

From M. L. Commons, R.J. Herrnstein, & A.R. Wagner (Eds.). 1983. *Quantitative Analyses of Behavior: Discrimination Processes*. Cambridge, MA: Ballinger (Vol. IV, Chapter 1, pages 3-19).

## **A Memory Model for Decision Processes in Pigeons**

*Eric G. Heinemann*

Brooklyn College of the City University of New York

This chapter presents a model for the acquisition of discriminative behavior by pigeons together with a variety of research results that bear on this model. The model considered here is for the second stage of an acquisition process. The first stage of this acquisition process is the presolution period (PSP). During the PSP the subject detects the statistical association that the experimenter (or nature) has arranged among (1) the stimuli presented for discrimination, (2) the stimulation associated with the subject's behaviors, and (3) the "outcomes" that follow these behaviors. A model for the processes occurring during the PSP is presented in Chapter 2 of this text (Heinemann, 1983). The decision processes assumed for the PSP are those of an optimal detector of statistical associations and require a perfect, but very simple memory. The subjects need only store the cumulative sum of the observations made (Wald, 1947). The theory presented in the present chapter assumes that the processes occurring after the end of the PSP involve a memory that has a limited storage capacity. Records of events that have occurred on individual trials are preserved in this memory.

The principal assumptions of the model concern (1) the storage and representation of events in the subject's memory, (2) retrieval of stored items, and (3) response selection. Each of these subjects will be treated in one of the following sections.

### **Storage and the Representation of Events in Memory**

Information gathered on each trial of an experiment is assumed to be placed in a relatively long-term memory, that we refer to here as limited-capacity memory (LCM). Each record placed in this memory during the course of discrimination training contains information concerning the *discriminative stimuli* that analysis during the PSP has shown to be predictive of the outcomes of behavior, the *responses* made, and the *reward* received. One such record is entered into LCM on every trial and is said to occupy a "storage location." The LCM is assumed to have a fixed number of storage locations and the location to which each record is sent is selected randomly.<sup>1</sup> Any record occupying a storage location will be destroyed ("overwritten") when a new record is entered in that location.

Most of the experiments considered in this chapter involve the presentation of two stimuli from a simple continuum such as sound intensity. It is assumed that, over repeated presentations, these stimuli induce sensory effects that are normally distributed with different means,  $\mu_1$  and  $\mu_2$ , and a common standard deviation  $\sigma$ . The sensory effect experienced on each trial is stored in the LCM, along with information about the nature

of the continuum. That is, the subject remembers not only the intensity of the stimulus, but also whether it was a sound or a light, etc.

The responses made just before receipt of reward are assumed to be represented in memory in the form of sensory information associated with these responses, such as the visual characteristics of the key that was pecked: its color, position, etc. Some important consequences of this way of representing information concerning the response are developed by in Chapter 3 of this book (Chase, 1983). In the present chapter, however, the responses will be treated as though they were represented in memory simply by the labels  $R_1$ , and  $R_2$ . It is possible to show that, for the experimental conditions to be considered in this chapter, this treatment is equivalent to that derivable from the more general model considered by Chase.

Information concerning the response outcomes (reward or non-reward, etc.) is treated in a comparable fashion. The internal effects of rewards delivered repeatedly are assumed to be normally distributed, and the sensory effect of the reward delivered on each trial is entered on each trial record. As with the discriminative stimuli, the sensory effects of reward are represented by normal deviates in the formal treatment. However, since in the applications to be considered in this chapter, the amount and kind of reward are not varied, the outcomes may be treated as though they were represented in memory simply as reward (positive) or non-reward (negative).

### **Retrieval**

It is assumed that during each trial the subject draws a small sample of records from the LCM. The choice of response is based solely on the information contained in this sample of records. The samples are assumed to be independent random samples with the following restriction: the subject is assumed to draw records from the LCM one at a time until a fixed number of positive records (records showing that reward was received) has been obtained.<sup>2</sup>

To say that the subject draws a sample of records does not imply that the records in questions are removed from the LCM. The idea is that these records are “looked at” or are “copied,” so to speak, for use in a working memory.

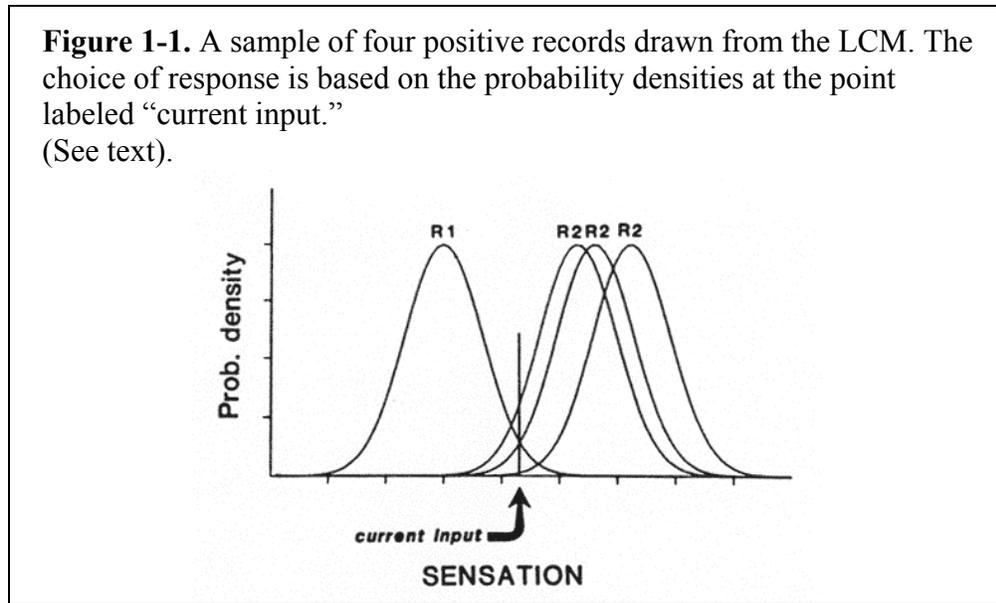
### **Response Selection**

The choice of response on each trial involves not only a sample of records from memory, of course, but also the sensory effect present on the continuum that the analysis done during the PSP has shown to be predictive of outcomes. An example would be the currently experienced “loudness” of white noise. This sensory effect will be called the “current input.”

Before describing how the current input and the records drawn from the LCM are used in response selection, it is necessary to state one further assumption. The sensory effect represented on each record drawn from the LCM is assumed to fluctuate rapidly over time, momentary values being described by a normal distribution, the mean of which represents the sensory effect experienced on the trial on which the record was formed.

