.0-W white light within the hopper tray. General illumination was given by a 3.5-W yellow-lensed house light situated 12 cm above the center key. Masking noise was generated by a ventilation fan and also provided by white noise relayed via a 35-ohm loudspeaker mounted on the back of the intelligence panel. A separate computer was used to generate the experimental stimuli for each test chamber. Both chambers and their stimulus generation computers were housed in a darkened testing room. A further PC-compatible computer (Pentium 133 MHz, Opus Technology PLC) running under the Windows 95 operating system was located outside this room and controlled all experimental events and recorded responses, using a program written in Borland Delphi; a network link enabled this computer to instruct the computers attached to experimental boxes in the testing room to display the stimuli. The pigeons' behavior during experimental sessions could be regularly monitored via video cameras, fitted with a wide-angle lens, mounted on the right wall of each chamber. Each pigeon was assigned to a single test chamber for all stages of the experiment.

Stimuli.

Training stimuli were photographic images of 10 faces of dogs and 10 faces of cats, shown full-face. These stimuli had been used in previous studies in our laboratory (Ghosh et al., 2004); they were originally scanned from magazines or books and their background was removed and converted to plain black. The images were then converted to 8-bit bitmap (256 colors) format. All the images were 192 pixels square on the monitor including any border, and subtended 22.5 degrees of visual angle square at the eye of a pigeon in typical pecking position. There were 200 test stimuli, produced from the 20 training stimuli by manipulating each in two different ways and at five different levels of distortion. For the mosaic manipulation, color was averaged within tiles of size of 4, 8, 12, 16 and 24 pixels square so that fine detail was lost. For the scrambling manipulation, the image was fragmented into equal size square tiles, 48, 32, 24, 16 or 12 pixels square and then the tiles were randomly scrambled. The levels of the two manipulations used were chosen as roughly matched in their effect on discriminability by humans in pilot tests. Samples of these stimuli are shown in Figure 1.

Procedure.

The pigeons that had experienced earlier experiments were not given any pretraining in this experiment. For the other three pigeons, pecks to the center key were autoshaped by repeatedly presenting food shortly after the shutter was opened to show the monitor with a plain white screen. After the acquisition of center key pecking, a fixed interval (FI) schedule was introduced and the value of the FI was gradually raised to 12 s. It was then followed by start key training, in which pecking to the right side key (start key) was shaped when the start key was illuminated to open the center shutter. The start key training was repeated for at least three sessions to stabilize the birds' behavior.

The pigeons were then divided into two groups of 6 each. Half of the pigeons were trained with cats as positive and dogs as negative stimuli (Cat+ group; three pigeons were naive and the other three pigeons had previous experience of an unrelated experiment), and the contingencies were reversed for the other half (Dog+ group; all the birds that had previous experience of the stimuli were in this group since they had received Dog+ training previously). The session started with 3 s of presentation of the food hopper followed by an intertrial interval. Each trial began with a start key peck which opened the shutter to reveal a stimulus on the monitor. Pecks to the center key were counted but not reinforced during the first 10 s of each trial (fixed time; FT). Following the completion of the FT 10 s, pecks were reinforced by a 3 s presentation of the food hopper on a variable interval 4 s schedule on positive stimulus trials; negative stimulus trials ended without reinforcement after the same variable interval. For the Dog+ group, dogs were presented on positive trials and cats on negative trials, whereas the contingencies were reversed for Cat+ group. The trials were separated by an intertrial interval that varied between 5 and 15 s. Each session consisted of 80 trials, and was divided into 4 blocks of 20 trials containing 10 positive and 10 negative trials. The order of stimulus presentation was chosen pseudorandomly by the computer program, and constrained so that no more than three positive or negative stimuli were presented consecutively. Training sessions were normally given once per day, 6 days per week. Performance was assessed by the number of center-key pecks during the FT period, using the ρ statistic of Herrnstein, Loveland and Cable (1976). This value is a linear transformation of the Mann-Whitney *U*-statistic, and it estimates the probability that any positive stimulus is ranked above any negative stimulus, using rates of pecking as the measurement. When discrimination is perfect, ρ is 1.0; when there is no discrimination, it is 0.5. When a pigeon reached or exceeded a ρ value of 0.80 in each of three consecutive sessions, training was ended and generalization tests were begun.

In the generalization tests, 200 test images were presented over five test sessions in addition to the 20 training stimuli. These test images had never been shown to the subjects during the discrimination training. Each test session consisted of 80 trials including 40 training stimuli and 40 test stimuli. The test sessions started with 20 training trials. In the succeeding 60 trials, two out of three trials were test trials. The test stimuli were distributed across the 5 test sessions, controlling for the numbers of dogs or cats, mosaic or scrambled stimuli, and the five different levels of manipulations. Reinforcement contingencies were applied to all of the test stimuli in the same manner as during training so as not to discourage the pigeons from responding to test stimuli. Between the test sessions, training sessions were given until discrimination performance again reached 0.80 or above as measured by the ρ statistic.

Results

All twelve pigeons attained the criterion. The pigeons in the Dog+ group had experienced the same discrimination task before, so they were experiencing reacquisition. The number of training and reacquisition sessions required ranged from 3 to 12. To compare original acquisition rates, therefore, a comparison was made between the number of sessions taken to reach the criterion by the Dog+ pigeons in the previous experiment (Ghosh et al., 2004) and the pigeons in the Cat+ group in the present experiment. There was no significant difference of acquisition speed between Cat+ and Dog+ groups (Mann-Whitney $U_{6,6} = 18$; all statistical tests of the data in this paper were evaluated using an alpha level of 0.05). Although the group factor was included as a between-subject factor in the subsequent analysis, all the subsequent analyses showed no significant difference between the two groups. The data from the two groups were therefore merged into one in preparing the figures shown below.

Table 1 shows mean response rates to training S+, S- and each type of test stimuli, together with 95% naïve confidence intervals; in this and all subsequent figures and tables, confidence intervals are calculated following the recommendations of Masson and Loftus (2003). To capture the effect of the stimulus manipulation, relative response rates to the training positive stimulus were used in the subsequent analyses. Mean response rates to test stimuli did not change substantially or consistently over the 5 test sessions. Accordingly, test session number was not included as a factor in the subsequent analyses. Figure 2 (top) shows mean response rates, relative to the rate to the positive stimulus under training conditions, under the five different levels of mosaic and scrambling manipulations (level 0 indicates mean response rates for training stimuli during test sessions; this level was not included in the statistical analyses). The significance of the trends that can be seen in Figure 2 was assessed via a mixed design ANOVA with three within-subject variables (distortion type [mosaic vs. scrambled], stimulus valence [positive vs. negative], and distortion level [1 to 5]) and one between-subjects variable (Dog+ training vs. Cat+ training). Greenhouse-Geisser and Fisher LSD corrections were applied where appropriate. Discrimination was maintained across the test stimuli overall (significant main effect of stimulus valence overall, $F_{1,10}$ =82.53,for mosaicization only, $F_{1, 10} = 78.25$, and for scrambling only, $F_{1,10} = 51.94$), and performance was worse at higher levels of degradation (significant interaction between stimulus valence and manipulation level, $F_{4,40} = 14.60$). The simple main effect of valence was individually significant at all five levels of mosaicization ($F_{1,10} > 12.58$), and at levels 1 to 4 of scrambling ($F_{1,10} > 5.79$).

