Title	Effects of various spatial transformation on pigeons' recognition of visual patterns : some applications of the Heinemann-Chase pattern recognition model
Sub Title	
Author	Heinemann, Eric G.
Publisher	慶應義塾大学大学院社会学研究科
Publication year	1997
Jtitle	慶応義塾大学大学院社会学研究科紀要 : 社会学心理学教育学 (Studies in sociology, psychology and education). No.45 (1997.) ,p.1- 10
JaLC DOI	
Abstract	
Notes	シンポジューム : Pattern recognition in humans and animals
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN0006957X-00000045-0001

EFFECTS OF VARIOUS SPATIAL TRANSFORMATIONS ON PIGEONS' RECOGNITION OF VISUAL PATTERNS: SOME APPLICATIONS OF THE HEINEMANN-CHASE PATTERN RECOGNITION MODEL.

*Eric G. Heinemann**

How pigeons discriminate between, recognize, categorize, or identify visual patterns has been examined in numerous experiments. The methods used depend on which of these processes is under investigation. To avoid cumbersome terminology, I shall use the term "recognition" whenever the process referred to is clear from the context.

The purpose of this paper is to review some experiments that deal with the way pigeons' recognition of patterns is affected by translation and by some specified rotations, and to interpret these findings in terms of the Heinemann-Chase information-processing theory. It is not my purpose to review other theories that deal with some of the phenomena to be discussed. One reason for this decision is that an adequate review would be inordinately long. Another reason is that our's is an extremely general theory of the pigeon's cognitive processes. Its strength is that it deals, in quantitative detail, with a wide range of phenomena including probability learning, discrimination learning involving unidimensional and multidimensional stimuli, the phenomena of blocking and overshadowing, as well as an account of the "range effect" (Pollack; 1956; Durlach & Braida, 1969) and an account of the limitations on information transmission discussed in G. Miller's well-known article on the "Magical number seven" (Miller, 1956; Chase, 1983). For a discussion of the significance that Chase's account of the last-mentioned phenomenon has for results obtained with human subjects see Sperling, (1988).

The theory has also been applied, with considerable success, to problems of visual patternrecognition. For example, it has yielded success-

---

* Hunter College, City University of New York, New York.

ful quantitative predictions of confusion matrices obtained by D. Blough (1985) for the 26 letters of the alphabet and for patterns of randomly arranged dots.

Does training to recognize stationary two-dimensional projections of three-dimensional objects result in viewpoint-independent recognition of these objects?

When the stimuli that pigeons are trained to recognize are visual patterns, these are most often presented in the frontal-parallel plane and may represent particular perspective projections of three-dimensional objects. In their everyday lives pigeons are certainly required to identify three-dimensional visual objects. An important question is whether, and under what conditions, pigeons can learn to recognize three-dimensional objects by being trained to recognize two-dimensional projections of these objects.

According to Cerella (1990), training with one perspective projection does not lead to recognition of other projections. In one of his experiments pigeons were trained to peck on a response key when shown an outline figure in the form of a chevron, called the "prototype" This two-dimensional form may be considered as a projection of a chevron made of wire. During a subsequent generalization test pigeons were shown other two-dimensional shapes that were produced by rotating the wire form that corresponded to the prototype about its x, y, or z-axis. Cerella also tested for generalization to prototypes that were reduced or magnified in size, or distorted randomly. The results of the generalization tests are shown in Fig. 1 (solid lines), together with the results of simulations based on our theory (dashed lines). If the pigeons were able to recover the three-dimen-
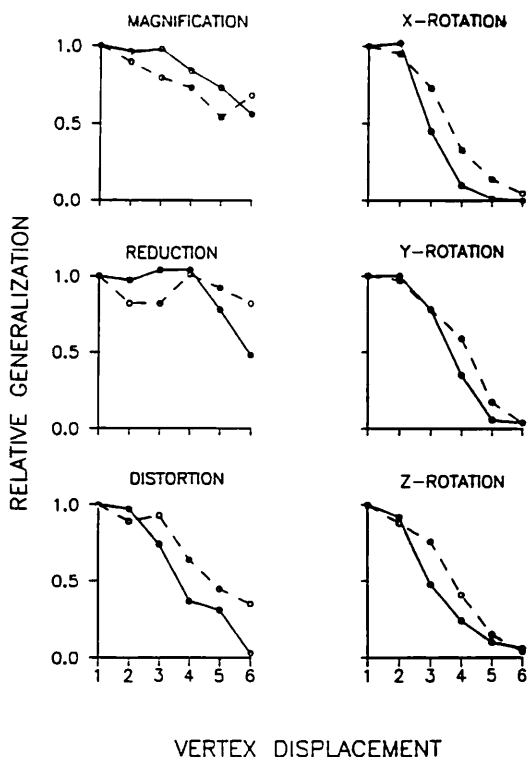
VERTEX DISPLACEMENT