The response selected by the subject is the one most likely to be rewarded on the basis of the evidence contained in the sample of records drawn from the LCM. To find it,

the subject considers the records containing sensory effects in the presence of which  $R_1$  was rewarded and sums the probability densities at the current input value. The subject then performs the same operations for the records with the sensory effects for which  $R_2$  was rewarded. The response selected is the one with the larger total probability density. The process is illustrated in Figure 1-1, which shows the sensory-effect distributions represented on four records, three of which show that  $R_2$  was rewarded while one shows that  $R_1$  was rewarded. The process amounts to summing the heights of the  $R_2$  curves above the point representing the current input, doing the same for the  $R_1$  curves, and then determining which sum is the larger.



If the subject were in the presolution period or in an experiment on probability learning, the records drawn from the LCM would contain no information concerning discriminative stimuli. Given a sample of this sort, the subject simply chooses the response with the greater probability of reward. The rule that accomplishes this is a simple one: If the sample contains more records labeled  $R_2$  than  $R_1$ , make response  $R_2$ ; otherwise, make response  $R_1$ .

The model described here, together with the model for the PSP described in Chapter 2 (Heinemann, 1983), has been written as a computer program that simulates results to be expected in a variety of experimental situations.

## RESULTS OF EXPERIMENTS AND SIMULATIONS

A number of experiments designed to check on predictions of the LCM model, to evaluate its parameters, and to find out whether parameter values can be recovered across experiments will be presented in the sections that follow. All of these experiments were done with pigeons and used the same basic method.

Training chambers contain three response keys located in a horizontal row at the approximate height of the pigeons' heads. Daily training sessions consist of eighty trials, separated from each other by intervals of ten seconds. At the start of a trial, the center key

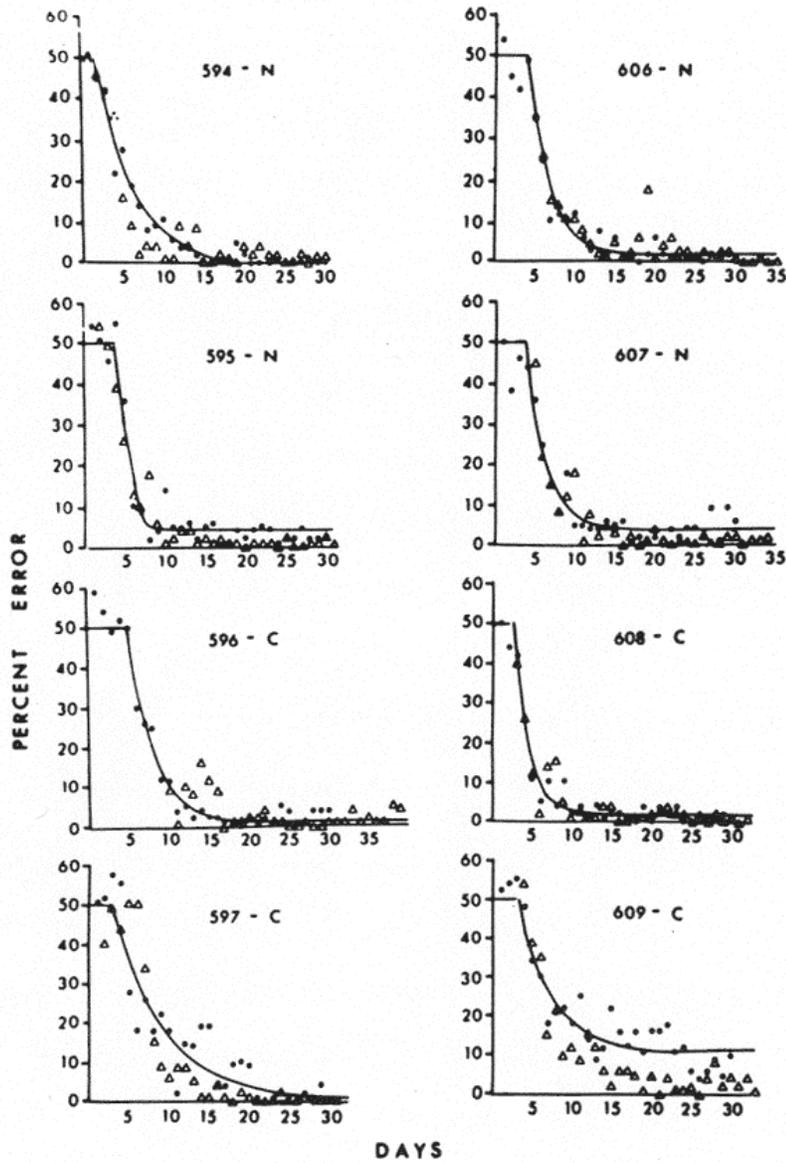
becomes illuminated, and a subsequent peck on the center key activates a stimulus. Acoustic stimuli are delivered through a loudspeaker located in the response panel, whereas visual stimuli are presented on the center key. The center-key peck also causes the two side keys to become illuminated. A single peck on one of the illuminated side keys is considered a correct response if it occurs in the presence of one of the two discriminative stimuli, and a single peck on the other side key is considered correct for the other stimulus. Correct responses are rewarded by a 2.5-second period of access to food in a tray. Incorrect responses are followed after ten seconds by a repetition of the trial.<sup>3</sup>