Overall the relative response rates were higher under scrambling, especially for negative stimuli (significant main effect of manipulation type, $F_{1,10} = 37.84$; significant interaction between manipulation and stimulus valence, $F_{4,40} = 36.10$). Since the manipulations tended to affect response rates to S- more than to S+, with an increment in response to S- as the manipulation level increased but little change in response to S+, these effects imply that the disruptive effect of scrambling was greater than that of mosaicization. Furthermore the pattern of disruption as manipulation level increased was different for the two manipulations (significant interaction between manipulation type and manipulation level, $F_{4,40} = 4.08$; significant three-way interaction between manipulation, stimulus valence and manipulation level, $F_{4,40} = 13.62$). No other main effect or interaction was significant, and in particular no overall difference was found between the Cat+ and Dog+ groups, nor was there any significant interaction between group and any factor (Fs < 0.97).

Discussion

The results of tests with mosaic and scrambled forms of the training stimuli were qualitatively consistent with previous work on picture perception in pigeons, which has shown generalization with some decrement following both mosaicization (e.g., Troje et al., 1999) and scrambling (e.g., Aust & Huber, 2001). In both manipulations, the changes of mean response rates were largely confined to negative stimuli, as is commonly the case following go/no-go discrimination training in pigeons. However, the patterns of generalization decrement were different between the two manipulations. While response rates to negative stimuli increased in an orderly way according to the level of mosaicization, the increase largely took the form of a step function with scrambling. Discrimination was not completely lost, since significant discrimination was seen at intermediate levels of scrambling. However, disruption of the discrimination was severe even at the lowest level of scrambling, whereas effectively no disruption was observed at the lowest level of mosaicization, though disruption increased steadily after that point; at the highest levels of mosaicization, discrimination was at about the same level as at the first level of scrambling.

In terms of pixel-based similarity, scrambling distorts our stimuli more than does mosaicization (see Supplementary Materials). This difference predicts some of the gross features of the generalization gradients shown in Figure 2. In particular, it accounts for the more severe degradation produced by scrambling

In terms of spatial frequencies, the spatial frequency ranges that can escape disruption as a result of the mosaicization and scrambling manipulations are illustrated in Figure 3. The response to mosaicization shown in Figure 2 implies that there was no information critical to the discrimination at spatial frequencies above around 24 cycles/picture, since the discrimination was essentially unaffected by the removal of frequencies in that range. Removal of frequencies in the range 24 down to 4 cycles/picture by increased levels of mosaicization disrupted discrimination but did not eliminate it, suggesting that information at frequencies in this range contributed to the discrimination. However, mosaicization also introduces spurious high frequency information at tile boundaries, and this is known to cause some generalization decrement in humans (Gordon & Field, 1978). It is therefore more conservative to focus on the evidence of sustained discrimination rather than the evidence of disruption, even though such high frequency noise appears to be small in extent (see Supplementary Materials). Even on this basis, however, we can draw the conclusion that the information at spatial frequencies below 4 cycles/picture (mosaicization level 5) was sufficient to maintain the discrimination. Correspondingly, the response to scrambling demonstrates that spatial frequencies above 12 cycles/picture (scrambling level 4) are also sufficient to maintain the discrimination. Overall, these results suggest that pigeons' discrimination of these pictorial stimuli is not a simple matter of dependence on any one spatial frequency range rather than another. Both higher (above 12 cycles/picture) and lower (below 4 cycles/picture) spatial frequency information were sufficient to sustain the pigeons' discrimination of these pictures.

Experiment 1b: Mosaic and scrambled stimuli in humans

Previous pigeon studies using the scrambling manipulation have not studied human participants simultaneously, so no corresponding data to that obtained from pigeons has so far been obtained (Aust & Huber, 2001; Matsukawa, Inoue & Jitsumori, 2004). Watanabe (2001) compared the effect of mosaicization on picture discrimination in pigeons and humans. However, the methodological and procedural differences between the human and pigeon tests mean that his comparisons are hard to interpret. For example, discrimination was trained on a go/no-go discrimination in pigeons, but it was trained on a binary choice task in humans. Moreover, the stimuli were different in size for the two species, and this is crucial if the absolute spatial frequencies involved are important (Goto, Wills & Lea, 2004). In Experiment 1b, therefore, we tested humans' responses to the same stimuli as presented to pigeons in Experiment 1a, using a go/no-go discrimination procedure analogous to that used with pigeons with stimuli that subtended the same visual angle as those used with the pigeons.

Method

Subjects.

Twelve human adults (*Homo sapiens*) from the University of Exeter were recruited as subjects. Their participations were rewarded either by 0.5 credits for their course requirement or the payment of 2 GBP. All subjects had normal or corrected-to-normal vision. All the subjects were naïve to the present categorization task.

Apparatus

Participants were individually tested in a quiet testing room using a PC-compatible computer (Pentium III 500 MHz, Tiny). E-prime (Psychology Software Tools, Inc.), running under the Windows 95 operating system, was used to control experimental events and record participants' responses. The distance between the participant's face and the 17-inch cathode ray tube monitor was approximately 30 cm.

Stimuli

The stimuli were the same as those used in Experiment 1a. The size of stimuli was adjusted to 22.5 degrees square, the same visual angle as used for pigeons in Experiment 1a.

Procedure

Participants were tested in a single session of three blocks. The participants were instructed to learn to press a key only when a category described as "N" appeared on the monitor (the nature of the category was not specified in the instructions). For half the participants, category 'N' comprised the dog face images; for the other half it comprised the cat face images.

The session started with a training block, during which the start of each trial was signaled by presenting a small fixation cross for 500 ms, followed by the presentation of the stimulus for 2 s. Participants were asked to press the "n" key on the computer keyboard if they thought the stimulus belonged to the category “N”, and not to press anything if they thought the stimulus did not belong to the category “N”. The stimulus disappeared after 2 s regardless of when or whether the participant made a response. When (and only when) the participants had made at least one response to a category “N” stimulus, the offset of the stimulus was followed by the feedback message “Correct”, which was displayed for 500 ms. No feedback was given if a participant failed to respond to a category “N” stimulus, and no feedback was given if a participant incorrectly responded to a stimulus not in category “N”. Therefore, the feedback situation mirrored closely that of the pigeons in Experiment 1a. The presentation or omission of feedback was followed by a 500-ms intertrial interval, and then the next trial was started. The order of stimulus presentation was chosen randomly with the following constraints: each training stimulus was presented only once in every 20 trials, and dogs (or cats) were presented in no more than three consecutive trials. The training block continued for at least 20 trials, during which all the training stimuli were exposed to the participant, and was then scheduled to continue until either the participant had made 10 consecutive correct responses or had completed 80 training trials in total. Following the training block, the two test blocks were started after a break of a few seconds.