Fig. 1. Relative generalization to various transformations of a two-dimensional shape measured by Cerella (1990). The solid lines represent the mean results for Cerella's pigeons, the dashed lines the results of simulations. The horizontal axes of the various panels have a common scale, a measure of prototype distortion in terms of vertex displacement.

sional form of the wire figures from the two-dimensional projections shown during the generalization tests then, ideally, all the generalization curves for rotation would be horizontal lines. Instead, all these curves drop steeply in the manner expected if responding were based simply on the similarity of the two-dimensional forms to the prototype, as specified by our simulations.

The curves for magnification and reduction fall less steeply than those for rotation. This is true for the empirical data and our simulations. Our treatment of this matter is grounded in some details of the pigeons pecking behavior. Using high-speed cinematography, Goodale (1983) found that, before pecking at food or at a response key pigeons make two head-fixations, one about 55 mm from the target (the F2 position) the other about 80 mm from the target (the F1 position). On the basis of the variance of the trial-by-trial distributions of head fixations it is possible to compute the variance of the trial-by-trial distribution of retinal image sizes. (For a somewhat more detailed discussion of this matter see Heinemann & Chase, 1990). When the fact that the retinal image varies from trial to trial is taken into account our theory yields the fits to Cerella's data shown in Fig. 1.

I now return to the question of whether pigeons can learn to recognize a solid object rotated through any angle by learning to recognize two-dimensional projections of that object. Perhaps Cerella's pigeons would have generalized to all projections if they had been trained to recognize several perspective transformations of the prototype rather than just one. This possibility was examined in an experiment by Ronald Stevens (1990). Stevens trained three groups of pigeons to discriminate between isometric projections of cubes and distortions of these projections. The cubes were produced by rotating an imaginary wire cube simultaneously about the x, y, and z axes in steps of 10 degrees (measured from a starting position in which one side faced the observer so that the cube appeared as a two-dimensional square). A distorted cube was produced by randomly displacing the vertices of a cube that is in its starting position ". . not more than half the length of a side in the x,y, or z planes, under the constraint that the mean displacement of the eight vertices would be zero. This produced an irregular three dimensional figure which was then rotated by 10 degree increments simultaneously about the x, y, and z axes through 360 degrees" (Stevens, 1990).