### **Retroactive Interference**

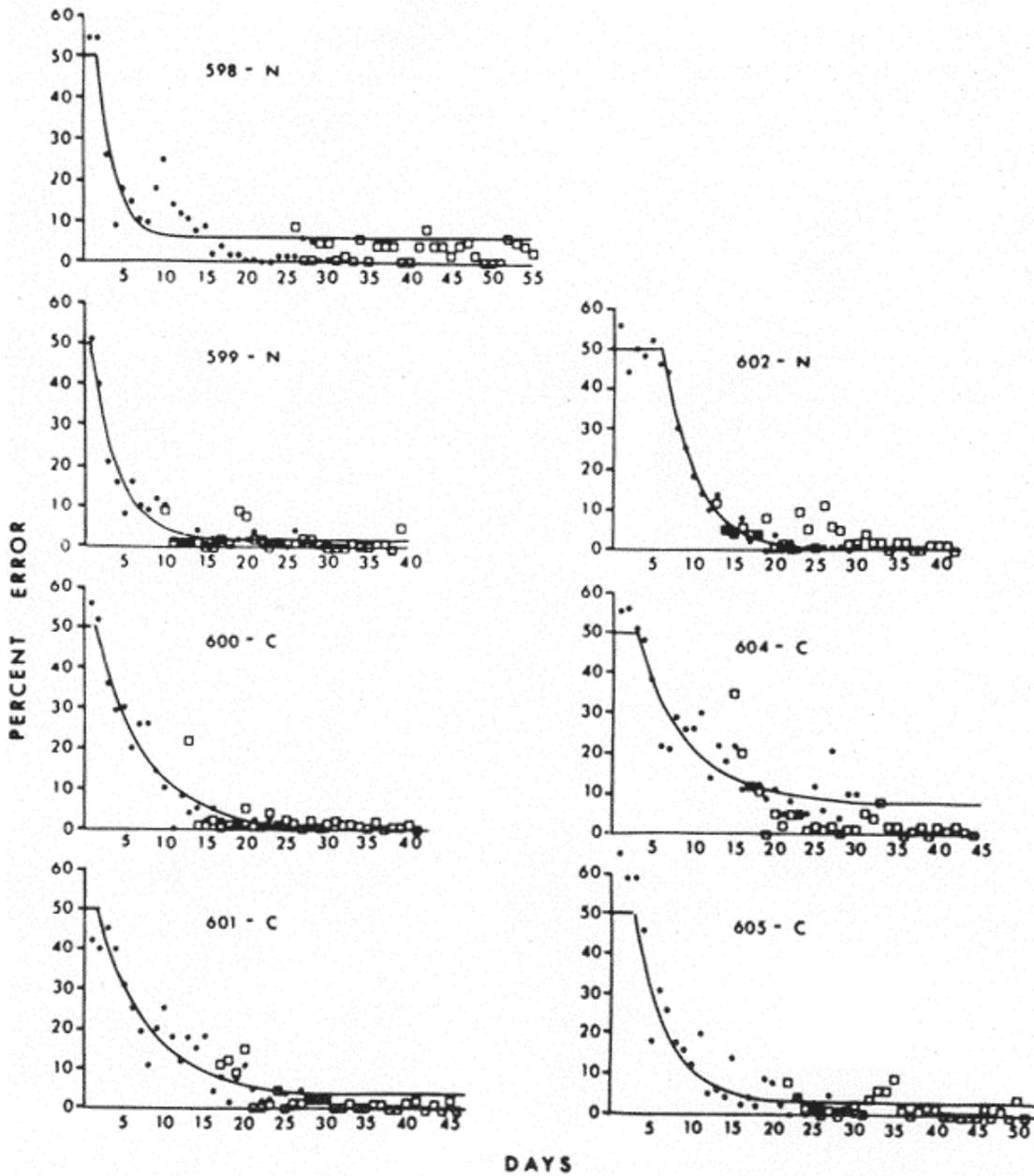
The assumption that new information placed in the LCM destroys information already stored leads to the prediction that prolonged training in any task should result in complete forgetting of previously learned discriminations. Several experiments have been done to check on this matter. In the first of these (Heinemann, Sage-Day, and Brenner, 1981), one group of pigeons was initially trained to discriminate between an orange and a red light presented on the center key, while a second was trained to discriminate between two levels of white noise differing by 30 dB. After the completion of this first phase of training, both groups were shifted to a probability learning task ( $\pi = 0.5$ ) for approximately 40,500 trials spread over forty-seven days. In the final stage of the experiment, the birds were retrained on the original discrimination. Control groups received the same initial and final stages of training as the experimental groups, but remained in the colony room during the time the experimental groups were engaged in the probability learning task.

The course of acquisition and reacquisition is shown in Figure 1-2a for the experimental groups and in Figure 1-2b for the control groups. It can be seen that six of the eight birds in the experimental groups performed essentially at chance levels at the beginning of the reacquisition period and then improved in a fashion that does not appear to differ systematically from the course of original acquisition. These subjects seem to have forgotten completely what response to make in the presence of which stimulus, but analysis of the data suggests that they did not pass through a presolution period during reacquisition. Although some of the subjects in the control groups also show some forgetting, several statistical tests showed the differences between the experimental and control groups to be highly significant.

**Figure 1-2a.** The small, filled circles represent the original acquisition of a discrimination between two intensities of white noise (N) or between two lights differing in color (C). The open triangles represent reacquisition. They were moved parallel to the horizontal axis so as to superimpose them on the smooth curve fitted to the original acquisition data. From Heinemann, Sage-Day, and Brenner (1981).



**Figure 1-2b.** Acquisition data (small, filled circles) and reacquisition data (open rectangles) for the subjects in the control groups. From Heinemann, Sage-Day, and Brenner (1981).



The basic experiment described above was repeated and extended by Brenner (unpublished). Brenner's pigeons were trained to discriminate between orange and red lights in the initial and final phases of the retroactive-interference design. In addition to the control group, there were three experimental groups. The training interpolated between acquisition and reacquisition consisted of probability learning for one of the experimental groups (a replication of the previous experiment), a discrimination between

two luminance levels of white light differing by 2 log-units for the second experimental group, and a discrimination between two levels of white noise differing by 30 dB for the third group. The results obtained for the three experimental groups did not differ significantly from each other and are virtually identical to those shown in Figure 1-2a. In other words, these pigeons also showed essentially complete forgetting of the relation between the values of the discriminative stimuli and the responses to be made.

The results suggest that “overwriting” may be indiscriminate in the sense that the destruction of old information by new may not depend on the relation between the old and new information. Such a state of affairs would be surprising in view of the well known effects of task similarity found in retroactive interference experiments done with human subjects (see Osgood, 1949). However, the research considered above shows only that various interpolated activities produced equal degrees of forgetting after a fixed number of trials. It has not been shown that the various interpolated activities cause previously learned material to be forgotten at identical rates.

### **Probability Learning**

One of the situations to which the LCM model has been applied is probability learning. In experiments on probability learning, discriminative stimuli are not available to the subject. Reward is scheduled with probability  $\pi$  for one of the responses,  $R_1$ , and with probability  $(1 - \pi)$  for the alternative response,  $R_2$ . The results obtained depend strongly on details of the procedure.

In experiments in which trials are scheduled independently of the subjects’ behavior (“non-correction” procedures), subjects usually end up choosing one of the alternatives exclusively. In terms of the LCM model this “absorption” is the result of a positive feedback system that involves the contents of the LCM and the response-selection process. The more often the subject is rewarded for one of the choices, say  $R_2$ , the greater will be the number of positive  $R_2$  records in the LCM. The more positive  $R_2$  records there are in the memory (in comparison to  $R_1$ , records), the more likely it is that response  $R_2$  will be made.

In the “correction” method used in most of the experiments described in this chapter, trials on which the subject does not collect the scheduled reward are repeated, so all scheduled rewards are held until collected. As a result, the average proportions of positive  $R_1$  and  $R_2$  records will come to equal  $\pi$  and  $(1 - \pi)$  after sufficient training.

The asymptotic distribution of response choices depends only on the size of the sample of records the subject retrieves on each trial. The symbol  $\theta$  will be used to denote this sample size. If  $\theta = 1$ , then the asymptotic response proportions  $p(R_1)$  and  $p(R_2)$  will match  $\pi$  and  $(1 - \pi)$ , respectively. The reason for this is simply that the subject is drawing single records randomly from a pool in which  $R_1$  and  $R_2$  records are presented in proportions  $\pi$  and  $(1 - \pi)$ .