The procedure in the test blocks was the same as the training procedure except that no feedback was given following the participant’s response. The 200 test stimuli used in Experiment 1a were presented across two test blocks. During test blocks, no more than three dogs (or cats) were presented consecutively. No training stimuli were presented during test blocks. When the participant had completed 100 test trials, a further short break was given. The second test block was started by the participant pressing a key. The procedure for the second test block was the same as the first test block. When the participant had completed the other 100 trials, the session was ended by the message ‘Thank you for your participation’.

Results

Figure 2 (bottom) shows mean response probability to each type of test stimulus. The significance of the differences apparent in the figure was tested via ANOVA, with one between-subjects factor (Dog+ training vs. Cat+ training) and three within-subject factors (distortion type [mosaic vs. scrambled], stimulus valence [positive vs. negative], and distortion level [1 to 5]), using data from test trials only. Discrimination was maintained across the test stimuli as a whole (significant main effect of stimulus valence overall, $F_{1,10} = 12.81$, for mosaicization, $F_{1,10} = 12.80$, and for scrambling, $F_{1,10} = 12.00$), but the discrimination performance was impaired by the manipulations (significant interaction between stimulus valence and manipulation level, $F_{4,40} = 7.66$). The loss of discrimination mainly took the form of a fall in response probability to S+ as the manipulation level increased, leading to a significant main effect of manipulation level ($F_{4,40} = 4.15$). Increasing levels of mosaicization differentially affected response probability to S+ and S- to a greater extent than did increasing levels of scrambling (significant three-way interaction between stimulus valence, manipulation type and manipulation level, $F_{4,40} = 4.13$). The simple main effect of valence was significant at each of the five levels of mosaicization ($F_{1,10} > 6.34$) and scrambling ($F_{1,10} > 10.24$). No other main effect or interaction was significant (Fs < 3.60).

Discussion

The impact of the two manipulations on the human participants did not exactly mirror their impact on the pigeons in Experiment 1a. The pigeons showed a greater impact of scrambling than mosaicization, but the humans showed, if anything, an opposite trend. Given that, in terms of pixel-based image similarity, scrambling distorts these images more than mosaicization, these results suggest that the humans were less affected than the pigeons by simple pixel-by-pixel similarity. Nevertheless, for the humans, as for the pigeons, the effects of mosaicization and scrambling imply that discrimination of these pictorial stimuli is not a simple matter of dependence on one spatial frequency range rather than another. Both high and low spatial frequency information were sufficient to sustain above chance performance to test stimuli, and therefore played some part in the humans' discrimination of these pictures. Information at spatial frequencies above around 16 cycles/picture (scrambling level 5) is sufficient but not necessary to sustain some discrimination, and so is information at frequencies below around 4 cycles/picture (mosaicization level 5).

A possible explanation of these results would be that relative to humans, pigeons are more affected by the disruption of low spatial frequencies (through scrambling) than by the disruption of high spatial frequencies (through mosaicization). However, an alternative explanation is that pigeons may be particularly sensitive to the introduction of greater levels of high-frequency noise in the scrambling manipulation than in the mosaicization manipulation.

Experiment 2a: Morph and cell-exchange stimuli in pigeons

Experiments 2a and 2b examined the effects of stimulus degradation further, by considering the impacts of morphing and cell exchange on learned concept discrimination. Experiment 2a used pigeons and Experiment 2b used human participants. Like the manipulations used in Experiments 1a and 1b, morphing and cell exchange affect the stimuli in different ways. However, whilst mosaicization and scrambling destroy the information in the stimuli, morphing and cell exchange dilute the information with conflicting information from the opposite stimulus class. In consequence, and unlike the two manipulations used in Experiments 1a and 1b, morphing and cell exchange are directly comparable: both produce probabilistic mixtures of the positive and negative stimuli. Both manipulations exchange portions of stimuli between a positive stimulus and a negative stimulus (in the present case, cat and dog faces). As with mosaicization and scrambling, the two manipulations have different effects in the spatial frequency domain. In particular, cell exchange introduces new sharp edges into the stimuli, and these will contribute some high frequency noise which is unlikely to occur in morphing. There are also differences are in terms of the effects on pixel-based similarity. Cell exchange produces a relatively gradual reduction in pixel-based image similarity, as material from one stimulus is moved into the other while other material remains in its original position. In morphing, however, everything is somewhat changed even at the lowest level of manipulation, so the change in pixel-based similarity is more abrupt. A more detailed analysis of the effects of morphing and cell exchange on our stimuli is provided in the Supplementary Materials.

Method

Subjects.

Twelve racing pigeons obtained as discards from local fanciers were used as subjects in this experiment. One pigeon was naïve and 11 had previously been used in an experiment involving a go/no-go discrimination procedure. None had been used in Experiment 1a of the present paper. Housing and maintenance were the same as in Experiment 1a.

Apparatus.

The apparatus consisted of the four identical three-key operant chambers used in Experiment 1a. The setup of the boxes and the data collection system was the same as in Experiment 1a. Each pigeon was assigned to a single test chamber for all stages of the experiment.

Stimuli.

The training stimuli were digital photographic images of eight faces of cats and eight faces of dogs. They were chosen from the stimuli used in Experiment 2 of Ghosh et al. (2004), selecting those whose size and color would be most suitable for producing test stimuli of the kind to be used in the present experiment. All images were resized to fit into a 6.0-cm square in order that the cell-exchange stimuli could be readily produced. Stimuli were presented at a size of 20.0 degrees square visual angle on the monitor. In addition to the 16 training stimuli, 144 images were prepared as test stimuli. Each cat face was paired with the dog face that was most similar to it size and overall color, as assessed subjectively by two raters. Two different sets of test stimuli were produced. In the morph set, the pairs of a cat face and a dog face were morphed using the program WinMorph 2.01. Morphing was correspondence-based: Between 80 and 100 marker points were used to establish correspondences between the pictures. The marker points were placed on the contour of the face and around the eyes, nose and mouth regions. The pictures were morphed progressively from 100% cat (0% dog) to 100% dog (0% cat), generating nine different levels of morphed pictures (90% cat / 10% dog to 10% cat / 90% dog, in 10% steps) to be used as test stimuli. In the cell-exchange set, the pictures were fragmented into 100 equally sized square tiles, and tiles in equivalent locations were exchanged between the paired cat and dog face images. The proportion of tiles exchanged was progressively increased so as to generate nine different levels of intermediate pictures (from 90% cat / 10% dog to 10% cat / 90% dog at 10% steps), to be used as test stimuli. Samples of these stimuli are shown in Figure 4.

Procedure.

Apart from one pigeon, pretraining was not required as the birds had experience of a previous discrimination task. Instead, they were given three sessions of recovery training, using the same schedule as the start key training used in Experiment 1a, to confirm that they could be immediately transferred to discrimination training. The naïve pigeon was pretrained in the same ways as the naïve pigeons used in Experiment 1a.

Once pretraining and recovery training were complete, the pigeons were arbitrarily divided into two groups of six. Half of the subjects were trained with cat faces as positive and dog faces as negative stimuli (Cat+ group), and the contingencies were reversed for the other half (Dog+ group). The discrimination was trained using a go/no-go discrimination procedure as in Experiment 1a. Each session consisted of 80 trials, and was divided into five blocks of 16 trials containing 8 positive and 8 negative trials. The order of stimulus presentation was chosen pseudorandomly by the computer program, and constrained so that no more than three positive or negative stimuli were presented consecutively. Training sessions were normally given six times in a week, one per day. The performance was assessed by the ρ statistic to the positive stimuli (Herrnstein et al., 1976), based on response rates during the first 10 s of stimulus exposure in each trial, during which no reinforcement could occur. When the bird reached a ρ value of 0.80 in three consecutive sessions, training was ended and generalization tests were begun.