Four groups of four pigeons participated in a categorization experiment. Group 1 was trained to discriminate between the members of a single cube/distortion pair, Group 2 was trained with two pairs, Group 3 was trained with four pairs and Group 4 with eight pairs. The apparatus had two response keys. Pecks on one of these keys were reinforced when the stimulus presented was a cube, pecks on the other key were
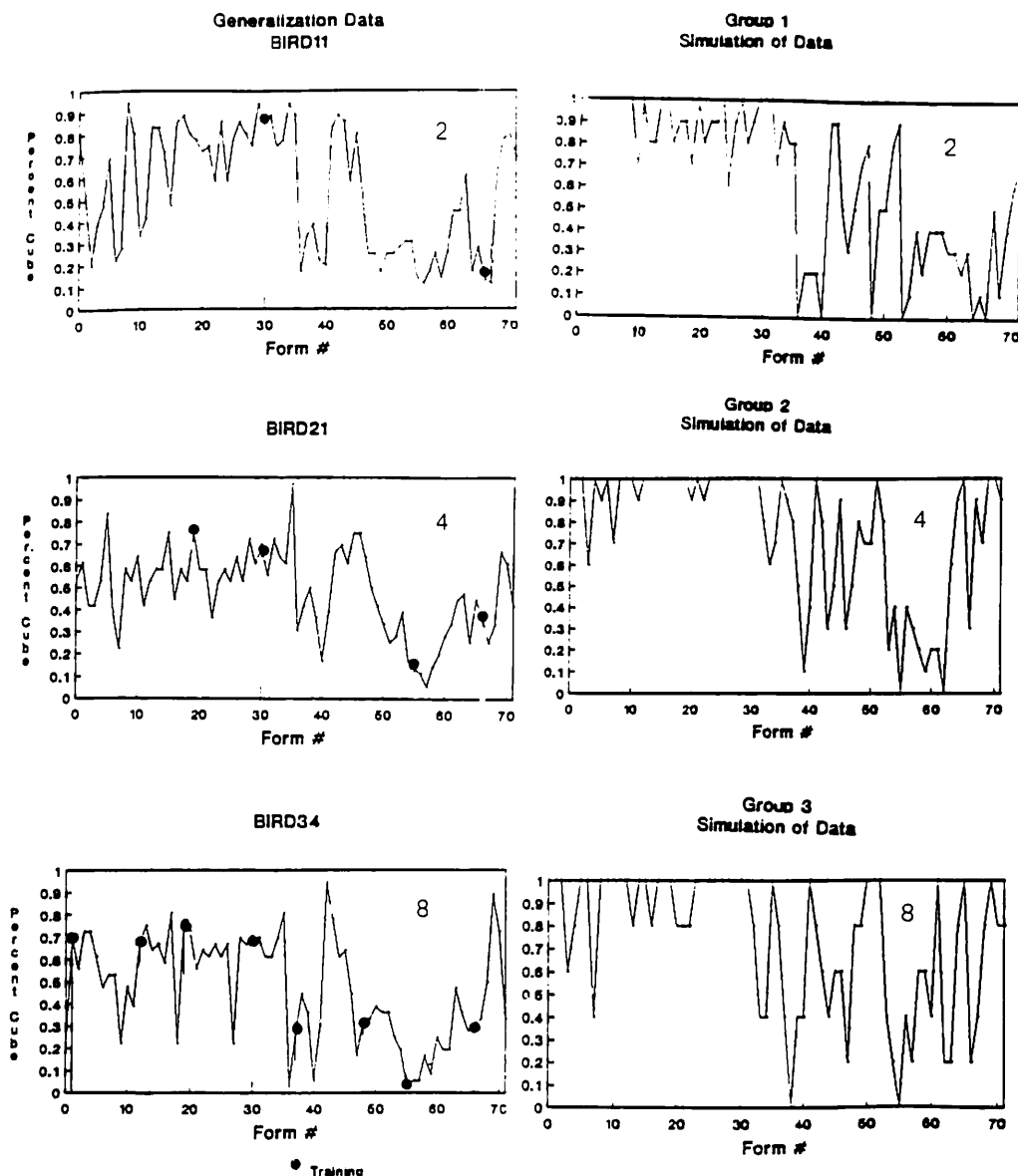
Fig. 2. Comparison of data obtained from three of Steven's (1990) birds and simulations of these conditions. The figures on the left are generalization test data for three birds-one from the group trained with two stimuli (top panel), one from the group trained with four stimuli (middle panel) and one from the group trained with eight stimuli (bottom panel). The training stimuli are shown as dots. Simulations of generalization test data for each training condition are shown in the left panel.

reinforced when the stimulus was a distortion.

After 120 daily training sessions only six pigeons had learned to categorize the stimuli at better than chance levels. Two were from Groups 1, one from Group 2, three from Group 3, and none from Group 4. The remaining birds were trained for an additional 40 days. When all of these birds failed to perform at better than chance levels after the additional training, they were withdrawn from the experiment.

After training, the six pigeons who categorized at better than chance levels were given generalization tests. During these tests they were presented with unique stimuli produced by the method described. There were 33 unique projections of the cube and 36 unique projections of the distorted cube. This set of stimuli included the training stimuli used for the various groups.

Fig. 2 (left panel) shows the proportion of trials on which the "cube-key" was pecked when the stimulus was a projection of a cube (forms 0 to 35), and when the stimulus was a projection of a distorted cube (forms 36 to 71). Results are shown for one bird from groups 1, 2 and 3. These generalization curves seem to be incompatible with the notion that the pigeons extracted the three-dimensional form of the cubes and are generalizing along the continuum of rotation for these forms. If they were generalizing in this fashion then the generalization functions would be horizontal lines. If, instead, they were generalizing only to stimuli that had a two-dimensional form similar to that of the training stimuli, then there would be a sharp, peaked generalization curve centered on each of the training stimuli. Stevens found no such trends in his results. It seems likely that the absence of a decremental gradient around the training stimuli reflects the fact that even the immediate neighbors of the training stimuli had two-dimensional forms that differed greatly from those of the training stimuli. Recall that neighboring forms were projections that were rotated simultaneously in x, y, and z dimensions using step sizes of 10 degrees.