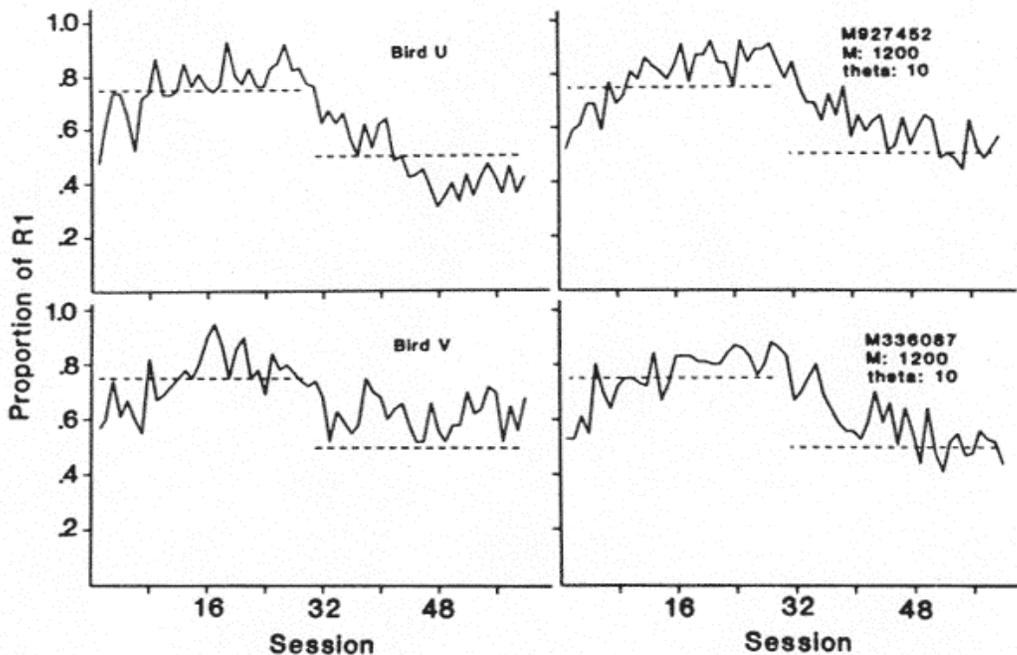
On the basis of a large number of experiments done in my laboratory, I have estimated that the LCM has approximately 1200 storage locations. If the subject were to examine the whole memory on each trial ( $\theta = 1200$ ), and if  $\pi$  were not equal to or very close to 0.5, then a well trained subject would always choose the more frequently rewarded response. As  $\theta$  takes on values progressively larger than 1.0, the subject should “overshoot” the matching level more and more. Thus the asymptotic level of performance

observed in a probability-learning experiment with a correction procedure provides an estimate of the size of the sample.

The rate of acquisition in a probability-learning experiment (as well as in a discrimination experiment) is primarily determined by the size of the LCM. The smaller the size of the memory, the faster the acquisition. Sample size has only a minor effect.

Figure 1-3 shows results obtained in the laboratory for two birds trained with the correction method (80 trials per day). The value of  $\pi$  was 0.75 during the first 30 daily sessions and was then shifted to 0.5. Also shown are the results of two simulated “birds” given identical training. The values of memory size  $M$ , and sample size  $\theta$ , for which the simulation program produced results judged to be comparable to those of the real birds shown in Figure 1-3, were determined by trial and error. As shown in Figure 1-3, these values were  $M = 1200$  and  $\theta = 10$ . The results obtained for other pigeons that participated in the probability-learning experiment could also be nicely approximated by simulations based on a memory size of 1200 locations, but the requisite value of  $\theta$  varied among birds, the extremes being  $\theta = 3$  and  $\theta = 12$ .

**Figure 1-3.** Performance of two pigeons (U and V), and two simulated pigeons in probability learning task. On day 30 the value of  $\pi$  was shifted from 0.75 to 0.50. The dashed horizontal lines indicate the “matching levels.”



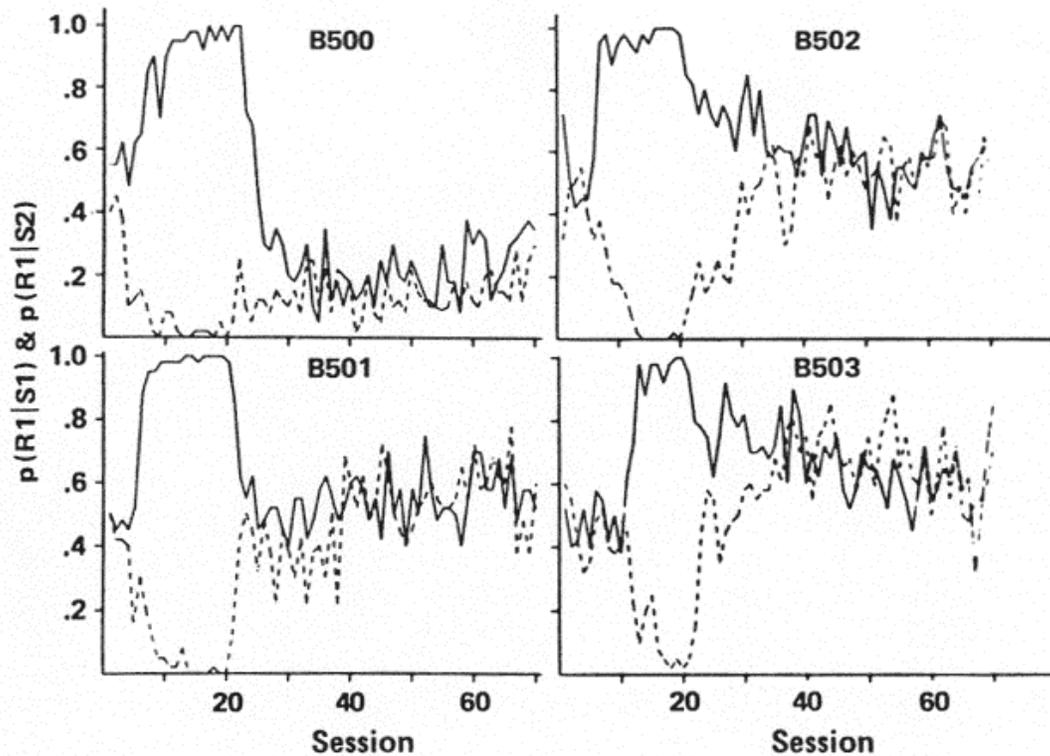
### Discrimination Learning

*Effects of Changing Contingencies.* The experiments described in this section were done in collaboration with Julie Sage-Day. In the first experiment, pigeons were trained to discriminate between two levels of white noise differing by 30 dB. After twenty days of training, the relation between the two stimuli and the response outcomes was changed