In generalization tests, the 144 test stimuli were each presented once in the course of four test sessions; training stimuli were also presented. Each test session consisted of 80 trials of which 44 used training stimuli and 36 used test stimuli. The test sessions started with 20 trials using training stimuli. In the succeeding 60 trials, 3 out of each 5 trials were test trials, with the constraint that no more than two test trials occurred consecutively. Each of the four test sessions contained equal numbers of stimuli using morph and cell exchange and so far as possible equal numbers of the nine different levels of manipulation. The reinforcement contingencies applied to test stimuli were those appropriate to the manipulation type contributing the larger proportion of content: thus 60% cat / 40% dog test stimuli, and all those with a higher cat content, were reinforced for the Cat+ group but not reinforced for the Dog+ group. The reinforcement contingencies to test stimuli consisting of 50% cat / 50% dog were randomly assigned with the constraint that half of them had positive contingencies. Between the test sessions, additional training sessions were given until discrimination performance again reached 0.80 or above as measured by the ρ statistic.

Results

All 12 pigeons attained the criterion, requiring between 5 and 9 training sessions. The Dog+ group showed some tendency to reach criterion in fewer sessions than the Cat+ group, but the difference was not significant according to a Mann-Whitney test on number of sessions to criterion, $U_{6,6} = 7.5$.

Table 2 shows mean response rates to training S+, training S- and intermediate forms between the two produced by both morphing and cell exchange. To analyze the test session data, standardized peck rates were calculated by dividing the number of pecks to the test stimuli by each bird's mean number of pecks towards the positive stimuli over the 4 test sessions. Figure 5 (top) shows the standardized mean numbers of pecks during the first 10 s of exposure period to each type of test stimulus. The peck rates decreased as the proportion of the training S+ stimuli in the test stimuli decreased, but the pattern of generalization decrement was different between the two manipulations. An ANOVA with two within-subject variables (manipulation type [morph vs. cell exchange] and S+ proportion [90% -> 10%]) and one between-subject factor (Dog+ vs. Cat+) confirmed these effects by the presence of significant main effects of manipulation type ($F_{1,10} = 10.81$), and S+ proportion ($F_{1,10} = 62.09$), and a significant interaction between manipulation type and S+ proportion ($F_{8,80} = 11.91$). The main effect of group (Cat+ and Dog+) was not significant, and nor was any other interaction (Fs < 3.75).

Discussion

The results indicated that pigeons' response rate to intermediate forms decreased in an orderly fashion with progressive reductions of S+ proportion within a picture, whether that reduction was achieved by morphing or by cell exchange. Recall that, unlike mosaicization and scrambling, these two manipulations are directly comparable, because both produce probabilistic mixtures of positive and negative stimuli. The patterns of generalization gradients for morphing were similar to those previously reported (e.g., Ghosh et al., 2004; Makino & Jitsumori, 2001). However the generalization gradients under the two manipulations were not identical. More responses were evoked to the intermediate forms with the cell exchanges than to morphs, and examination of Figure 5 shows that the origin of this effect was that, when stimuli contained more S- than S+, pigeons responded more rapidly to cell-exchange stimuli than they did to morph stimuli.

Considering our stimuli in terms of pixel-based image similarity contributes little to the understanding of the results of this experiment. In pixel-by-pixel similarity terms, morphing distorts our stimuli more rapidly than does cell exchange. From this, one might predict that morphing would be more disruptive to the discrimination. In fact, the opposite effect was found.

.

Experiment 2b: Morph and cell-exchange stimuli in humans

Experiment 2b continued the comparison between pigeon and human performance by testing human participants with the morph and cell-exchange stimuli used with pigeons. If the effect of these stimulus manipulations is similar in the two species, a similar pattern to pigeons in Experiment 2b should be found in this experiment, with cell exchange leading to greater degradation of discrimination than morphing.

Method

Subjects.

Twelve students from the University of Exeter were recruited as subjects. Their participations were rewarded either by 0.5 credits for their course requirement or the payment of 2 GBP. All subjects had normal or corrected-to-normal vision. All the participants were naïve to the present categorization task and none had participated in Experiment 1b.

Apparatus.

Participants were individually tested in a quiet testing room using a PC-compatible laptop computer (PCG-505V, Sony). E-prime (Psychology Software Tools, Inc.), running under the Windows 98 operating system, was used to control experimental events and record participants' responses. The distance between the subject's face and the monitor was approximately 30 cm.

Stimuli.

The stimuli were the same as those used in Experiment 2a. The size of stimuli was adjusted to 20.0 degree square, the same visual angle as used for pigeons in Experiment 2a.

Procedure.

The procedure was the same as that used in Experiment 1b. Participants were tested in a single session of three blocks. Half of the subjects were allocated to press a key in the presence of cat face images (Cat+), whereas the other half were assigned to press when dog face images were presented (Dog+). The session started with a training block. The order of stimulus presentation was chosen randomly, and constrained so that each training stimulus was presented only once in every 16 trials. The training block continued for at least 16 trials, during which all the training stimuli were exposed to the participant, and was then scheduled to continue until either the participant had made 10 consecutive correct responses or 80 responses in total. When performance reached criterion, the two test blocks were started after a short break.

The procedure in the test blocks was the same as the training procedure except that no feedback was given following the participant's responses. Each test block consisted of 72 test trials, during which only the test stimuli were shown. When the participant completed the first test block, a further short break was given. The second test block was started following the participant pressing the space key. The procedure for the second test block was the same as for the first test block. When the participant had completed the other 72 trials, the session was ended by the message 'Thank you for your participation'.

Results

Figure 5 (bottom) shows response probability to morph and cell-exchange stimuli. The significance of differences in the probability dependent measure was tested by an ANOVA with two within-subject variables (manipulation type and S+ proportion), and one between-subject variable (Dog+ vs Cat+). Response probability decreased as the S+ proportion within a picture declined (significant main effect of S+ proportion, $F_{8,80} = 214.37$). However, the decrement began at higher S+ proportions in morph pictures than in cell-exchange pictures (significant interaction between manipulation and S+ proportion, $F_{8,80} = 3.14$). There was also a significant interaction between manipulation type and group ($F_{1,10} = 8.85$), although as this effect was not significantly modulated by proportion of S+ (i.e. no significant S+ proportion x manipulation type x group interaction) it seems to be of little consequence to the central questions addressed by this study. No other main effect or interaction was significant (Fs < 2.39).

Discussion

The overall effects of morphing and cell exchange were similar to those in pigeons. Response probability decreased progressively as the proportion S+ within a picture fell. However, significant differences in the pattern of degradation of discrimination were found for morphing and cell exchange.