In any case, the actual results appear disorderly and perhaps random at first glance. However, closer analysis reveals considerable order. All six of the pigeons performed at levels significantly better than chance when categorizing the generalization stimuli as well as the training stimuli. Further, the generalization curves for the six pigeons have remarkably similar shapes. As one measure of this similarity Stevens presents between-group correlations (Pearson r) based on average group data. These are 0.88 between Groups 1 and 2, 0.7 between Groups 1 and 3, and 0.74 between Groups 2 and 3.

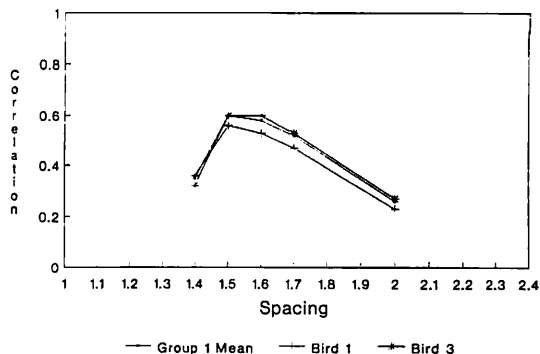It is difficult to think of an underlying dimen

Fig. 3. The correlation between the empirical and theoretical confusion matrices as a function of the spacing parameter used in the simulations (Stevens, 1990). Generalization test data are shown for Group 1, the birds trained with one view of the cube and one view of the distorted form.

sion along which "generalization" could be occurring to produce the results of Stevens' experiment. More fundamental: it is not clear what psychological process, if any, the word "generalization" refers to. An alternative approach is to treat the experiment as dealing simply with visual pattern recognition. I shall briefly present some simulated results based on the theory described in Sheila Chase's contribution to this volume. According to that theory, neither the cube nor its distortions are experienced by the pigeons as three-dimensional forms. All the simulations to be discussed were done by Stevens (1990). The plan is to compare the performance of the simulation program to the performance of actual pigeons. Before doing the simulation it is necessary to determine how the correlation between simulated and empirical values varies with a parameter we call "spacing." That parameter represents the distance between pixels of the form that is imaged on the retina-as represented in the internal representation of that form. It is simply a number by which all coordinates of the form are multiplied. Fig. 3 shows how the correlation between simulated generalization data and the actual generalization data for the two pigeons in Group 1, and the mean of these values, varies with spacing. Based on the mean value, the best correlation is obtained at a spacing of 1.5. It is
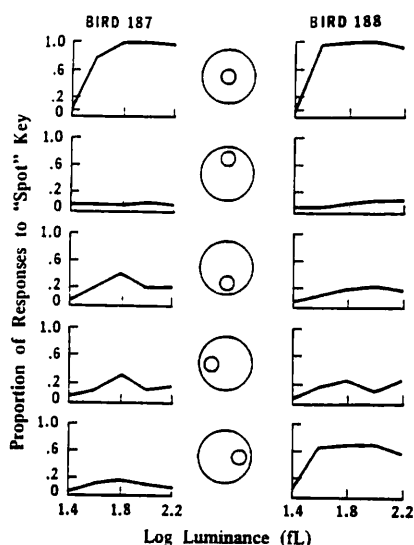
Fig. 4. Choice curves obtained after training to discriminate between a circular field containing a centrally located spot of light and an evenly illuminated field. The central figure in each row shows the position of the spot, but the relative sizes of the spot and background field are not drawn to scale. From Heinemann and Kadison (1976).

this spacing that is used in all simulations. The function shown in Fig. 3 is virtually identical to one published by Heinemann and Chase (1990) when they compared simulated confusion matrices for letters and random dot patterns to confusion matrices for pigeons obtained by D. Blough (1985).

Fig. 2 (right panel) shows simulated generalization curves for the three groups of pigeons. The simulation program categorized the cube stimuli more accurately than the pigeons and distortions less accurately. The best correlations between the mean proportion of correct responses made by the pigeons and by the simulation program (Pearson r) are 0.60 for Group 1, 0.61 for Group 2, and 0.44 for the single pigeon in Group 3. These correlations are lower than corresponding correlations between the groups of pigeons. Stevens suggests that this discrepancy may be a consequence of the fact that his simulation program assumed that performance was essentially perfect at the end of acquisition.

This was not true for the real pigeons.