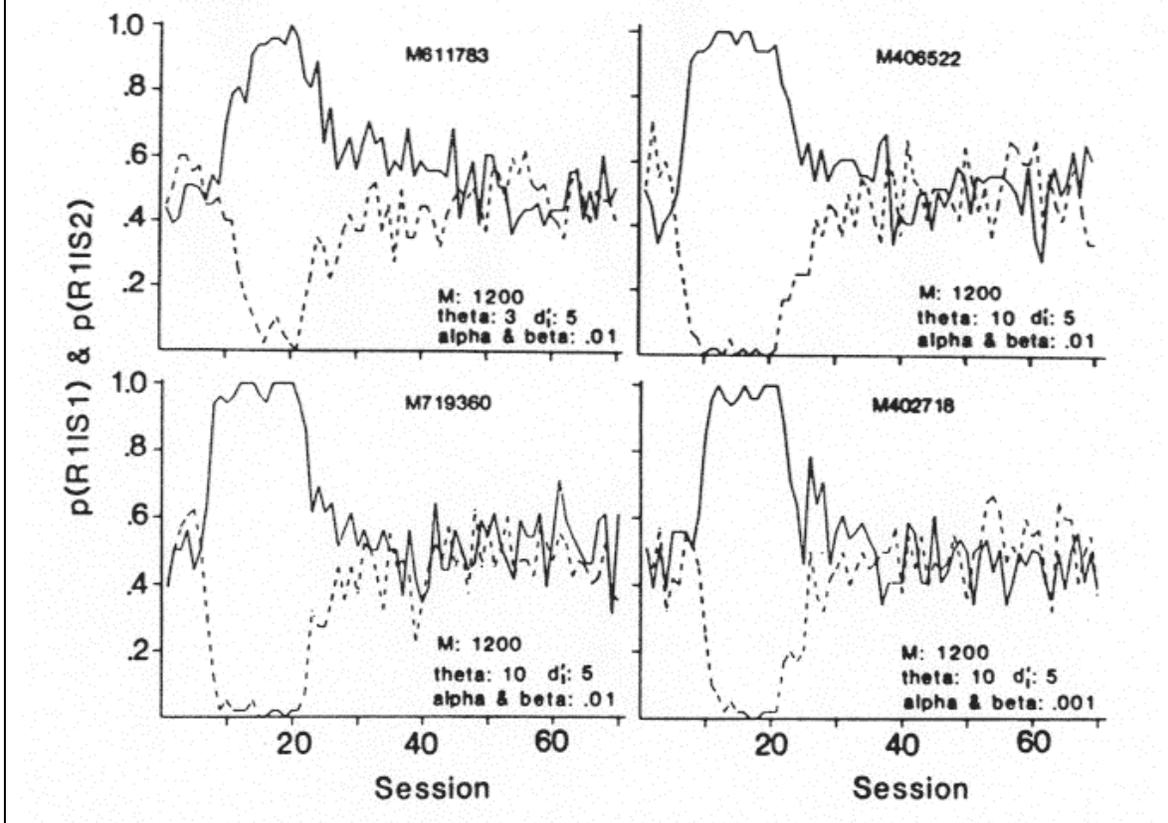
to a non-predictive one. Specifically, from day twenty-one on, the pigeons were rewarded equally often for making response  $R_1$  and  $R_2$  in the presence of each of the stimuli.

The results for four individual birds are shown in Figure 1-4. Figure 1-5 shows the results of four computer simulations. In this and in all other simulations reported in this chapter, memory size is assumed to be 1200 locations. M719360 and M406522 represent two runs of the computer program that differ only in the “seed” provided to the random number generator. M611783 differs from M719360 and M406522 also in the value of  $\theta$ , and M402718 differs also in the values assumed for  $\alpha$  and  $\beta$ , the parameters that define the strength of the sequential probability-ratio test that the subject is assumed to perform during the PSP.

**Figure 1-4.** Results of 20 days of training of subjects to discriminate between two levels of noise differing by 30 dB followed by abolition of the discrimination. The solid lines represent the proportions of  $R_1$  responses made in the presence of  $S_1$  and the dashed lines the proportion of  $R_1$  responses made in the presence of  $S_2$ . Each panel shows the results for a single pigeon.



**Figure 1-5.** Results of four simulations of “birds” trained under the conditions described for Figure 1 - 4. (See text).



When simulating discrimination training, it is necessary to specify the discriminability of  $S_1$  and  $S_2$  in terms of  $d'_1 = (\mu_1 - \mu_2)/\sigma$ , where all symbols have the meanings previously assigned. This value of  $d'_1$  represents the discriminability of the sensory effects and is somewhat larger than the value of  $d'$  that can be estimated from the asymptotic response proportions of a simulation of bird behavior under the given conditions. The reason for this discrepancy is that the sampling- and response-selection processes add some variance to the sensory-effect variance.

Figure 1-6 shows the mean results for nine birds that were in the experiment just described (points) together with the mean results for nine simulated birds. Two points should be kept in mind when considering this and other comparisons of real and simulated data. First, the parameters used in the simulation were determined by trial and error. The curve showing the results of the simulations represents an approximation, not a best fit. Second, the simulations that were averaged differed only in the seeds provided to the random-number generator, not in the values of  $M$ ,  $\theta$ , etc., whereas data from real birds show individual differences in some parameters.

The principal systematic difference between the real birds and the simulated ones seems to be that at least some of the real birds have strong key biases, as shown by the fact that after thirty days, the majority of the empirical points lie above 0.5 on the vertical axis in Figure 1-6. Key biases of this sort also distinguish real birds from simulated ones in a variety of other situations. It is possible that these biases arise from some

uncontrolled asymmetry in the experimental situations. For example, it might be that, because a particular bird stands in a given position, it obtains more food when approaching the food hopper from the left than from the right.

**Figure 1-6.** Mean proportion of  $R_1$  responses made in the presence of  $S_1$  and  $S_2$  by nine simulated birds (lines) and nine real birds (data points).