Superficially this result is consistent with the pattern found in pigeons in Experiment 2b, but comparison between pigeon and human performance in Figure 5 shows that the two results are in fact contradictory. The pigeons' high response rates to cell-exchange stimuli occurred to stimuli with a high proportion of S-, and therefore represents poor discrimination, whereas the humans' high response probability occurred to stimuli with a high proportion of S+ and thus represents sustained discrimination. To put it another way, under cell exchange but not under morphing, the pigeons generalized their responding even to stimuli that contained a very low proportion of S+ material. Humans showed no such tendency.

The human data are what one might expect from pixel-based image similarity – morphing degrades this similarity more rapidly than cell exchange, and humans response to the S+ reduces more rapidly for morphing than for cell exchange. As noted earlier, the pigeon data with these manipulations cannot be predicted in this way. This contrasts with the conclusion drawn about pixel-based image similarity in Experiment 1, where pigeon, but not human, generalization performance could be accounted for with this metric.

General Discussion

In this study, pigeons and humans were tested to investigate the effects of four types of image manipulations – mosaicization, scrambling, morphing, and cell exchange. In Experiment 1 we degraded positive and negative stimuli by mosaicization (a fairly pure low-pass spatial frequency filter) or by scrambling (a high-pass spatial frequency filter, but contaminated by induced broad spectrum noise). The results of the mosaicization and scrambling manipulations suggested that, for both humans and pigeons, both high and low spatial frequency information were sufficient but not necessary to maintain the discrimination.

In Experiment 2 we examined transfer to stimuli intermediate between the S+ and S-, with the intermediate forms generated either by morphing, or by cell exchange. Cell exchange preserves some of the high frequency detail that morphing damages. As others have previously reported, morphing led to an orderly loss of discrimination as the proportion of S+ in the stimulus declined (see also e.g. Makino & Jitsumori, 2001; Ghosh et al., 2004). Relative to the morphing manipulation, pigeons were more adversely affected than humans by cell exchange, with pigeons continuing to respond to stimuli that contained a very low proportion of the S+. This result is consistent with the idea that pigeons' generalization performance is relatively more under the control of high spatial frequency information than is that of humans.

Hence, whilst Experiment 1 suggests that both high and low spatial frequency information can be sufficient to maintain discrimination in both species, Experiment 2 suggests that high-spatial frequency information has greater control over responding in pigeons than in humans. Such a conclusion is consistent with the view that there are quite general species differences in visual processing. For example, Cavoto and Cook (2001), argued that pigeon visual processing of hierarchical stimuli is more dependent on local than global features, whereas Navon (1977), using similar hierarchical stimuli, argued for configural feature dominance in humans.

In terms of pixel-based image similarity (as measured by cross-correlation analyses), our cell exchange manipulation distorted the images more than our morphing manipulation, and our scrambling manipulation distorted the images more than our mosaicization manipulation. These differences did not, however, have any consistent relationship to performance in either species. Other forms of image analysis (in addition to the spatial frequency and pixel-based similarity analyses we have performed) might have provided further insights in the behavior of our participants, but this remains an issue for future research.

Inevitably, the procedures for humans and pigeons were not identical, and even if they were, there would be no way of knowing that their impacts on the two species were identical. However, since the purpose of the experiments was to examine the relative rather than the absolute impacts of different stimulus manipulations, it is not obvious how the residual procedural differences could have produced the results we observed. The procedures used here were similar to those used in successful pigeon/human comparisons, e.g. Spetch, Friedman and Vuong (2006).

One limitation of the current studies is the likely difference in expertise with the stimulus sets in humans and pigeons. Gibson, Wasserman, Gosselin and Schyns (2005) showed that (after training the pigeons for 60 days), pigeons and humans used similar regions of human face images to discriminate gender (and emotion). Nevertheless, some differences between humans and pigeons remained, and Gibson et al. (2005) attributed these to humans' greater expertise with faces. In a similar manner, the human participants in our experiments presumably had more experience of dog and cat faces than the pigeons. This greater expertise may have increased the humans' tolerance to the picture manipulations, and this may in turn have contributed to the species difference we observed. It may also be the case that the relatively small number of stimuli we employed made it difficult for the pigeons to gain much expertise in these categories during the experiment. In future studies, one might use images that are matched for expertise.

The image manipulation techniques employed in the current paper, with the exception of cell exchange, have been employed in previous studies of avian visual cognition (although they have not been compared on the same stimuli). Our pixel-based similarity and spatial frequency analyses of these manipulations indicate that their effects are quite complex. These complexities may prompt future researchers to consider alternative methods of image manipulation - such as purer forms of spatial frequency filtering (cf. Schyns & Oliva, 1994), or reverse-correlation methods (e.g. Gibson et al., 2005; Martin-Malivel, Mangini, Fagot & Biederman, 2006)

In summary, the present experiments have produced some evidence that a bias towards elemental processing in pigeons may cause some species differences in the perception of naturalistic images, as it does for simple geometrical stimuli. However, our results also show that the impact of different stimulus manipulations involves an interaction of species and stimulus factors. Further research is needed to investigate the complex and subtle differences between the visual cognition of pigeons and humans. It does seem clear, however, that such research needs to compare the two species' responses using stimuli and procedures that are as closely matched as possible, as was done in the present experiments.

References

Allan, S. E. & Blough, D. S. (1989). Feature-based search asymmetries in pigeons and humans, *Perception & Psychophysics, 46*, 456-464.

Aust, U. & Huber, L. (2001). The role of item- and category-specific information in the discrimination of people versus nonpeople images by pigeons. *Animal Learning & Behavior, 29*, 107-119.

Aust, U., & Huber, L. (2003). Elemental versus configural perception in a people-present/people absent discrimination task by pigeons. *Learning & Behavior, 31*, 213-224.

Cavoto, K. K., & Cook, R.G. (2001) Cognitive precedence for local information in hierarchical stimulus processing by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 27*, 3-16.

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

Delius, J. D., & Hollard, V. D. (1995). Orientation invariant pattern recognition by pigeons (*Columba livia*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 109*, 278-290.

Donis, F. J., & Heinemann, E. G. (1993). The object-line inferiority effect in pigeons. *Perception & Psychophysics, 53*, 117-122.

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

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A: Optics Image Science and Vision, 4*, 2379-2394.

Ghosh, N. Lea, S. E. G., & Noury, M. (2004). Transfer to intermediate forms following concept discrimination by pigeons: chimeras and morphs. *Journal of the Experimental Analysis of Behavior, 82*, 125-141.

Gibson, B. M., Wasserman, E. A., Gosselin, F., & Schyns, P. G. (2005). Applying Bubbles to localize features that control pigeons' visual discrimination behavior. *Journal of Experimental Psychology: Animal Behavior Processes, 31*, 376-382.

Gordon, I. E., & Field, D. (1978). Possible explanation as to why newly sighted commonly perform well on pseudoisochromatic color-vision tests. *Perception, 7*, 119-122.

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.

Harmon, L. D. (1973). The recognition of faces. *Scientific American, 229*, 71-82.

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

Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 36*, 405-445.

Huber, L. (2001). Visual categorization in pigeons. In R. G. Cook (Ed.), *Avian visual cognition* [On-line]. Available: www.pigeon.psy.tufts.edu/avc/huber/

Kirkpatrick-Steger, K., Wassermnan, E. A., & Biederman, I. (1996). Effects of spatial rearrangement of object components on picture recognition in pigeons. *Journal of the Experimental Analysis of Behavior, 65*, 465-475.