Effects on recognition of shifting the retinal image to new location

The effects of shifting (translating) an object to a physical location different from the one in which the pigeon was trained to recognize have been examined in several experiments done in our laboratories. Fig. 4 shows the results of a generalization test given to pigeons who had been trained to peck on one response key when presented with a small circular spot of light projected on a larger evenly illuminated disk, and to peck on a different response key when presented with the large disk without the added spot (Heinemann and Kadison, 1976). The visual display was located between the two response keys. The luminance and position of the spot were varied during the generalization test. Fig. 4 shows that the discrimination between presence and absence of the spot was essentially perfect when the luminance of the spot was high and the spot was in the center of the disk. For most of the other test positions, however, the birds behaved almost as though the spot did not exist.

We have obtained essentially similar results in several experiments on pattern recognition. In one of these, three pigeons were trained to identify each of four visual shapes by pecking on one of four different response keys. During training the single pattern shown on each trial (an outline figure consisting of black pixels on an otherwise white screen) was always presented near the center of the 6 cm square face of a small monochrome TV monitor. During a "translation test" that followed 130 days of training (100 trials per day), probe trials were interspersed among training trials. On each probe trial one of the four shapes was presented 3 cm above and to the right of the training position. During this test the birds averaged 72% correct when the stimuli were in their training position, but only 31% correct when the stimuli were in the new position. Two of the birds actually performed at chance (25% correct when the stimuli were in the new position.) The position change studied here was quite large—about 32 degrees of visual angle.

Several other unpublished experiments done

in our laboratories have yielded similar results. Even small changes in the position of visual patterns cause recognition of these patterns to deteriorate. However, it is important to note that these effects occur when the pattern is always in the same position during training. If the pattern is moved about during training then the pigeons learn to search for it (e.g. D. Blough, 1979; P. Blough, 1984; Jenkins & Sainsbury, 1969).

The effect of shifting the *retinal image* from a single position it occupied in training was investigated in an experiment by Yamashita (1991). Before discussing this work it is necessary to point out that my discussion is focused entirely on learning and behavioral processes, whereas Yamashita's principal focus was on sensory processes. With respect to the behavioral results, the general outcome of Yamashita's experiments was that pigeons did not recognize the training stimuli when these were first shown in new positions. However, after extended training, during which the stimuli were shown in a variety of new positions, the pigeons recognized them equally well in all these positions. The discussion that follows concerns the nature of the processes that may underlie this acquired "translational invariance."

The training procedures used in Yamashita's experiments cannot be presented in detail here, but it is necessary to describe the essential features.

After preliminary training, the pigeons were taught to peck on a fixation point (a 3-mm square located near the center of a monochrome monitor on which the experimentally manipulated stimuli were displayed). This training was followed by two phases of discrimination training. For convenience, Yamashita called the first phase "detection" and the second "discrimination." Both phases consisted of a series of positive and negative trials. Each trial began with the presentation of the fixation point. The first peck that landed on the fixation point (an "effective response") initiated a 5-sec "sampling interval". During positive trials each effective response was followed, after a delay of 100 ms, by a 50 ms exposure of a group of eight illuminated pixels in the form of an "x". If at least two
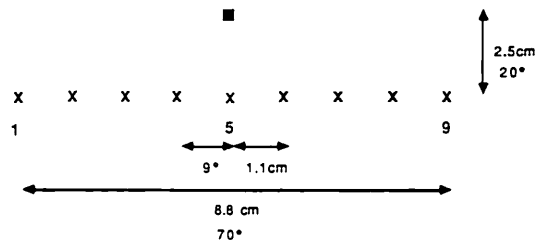

Fig. 5. Schematic illustration of stimulus positions. Estimates of visual angles are based on the assumption that the viewing distance is 6.7 cm. From Yamashita (1991) - reproduced with permission.

effective responses were made during the sampling interval, then the first effective response that was made after the end of the sampling interval was reinforced. Responses were not reinforced on negative trials. It is important to note that for all data analyses, the number of pecks on the fixation point (effective responses) is taken as the measure of response strength.

Whereas the stimuli present during sampling intervals of positive trials were identical in the detection and discrimination procedures, those present during negative trials differed. During the detection procedure effective responses were not followed by a stimulus presentation on negative trials; the area of the display where the "x" could appear remained blank. Thus, the subjects were trained to discriminate between the presence and absence of the positive stimulus.

During "discrimination training", the stimulus presented during the sampling intervals of negative trials consisted of eight illuminated pixels in the form of an "o".