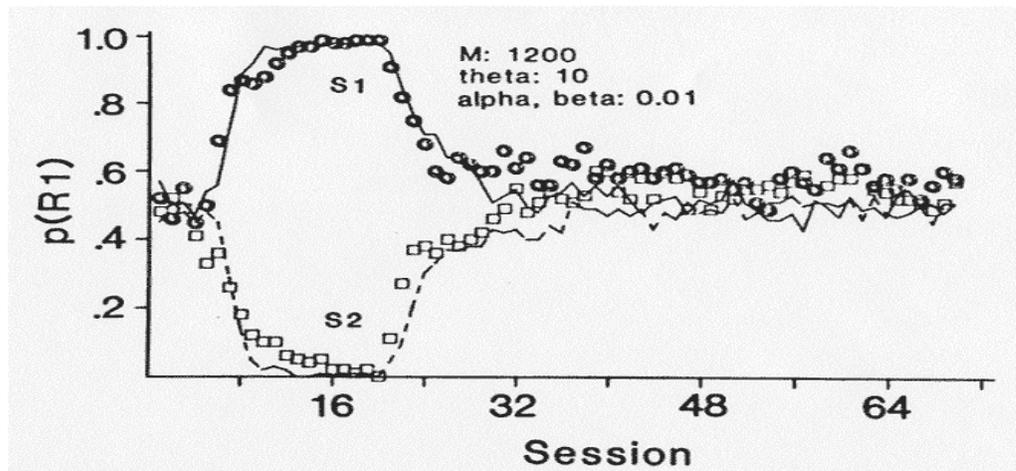
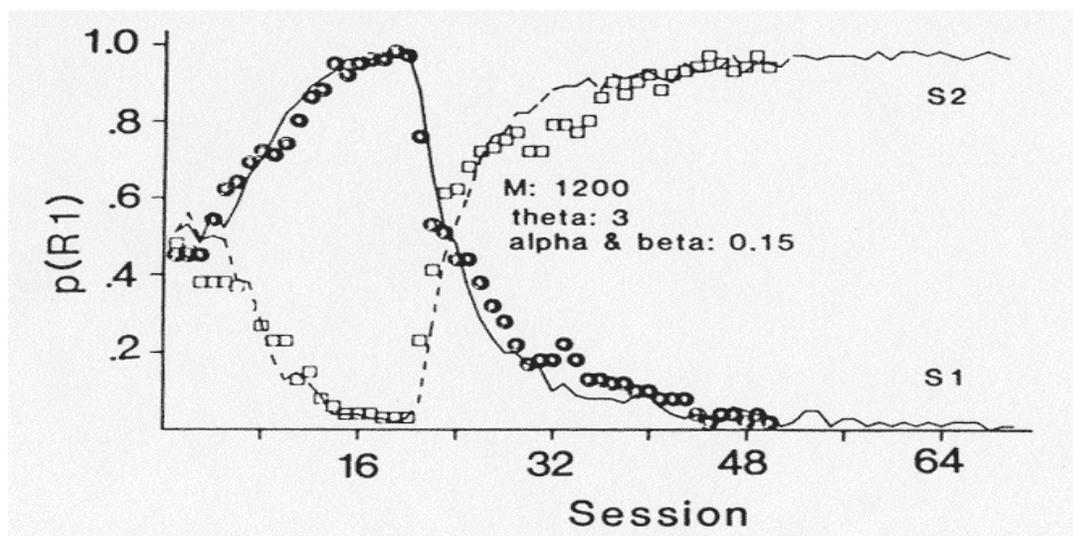


Figure 1-7 shows the results of an experiment that was identical to the one just considered except that the discrimination was reversed after twenty days of training. Again, the lines represent the mean results for nine simulated birds, while the points represent the mean results of nine real birds.

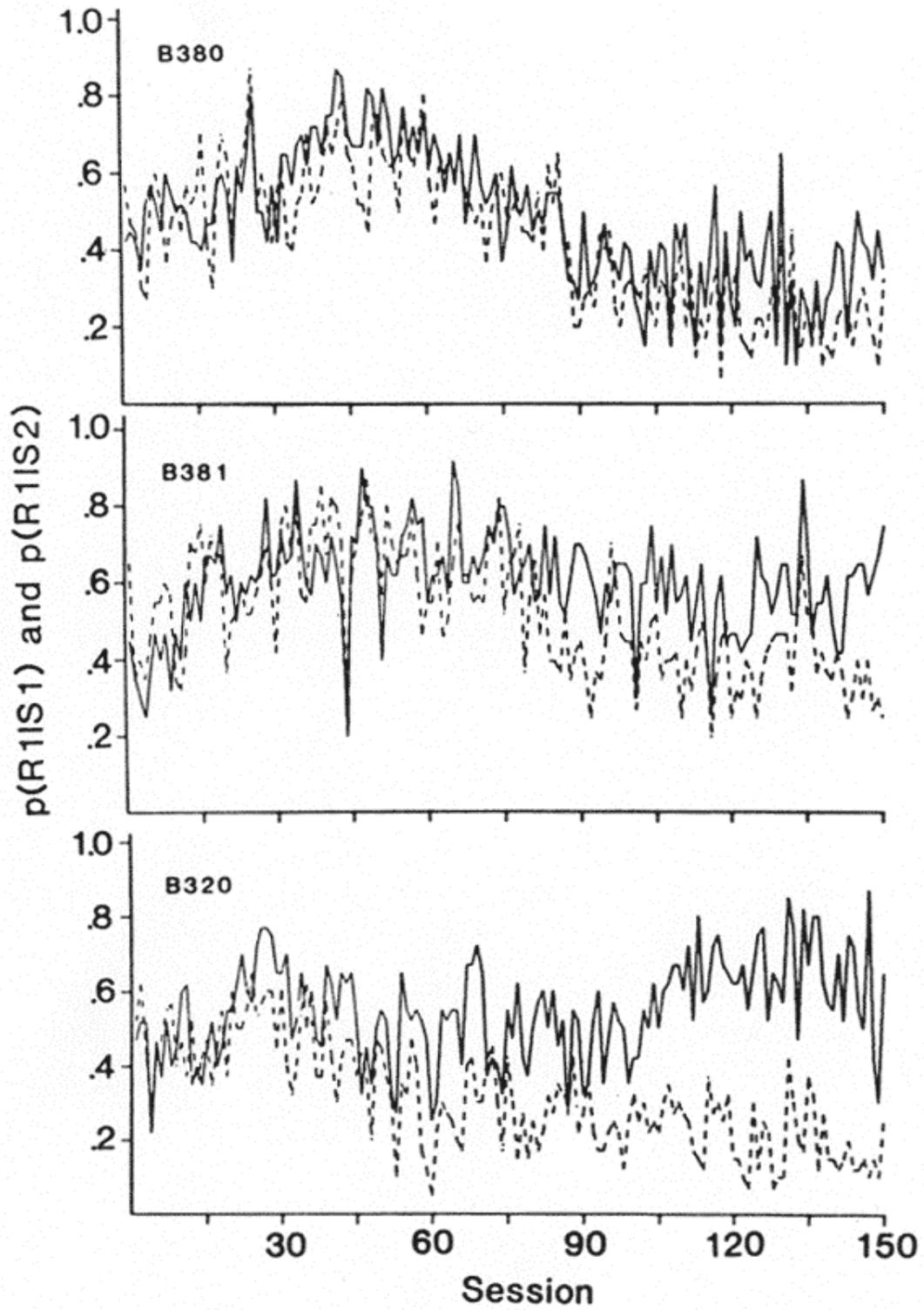
**Figure 1-7.** Mean proportion of  $R_1$  responses made in the presence of  $S_1$  and  $S_2$  by nine simulated birds (lines) and nine real birds (data points). The discrimination was reversed after 20 days.