Kirkpatrick-Steger, K., Wasserman, E. A., & Biederman, I. (1998). Effects of geon deletion, scrambling, and movement on picture recognition in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 24*, 34-46.

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

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

Makino, H., & Jitsumori, M. (2001). Category learning and prototype effect in pigeons: A study with morphed images of human faces. *Japanese Journal of Psychology, 71*, 477-485.

Martin-Malivel, J., Mangini, M. C., Fagot, J., & Biederman, I. (2006). Do humans and baboons use the same information when categorizing human and baboon faces? *Psychological Science, 17*, 599-607.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation, *Canadian Journal of Experimental Psychology, 57*, 203-220.

Matsukawa, A., Inoue, S., & Jitsumori, M. (2004). Pigeon’s recognition of cartoons: Effects of fragmentation, scrambling, and deletion of elements, *Behavioural Processes, 65*, 25-34.

Nakamura, N., Fujita, K., Ushitani, T., & Miyata, H. (2006). Perception of the standard and the reversed Muller-Lyer figures in pigeons (*Columba livia*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 120*, 252-261.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353-383.

Schyns, P.G. & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition, *Psychological Science, 5*, 195-200.

Sekuler, A. B., Lee, J. A. J., & Shettleworth, S. J. (1996). Pigeons do not complete partly occluded figures. *Perception, 25*, 1109-1120.

Spetch, M. L., Friedman, A., & Vuong, Q. C. (2006). Dynamic object recognition in pigeons and humans. *Learning & Behavior, 34*, 215-228.

Troje, N.F., Huber, L., Loidolt, M., Aust, U., & Fieder, M. (1999). Categorical learning in pigeons: the role of texture and shape in complex static stimuli. *Vision Research, 39*, 353-366.

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.

Watanabe, S. (2001) Van Gogh, Chagall and pigeons: Picture discrimination in pigeons and humans. *Animal Cognition, 4*, 147-151.

Authors' Note

The experimental work was conducted by Kazuhiro Goto in partial fulfillment of the requirements for the PhD degree at the University of Exeter, with support from Experimental Analysis of Behavior Fellowship from the Society for the Advancement of Behavior Analysis. The preparation of the manuscript was further supported by grants from the NIMH (MH068426) and the Japan Society for the Promotion of Science (KG) and by European Commission Framework 6 (NEST) Project 516542, 'From Associations to Rules' (SEGL and AJW). Housing of the animals and conduct of the experiments conformed with the requirements of the UK Animals (Scientific Purposes) Act, 1986. Correspondence concerning this article should be addressed to Kazuhiro Goto, who is now at Kokoro Research Center, Kyoto University, 46 Shimoadachi-cho Yoshida Sakyo-ku Kyoto, 606-8501, Japan. (kgoto@psy.flet.keio.ac.jp).

Figure captions

Figure 1. Experiments 1a and 1b: Examples of training and test stimuli. One stimulus of each manipulation from cats and dogs are shown above. Stimuli with no manipulation are shown on the left of each row (training stimuli). These upper rows for each stimulus show the effect of successive levels of the mosaic manipulation and the lower rows show the effect of scrambling. In the experiments, the stimuli were shown in full color.

Figure 2. Discrimination performance in mosaic and scrambling tests in pigeons (top) and humans (bottom). Pigeons' relative mean responses were calculated from the mean number of responses divided by the mean number of responses to the training positive stimuli during the test sessions for each subject. In these figures, the greater the difference of response rates or response probability between S+ and S-, the better the discrimination. Error bars are the 95% confidence intervals for this repeated measures design.

Figure 3: Informational effects of spatial frequency filtering. Bars show the range of spatial frequencies that can be preserved for each manipulation level. Black bars indicate the most severe degradation that still produces significant discrimination in pigeons. The derivation of this figure is provided in the Supplementary Materials,

Figure 4. Experiments 2a and 2b: Examples of training and test stimuli. The first and last columns are training stimuli (100 % cats and 100 % dogs, respectively). First two rows are the examples of intermediate forms by morphing at every 10 % step (10 % cat - 90 % dog to 90 % cat -10 % dog). The other two rows are the examples of the intermediate forms by cell exchange at every 10 % step.

Figure 5. Discrimination performance in morph and cell-exchange tests in pigeons (left) and humans (right). Pigeons' relative mean responses were calculated from the mean number of responses divided by the mean number of responses of the training positive stimuli during the test sessions for each subject. Error bars are the 95% confidence intervals for this repeated measures design.

Table 1. Mean number of responses in 10-s trials to mosaic and scrambling test stimuli in pigeons. Mean response rates to training stimuli were those recorded during the 5 test sessions in Experiment 1a. The 95% confidence intervals for mosaic and scrambling are ±1.8 and ±1.0 respectively.

		Level of manipulation					
Manipulation	Stimulus Valence	None (training)	1	2	3	4	5
Mosaic	S+	16.3	15.6	14.8	14.9	13.9	13.1
	S-	5.6	4.8	5.8	9.1	9.1	9.7
Scrambling	S+	16.3	14.6	14.5	14.8	13.3	13.8
	S-	5.6	11.6	10.4	10.1	11.4	11.2

Table 2. Mean number of responses in 10-s trials to morphing and cell-exchange test stimuli in pigeons. Mean response rates to training stimuli were those recorded during the 4 test sessions in Experiment 2a. The 95% confidence intervals for this repeated measures design is ±1.3.

	Proportion of S+ stimulus										
	100 (Training)	90	80	70	60	50	40	30	20	10	0 (Training)
Morph	15.9	14.7	15.2	14.1	13.3	11.8	8.4	5.0	5.2	4.5	4.4
Cell exchange		15.6	14.9	14.0	14.0	12.4	12.5	11.9	9.2	5.7	