Fig. 5 is a schematic illustration of the positions in which stimuli could be presented, together with estimates of distances in degrees of visual angle. The numbers 1 to 9 will be used to refer to these stimulus-positions. As mentioned, the principal focus of my discussion will be the effect that a shift in position of a trained stimulus has on recognition. However, to interpret data from shift experiments it is necessary
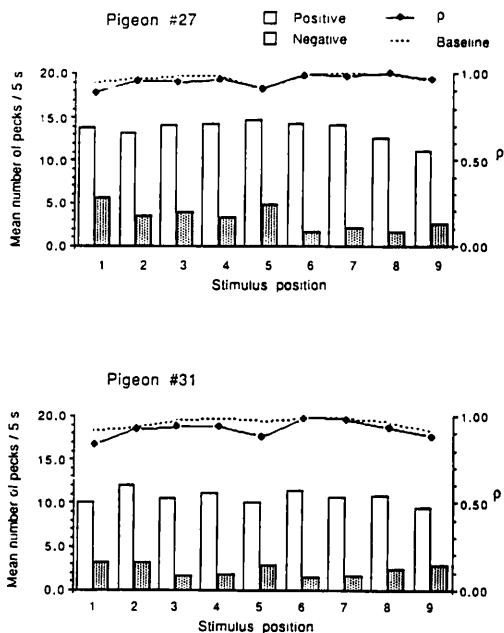
Fig. 6. Mean number of pecks per 5-s sample interval in positive (empty columns) and negative trials (filled columns) and rho (solid line) for the last three sessions with the 1 ms display of the stimulus as a function of stimulus positions. Average rho (dotted line) for the last three sessions of the preceding training with the 150 ms display (baseline) is also shown for comparison. From Yamashita (1991) - reproduced with permission.
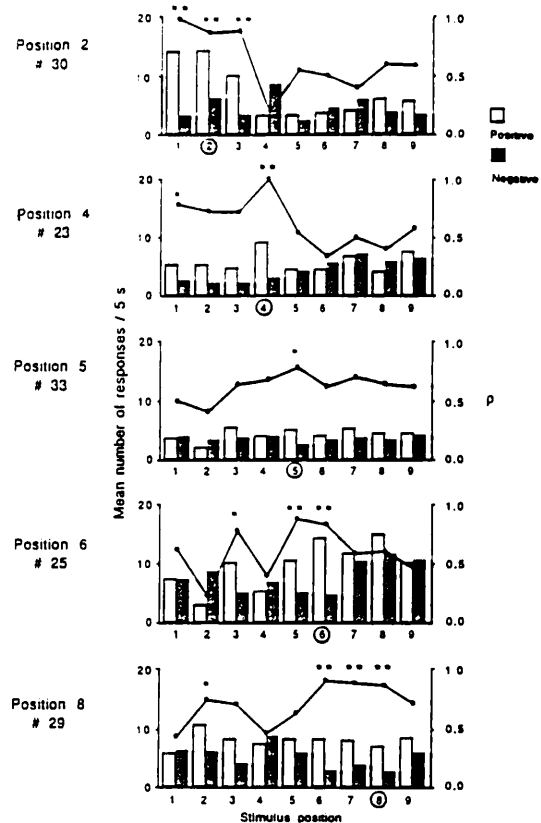

Fig. 7. Mean numbers of pecks per 5-s sample interval in test trials with the positive stimuli (empty columns) and the negative stimuli (filled columns) and rho (line) as functions of stimulus positions for each subject. Statistical significance of rho is indicated by asterisks:
* p<.05, ** p<.01. From Yamashita (1991) -reproduced with permission.

to examine the effects of imaging the stimuli on different portions of the retina when no shift is involved. Yamashita did this in his Experiment 1 in which the subjects received an equal number of training trials at all nine positions. The method was that previously described, with one important addition. Because the interpretation of the results obtained in all of Yamashita's experiments depends critically on the assumption that the pigeon continued to fixate the "fixation point" during exposure of the positive and negative stimuli, this exposure was reduced to 1 ms after the regular training sessions were completed. That duration corresponded approximately to a single refresh cycle of the computer display. Training with this short stimulus

duration continued for five more sessions. Yamashita checked on maintained fixation in a number of other ways, but these will not be considered here.

Three dependent variables were used in the analysis of the results. The following definitions are taken from Yamashita.

One of the dependent variables is Herrnstein's rho defined as

$$rho = 1 - U/(N_p x N_n)$$

where U is the Mann-Whitney U statistic, which is computed by ranking the numbers of responses in positive and negative trials. N_p and

N_n are the numbers of positive and negative trials, respectively. A rho equal to 1 indicates a perfect discrimination, a rho equal to 0.5 indicates chance performance.