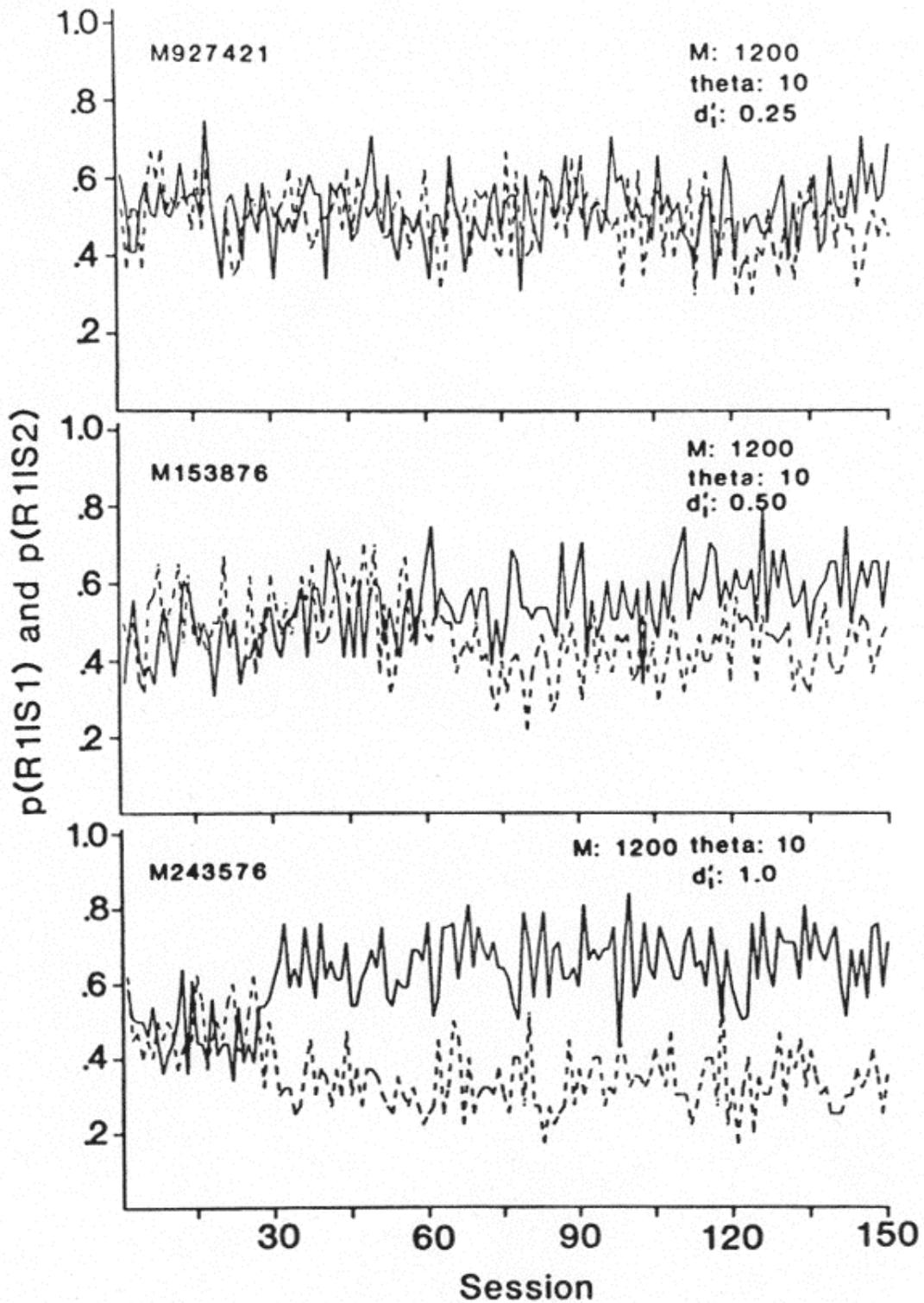
*Effects of Variations in Stimulus Discriminability.* The size of the physical difference between stimuli presented for discrimination affects the course of acquisition and the asymptotic performance of subjects. Nevertheless, there are large differences in the way different individuals are affected by the same stimulus difference. This may be seen in Figure 1-8, which shows acquisition curves for three pigeons trained to discriminate between two levels of white-noise intensity differing by 3 dB. A series of simulations done to find the (theoretical) source of such individual differences showed that they result primarily from differences in the response of the sensory system. Identical stimulus differences correspond to different values for the discriminability of the sensory effects ( $d'_1$ ).

Simulated data showing the effects of varying the value of  $d'_1$  are shown in Figure 1-9. The values of  $d'_1$  used in the simulations were chosen to produce data that are roughly comparable to those of the real birds shown in Figure 1-8.

**Figure 1-8.** Acquisition of discrimination between levels of white noise differing by 3 dB by three individual birds.



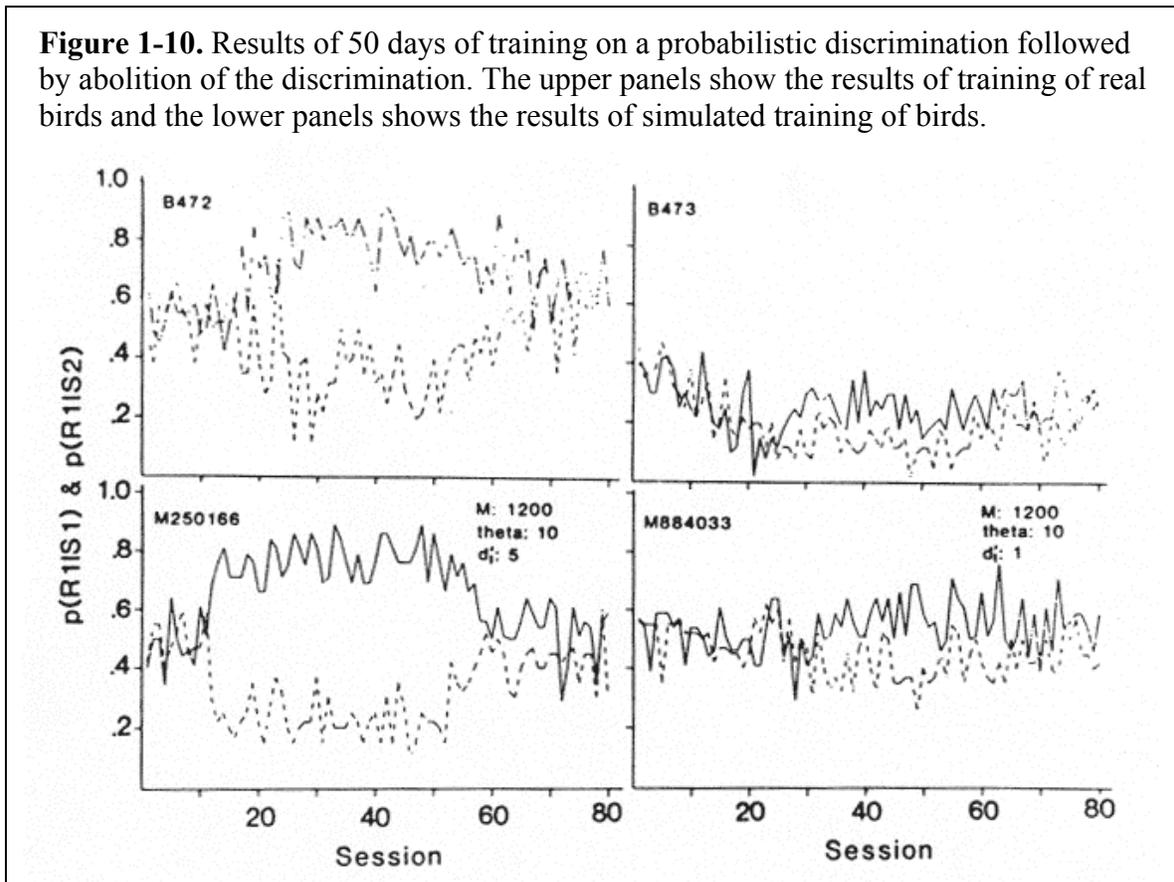
**Figure 1-9.** Acquisition of a discrimination by three simulated birds, each assigned the value  $d'$ , shown in the panel.