Figure1

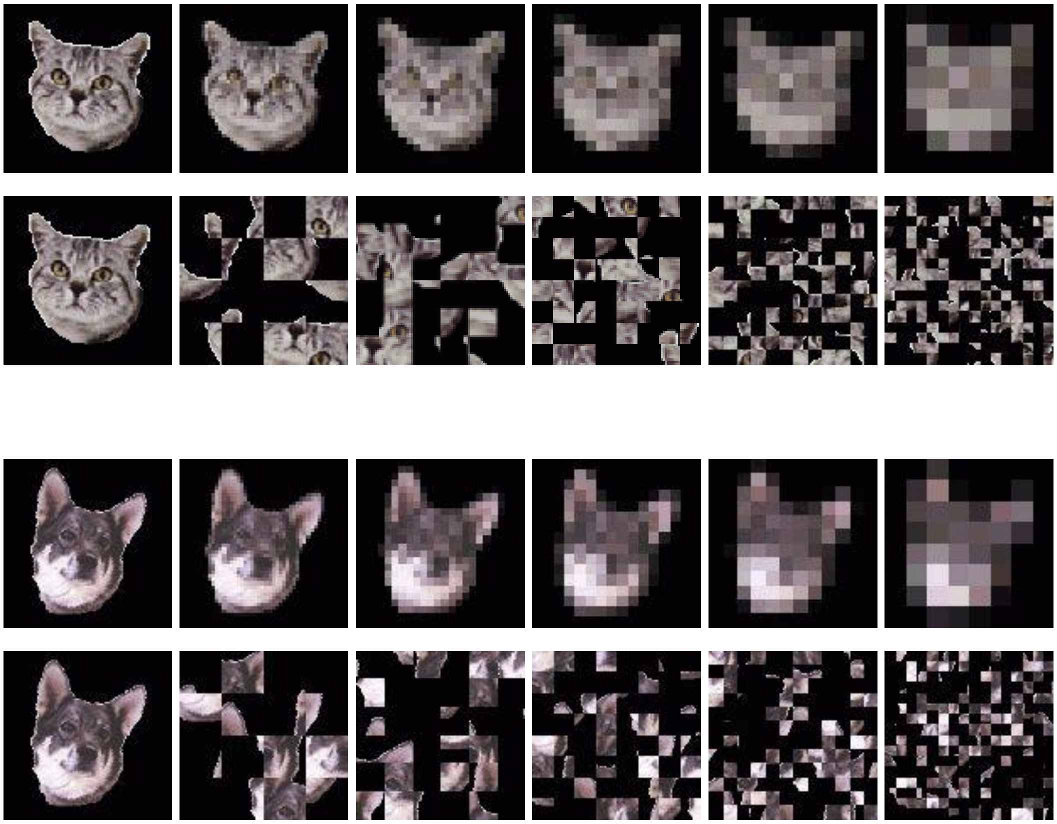
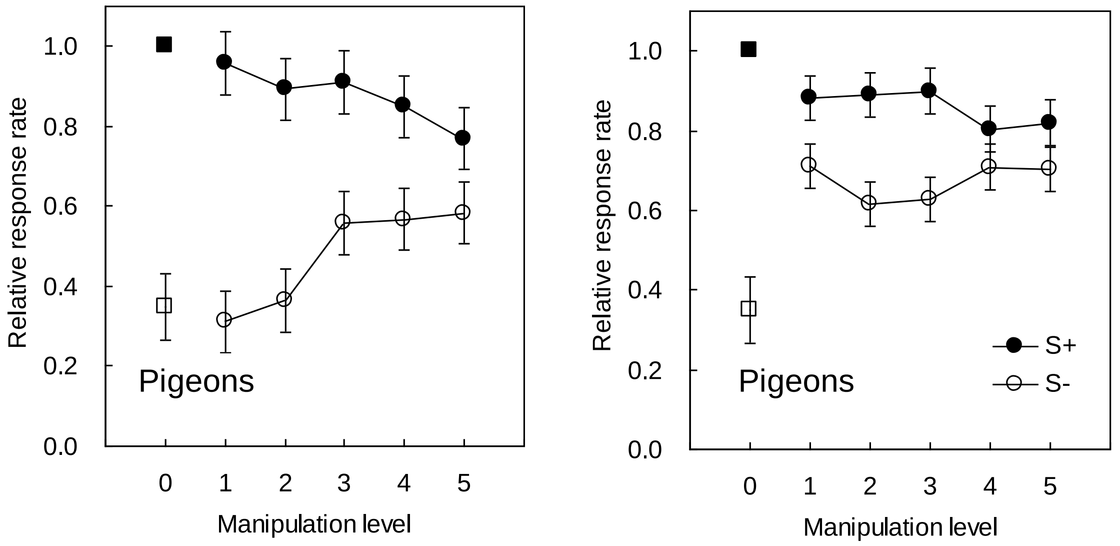
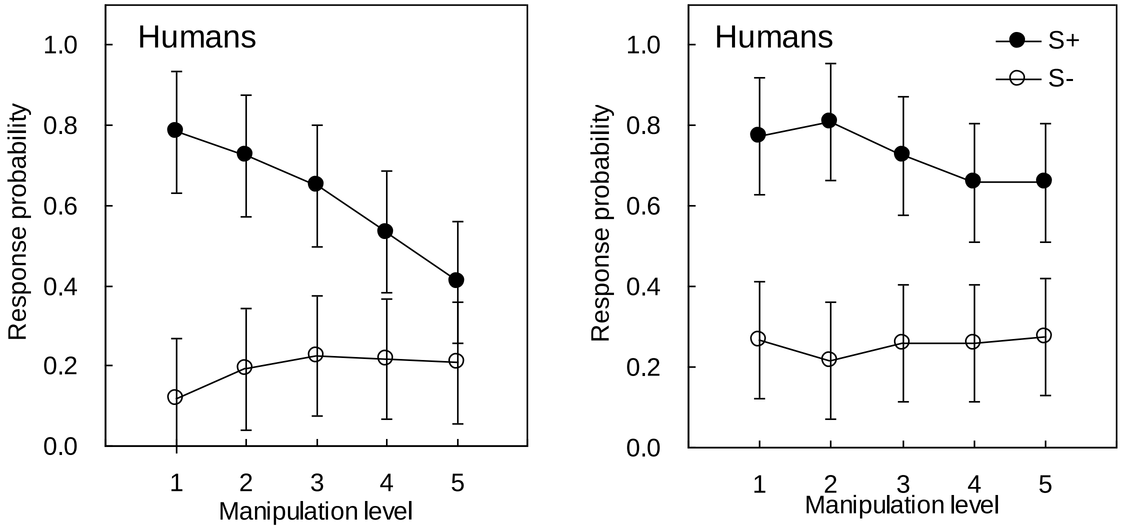


Figure 2



Mosaicization (low-pass filter)

Scrambling (high-pass filter)



Mosaicization (low-pass filter)

Scrambling (high-pass filter)

Figure 3.