Another dependent variable is the discrimination ratio, defined as the proportion of responses made in positive trials to the total number of responses. It was calculated using the average number of responses on each type of trial in each session.

A third dependent variable was the mean number of pecks per 5-sec sample interval.

The results of this experiment are shown in Figure 6. There appears to be little if any systematic effect of position upon recognition, and performances with the regular and short duration stimuli differ very little.

In the main experiments to be considered pigeons were trained with stimuli that were presented in only one of the 9 positions. After performance in each of the two phases showed little improvement with further training a translation test was done during which the pigeons were shown the stimuli in each of the nine positions.

The results are shown in Fig. 7. Visual inspection of the figure and statistical analyses support Yamashita's statement "It may be concluded that the transfer is limited to the vicinity of the position in which stimuli were shown during training." (p 57)

Some of the acquisition data reported by Yamashita make a further analysis possible. According to the Heinemann-Chase theory, there are two distinct phases during acquisition of a discrimination. The first of these is called the Presolution Period (PSP). It is a period, at the beginning of training, during which there is no evidence of a developing discrimination, i.e., performance remains at a chance level. The duration of this period (or number of trials) increases systematically with the physical similarity of the stimuli presented for discrimination. It is assumed that during the PSP the subject searches for significant statistical associations between the information arriving over specific sensory channels, the response made, and the consequences of that response (e.g. reinforcement or nonreinforcement). The formal model (Heinemann, 1983; Heinemann & Chase,
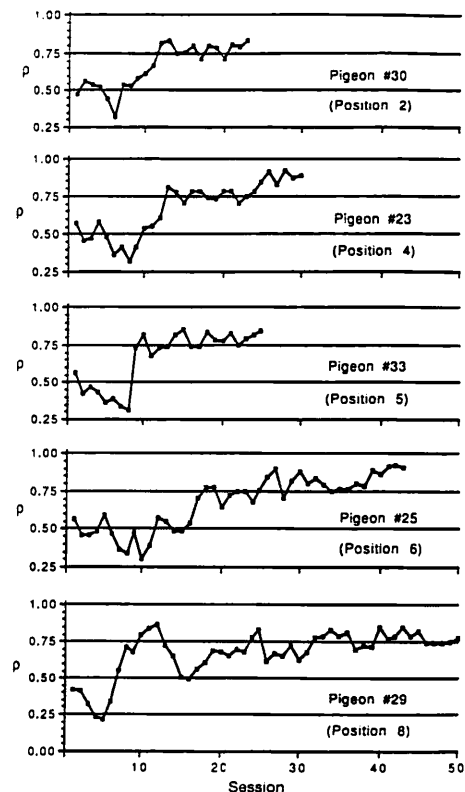

Fig. 8. Learning curves in terms of rho index for the detection training. From Yamashita (1991) - reproduced with permission.

1990) is an application of the Sequential Probability Ratio Test developed by Wald and his associates (Wald, 1947). A successful outcome of the analysis carried out during the PSP is that the subject discovers a sensory channel or channels that predict outcomes of behavior, at a specified level of significance. The PSP ends, and a second phase of learning begins, when the specified level of significance is attained.

Fig. 8 shows acquisition curves for detection training of the pigeons in Yamashita's Experiment 2b. These curves seem to show presolution periods in the range of 7 to 10 sessions. All but two of the birds on discrimination training also show PSPs. Interestingly, when the pigeons who served in Experiment 2b were later retrained with the x as S+ and the o as S-, now presented equally often in each of the nine positions, they showed PSP's at those new positions

that were reasonably far removed from the original training position. This strongly suggests that what is learned during the PSP is specific to a retinal position-which is related to our concept of "sensory channel". To be a little more specific: we cannot describe the entire channel in anatomical and physiological terms, but it is known that the first level at which information about visual stimuli appears for transmission to higher regions of the brain is a retinal ganglion cell that has a receptive field with mutually antagonistic centers and surrounds, or possibly a cluster of such cells.

In any case, our present theory is that recognition of trained spatial patterns that are shifted to new sensory channels is invariant only over those channels for which the PSP has been completed.