*Probabilistic Discrimination Training.* In another experiment done in collaboration with Julie Sage-Day, pigeons were trained to discriminate between two levels of white noise differing by 30 dB, but were rewarded probabilistically. Specifically, in the presence of  $S_1$ , response  $R_1$  was rewarded with probability 0.75, and response  $R_2$  was rewarded with probability 0.25. These reward probabilities were reversed in the presence of  $S_2$ . After fifty days of training, the probability of reward for each of the responses was shifted to 0.5, regardless of which stimulus was presented. This was done to abolish the discrimination.

The results for two birds are presented in Figure 1-10, together with roughly comparable results produced by the simulation program. The results for Bird 472 are quite representative of those obtained from other birds in this experiment, but those for Bird 473 are atypical and are shown to illustrate the range of individual differences. The sole difference in the two simulation data sets is the value of the parameter  $d'$ . Comparable differences in the results of simulations cannot be produced by manipulating any of the other parameters of the LCM model.

**Figure 1-10.** Results of 50 days of training on a probabilistic discrimination followed by abolition of the discrimination. The upper panels show the results of training of real birds and the lower panels shows the results of simulated training of birds.



### CONCLUDING REMARKS

The model under discussion has implications for a number of important psychological phenomena such as generalization, categorization and concept formation, absolute

identification of stimuli, and the psychophysics of intensity resolution. The last two topics are considered in detail by Chase (1983) in Chapter 3.

With respect to stimulus generalization, the response-selection process illustrated in Figure 1-1 leads to predictions for generalization that are quite similar to those expected on the basis of a criterion-decision rule of the type assumed in signal-detection theory (see Heinemann, Avin, Sullivan and Chase, 1969).

The situations considered in this chapter involve stimuli that vary within a single dimension, but the model can be extended easily to deal with stimuli that vary in a number of dimensions. This involves generalizing the response-selection rule so that the selection of the response is based on the joint probability densities of the various stimulus dimensions. Chase and I have used a program representing such a generalized version of the model in order to simulate generalization along two dimensions after training subjects to discriminate between compound stimuli (the “redundant case” described by Chase and Heinemann, 1972, and the “conditional case” described by Heinemann and Chase, 1970). We had previously analyzed these cases in terms of signal-detection theory. The simulation showed that, as in the one-dimensional case, the assumed response-selection process leads to results that are essentially the same as those expected on the basis of signal-detection theory. The relation between the signal-detection model and the present one needs to be examined in more detail, but it is clear that the two models lead to quite similar predictions for most of the phenomena of generalization.

## REFERENCES

- Brenner, N., Ph.D. dissertation, City University of New York.
- Chase, S. 1983. Pigeons and the magical number seven. In Commons, M. L, Herrnstein, R. J. & Wagner, A. R. (Eds.). (1983). *Quantitative Analyses of Behavior: Discrimination Processes*. Cambridge: Ballinger (Chapter 3, pages 37-58).
- Chase, S. & Heinemann, E.G. (1972). Decisions based on redundant information: An analysis of two-dimensional stimulus control. *Journal of Experimental Psychology* 92: 161-75.
- Estes, W.K., 1955. Statistical theory of spontaneous recovery and regression. *Psychological Review* 62: 145-54.
- Heinemann, E.G. (1983). The presolution period and detection of statistical associations. In Commons, M. L, Herrnstein, R. J. & A. R. Wagner (Eds.). *Quantitative Analyses of Behavior: Discrimination Processes*. Cambridge: Ballinger (Chapter 2, pages 21-35).
- Heinemann, E.G., Avin, E. Sullivan, M.A. & Chase, S. (1969). Analysis of stimulus generalization with a psychophysical method. *Journal of Experimental Psychology* 80: 215-24.
- Heinemann, E.G., & Chase, S., (1970). Conditional stimulus control. *Journal of Experimental Psychology* 84: 187-97.
- Heinemann, E.G., Sage-Day, J. & Brenner, N., (1981). Retroactive interference in discrimination learning. *Science* 214: 1254-57.
- Landauer, T.K., (1975). Memory without organization: Properties of a model with random storage and undirected retrieval. *Cognitive Psychology* 7: 495-531.

Osgood, C.E. (1949). The similarity paradox in human learning: A resolution.  
*Psychology* 56: 132-43.

### ACKNOWLEDGEMENTS

This article is reproduced from the book with some minor changes, e.g. in formatting and in references to other chapters in the book. The research reported here was supported by NIMH Grant No. MH 18246 and PSC-CUNY Grant No. 14002.

### NOTES TO CHAPTER 1

---

<sup>1</sup> In future versions of this model it may be useful to assume that only a limited number of the LCM's storage locations are available at any time, and that the loci of these storage locations drift slowly over time. Such an assumption might make it possible to deal with a variety of temporal effects such as spontaneous recovery and regression, trial-spacing effects, etc. This general concept is similar to that incorporated in the theories of Estes (1955) and Landauer (1975). The entire model presented here was influenced strongly by stimulus-sampling theory. It also has much in common with Landauer's model, but was developed independently of the latter.

<sup>2</sup> Although all the theoretical results that will be presented are based on these assumptions about the sampling process, the assumptions should be regarded as tentative. In the future, somewhat different assumptions may turn out to be preferable: For example, ( 1 ) that the sample also contains negative records, (2) that the sample size varies from trial to trial, (3) that sampling continues until the sample contains at least one record containing information about each of the discriminative stimuli, etc.

<sup>3</sup> The response proportions shown in Figure 1-3 and in all other figures reflect responding on trials that followed trials ending with a correct response. In other words, the correction (rerun) trials are excluded.