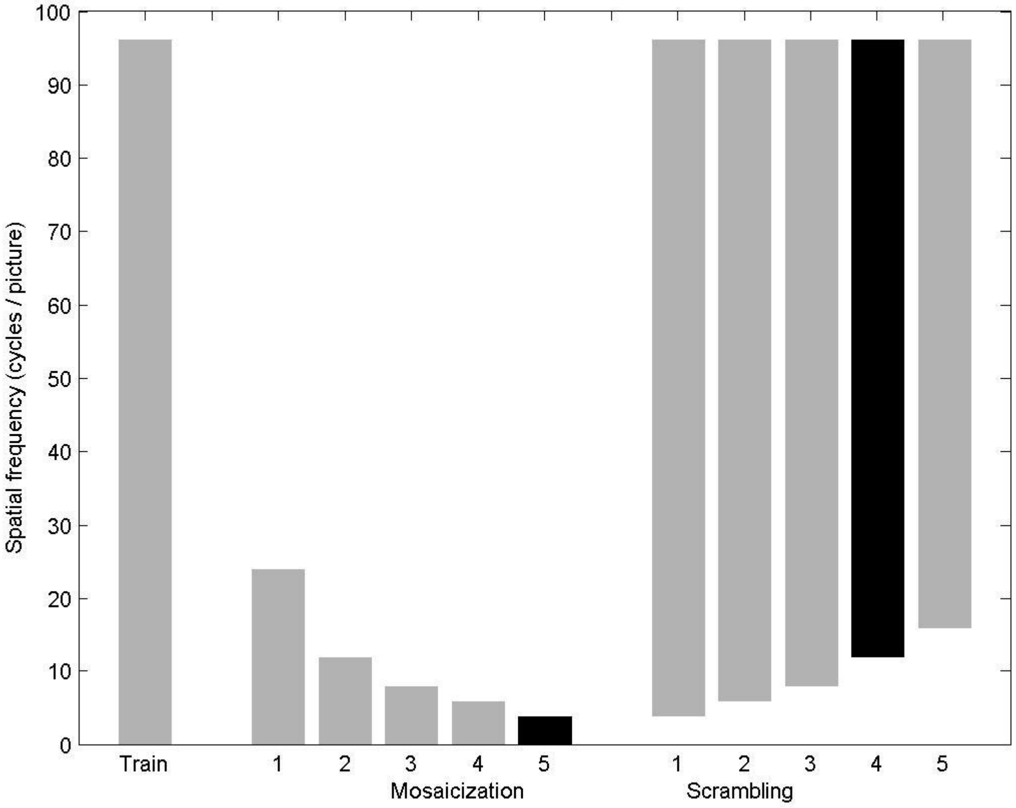
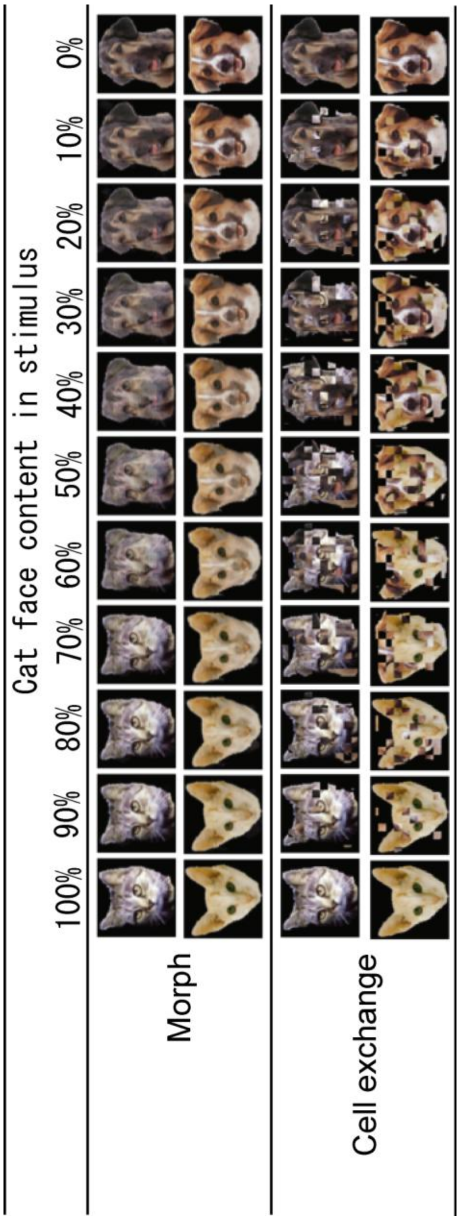
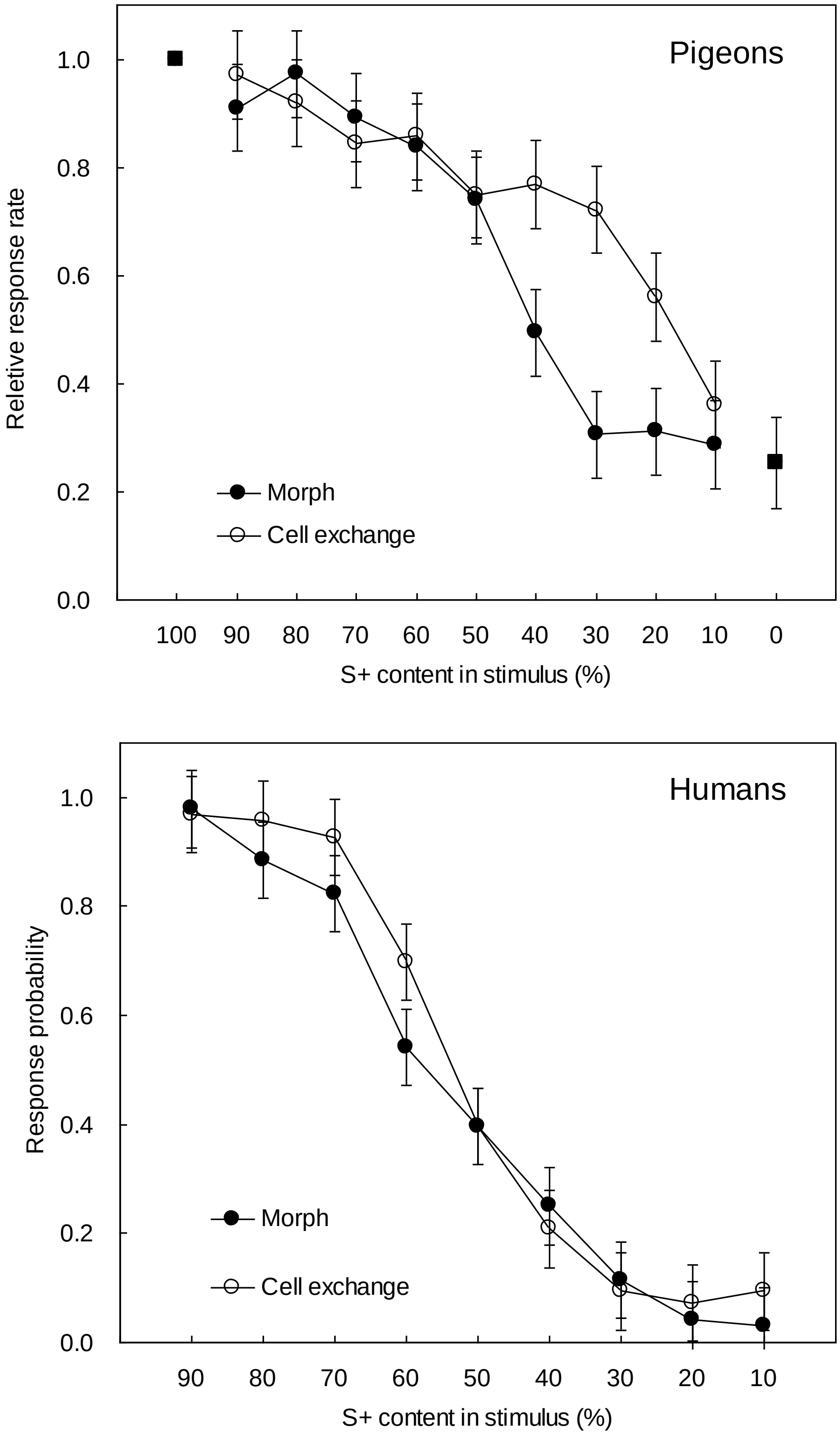

100
90
80
70
60
50
40
30
20
10
0
Spatial frequency (cycles / picture)
Train
1
2
3
4
5
Mosaicization
1
2
3
4
5
Scrambling

Figure 4


Cat face content in stimulus
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Morph
Cell exchange

935 Figure 5



936