Of the shift experiments that have been discussed, that of Yamashita was the only one in which an attempt was made to control fixation. The other experiments specify the shift of the physical image on the display surface, not the shift of the retinal image. It is not yet clear how the two sets of experiments may be related. Perhaps the pigeons in the experiments in which no explicit effort was made to control fixation were, nonetheless, fixating a certain feature of the visual displays. This could possibly come about through the "feature positive effect" first described by Jenkins and Sainsbury (1969; 1970) Briefly, these investigators reported that if a distinctive feature, such as the spot in the experiment of Heinemann and Kadison (1976), appears on positive trials of a discrete trial operant discrimination, the pigeons pecked directly at that feature. If the feature is small, this would constitute a kind of fixation. Additional research on this matter might well yield results of great importance to our understanding of visual pattern recognition.

Conclusions

The experimental findings discussed in this paper indicate that pigeons perceive stationary two-dimensional projections of three-dimensional objects as two dimensional. With respect to the effects of translating the retinal image of a known object over the retina, the experiments discussed indicate that translational invariance

is a result of a learning process. Both of the above mentioned phenomena are predicted by the theory of Heinemann and Chase.

REFERENCES

Blough, D. S. (1979) Effects of the number and form of stimuli on visual search in the pigeon. *Journal of Experimental Psychology: Animal Behavior Processes, 5,* 211-223.

Blough, D. S. (1985) Discrimination of letters and random dot patterns by pigeons and humans. *Journal of Experimental Psychology: Animal Behavior Processes, 35,* 344-352.

Blough, P. M. (1984) Visual search in pigeons: Effects of memory set size and display variables. *Perception & Psychophysics, 35,* 344-352.

Cerella, J. (1990) Shape constancy in the pigeon: The perspective transformations decomposed. In M. L. Commons, R. J. Herrnstein, S. Kosslyn and D. B. Mumford (Eds.), *Quantitative Analyses of Behavior: Behavioral Approaches to Pattern Recognition and Concept Formation* (Vol. 8, pp. 145-163), Hillsdale, NJ: Erlbaum.

Chase, S. (1983) Pigeons and the magical number seven. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (EDs), *Quantitative Analyses of Behavior: Discrimination Processes* (Vol. IV, pp. 37-57), Cambridge, MA: Ballinger.

Durlach, N. I., & Braida, L. D. (1969) Intensity perception I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America, 46,* 372-383.

Goodale, M. A. (1983) Visually guided pecking in pigeons (Columba livia). *Brain, Behavior and Evolution, 22,* 22-41.

Heinemann, E. G. (1983) The presolution period and the detection of statistical associations. In M. L. Commons, R. J. Herrnstein and A. R. Wagner (Eds.) *Quantitative Analyses of Behavior: Discrimination Processes* (Vol IV, pp. 21-35). Cambridge, MA: Ballinger.

Heinemann, E. G., & Chase, S. (1990) A quantitative model for pattern recognition. In M. L. Commons, R. J. Herrnstein, S. Kosslyn and D. Mumford (Eds.), *Quantitative Analyses of Behavior: Computational and Clinical Approaches to Pattern Recognition and Concept Formation.* (Vol 9, pp. 109-126). Hillsdale, NJ: Erlbaum.

Heinemann, E. G., & Kadison, K. (1976) Control of pigeons' choice behavior by the position and luminance of a spot of light. *Bulletin of the Psychonomic Society, 7,* 522-524.

Jenkins, H. M., & Sainsbury, R. S. (1969) The development of stimulus control through differential

reinforcement. In N. J. Mackintosh and W. K. Honig (Eds.), *Fundamental Issues in Associate Learning*. Halifax: Dalhousie University Press, 123-161.

Jenkins, H. M., & Sainsbury, R. S. (1969) Discrimination learning with the distinctive feature on positive or negative trials. In D. Mostofsky (Ed.), *Attention: Contemporary Theory and Analysis*. New York: Appleton-Century-Crofts, 239-273.

Miller, G. A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Pollack, I. (1956) Identification and discrimination of components of elementary auditory displays. *Journal of the Acoustical Society of America, 28*, 906-909.

Sperling, G. (1988) The magical number seven: Information processing then and now. In William Hirst (Ed.), *The Making of Cognitive Science: Essays in Honor of George A. Miller*. Cambridge, U.K.: Cambridge University Press, 71-80.

Stevens, R. (1990) An investigation of the concept of cube in the pigeon. Doctoral dissertation. City University of New York.

Wald, A. (1947) *Sequential Analysis*. New York: Dover Publications, Inc.

Yamashita, H. (1991) Frontal visual field position and visual discrimination learning in the pigeon (*Columba livia*). Doctoral dissertation. City University of New York.