NONHUMAN SHORT-TERM MEMORY: A QUANTITATIVE REANALYSIS OF SELECTED FINDINGS

JOHN T. WIXTED

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Intense interest in the experimental analysis of nonhuman short-term memory was occasioned by the "cognitive revolution" in the early 1970s. As researchers tested and refined their theoretical models, a sizable empirical literature on the subject rapidly emerged. Generally missing from that literature, however, is an integrated account of the empirical laws of delayed stimulus control. In its absence, cognitive theories have grown increasingly complex and their connection to established principles of learning more remote. The present article reviews a representative segment of the nonhuman memory literature in an effort to identify and tentatively quantify empirical laws that may underlie some of the major findings. Specifically, a delay-reduction principle adapted from the conditioned reinforcement literature and a proposed principle of serial stimulus compounds were evaluated against data from delayed matching to sample, serial probe recognition, differential outcome, directed forgetting, and surprisingness preparations. Although by no means offering a complete description of the data, the proposed analysis organizes an apparently disparate collection of empirical results and suggests several new lines of inquiry into the subject.

Key words: short-term memory, delayed stimulus control, delayed matching to sample, choice, conditioned reinforcement, animal cognition, serial probe recognition, directed forgetting, delay reduction

Empirical inquiry into the subject of nonhuman short-term memory has proceeded at a rapid pace since the emergence of cognitive theories approximately 20 years ago. Throughout this period, efforts to test a variety of cognitive models have generated a number of innovative methods of investigation as well as a host of interesting behavioral findings. Nevertheless, the ratio of theory to data often seems unacceptably high, and efforts to identify common empirical principles of memory are relatively rare. In their absence, individual cognitive theories have grown in number and complexity, perhaps occasionally exceeding reasonable limits of parsimony (cf. Branch, 1977; Grant, 1981a). An analysis of short-term memory in terms of basic empirical principles, as opposed to cognitive models, may help to condense what is now a large and diverse empirical literature and suggest fruitful lines of inquiry not readily derived from current accounts. In addition, for those who prefer theoretical models, such an analysis may provide boundary conditions within which those models must operate. In view of these consid-erations, the present article offers one possible functional analysis of some of the more important findings from the nonhuman memory literature.

BACKGROUND

For the last 30 years, the most widely used procedure for the study of short-term memory in pigeons and monkeys has been the delayed matching-to-sample (DMTS) task or some variation thereof (Blough, 1959). A DMTS trial consists of the presentation of a sample stimulus (e.g., a red or green light) on a center key, followed by a delay interval during which the sample is extinguished, followed by the concurrent presentation of the comparison stimuli (e.g., both the red and the green light) on two side keys. A response to the stimulus that matches the sample is reinforced, whereas a response to the nonmatching stimulus terminates the trial. In the majority of DMTS experiments involving pigeons, the postsample delay, or *retention interval,* falls within the range of 0 to about 20 s. The basic finding is that matching accuracy declines (i.e., pigeons become more likely to forget the sample) as the delay interval increases (Blough, 1959).

The first modern cognitive theory of this phenomenon was proposed by Roberts (1972; Roberts & Grant, 1976) and is known as trace strength theory. According to this account, the

presentation of the DMTS sample stimulus establishes an isomorphic memorial representation, or memory trace, of the sample's physical characteristics. The strength of the memory trace is assumed to increase gradually as a function of time spent in the presence of the sample and degrade as a negatively accelerated function of time in its absence. This relatively simple and straightforward theory prevailed for several years, but has now been replaced by a variety of more detailed cognitive models based upon instructional encoding, active and inactive memories, outcome expectancies, temporal tagging, and rehearsal (Kendrick, Rilling, & Denny, 1986). Some of these theories, and the data that bear on them, will be considered in a later section.

The common denominator of all cognitive theories of DMTS performance is the concept of representation. Thus, most contemporary models consist of hypotheses regarding the way in which the sample stimulus is initially represented (or coded) and how that representation is processed during the delay. By contrast, a functional analysis seeks to identify variables that govern the ability of a stimulus to exert discriminative control across a temporal distance (cf. Branch, 1977; Catania, 1984; Jans & Catania, 1980). As indicated above, the recent surge of cognitive theories has not been paralleled by an increased understanding of the functional principles that underlie the delayed response. Nevertheless, a review of the findings produced by those theories suggests several empirical laws that may warrant careful evaluation.

FUNCTIONAL ANALYSIS OF DMTS

Procedurally, the DMTS paradigm resembles an autoshaping or conditioned reinforcement arrangement in that the sample stimulus reliably predicts the impending availability of primary reinforcement (Roberts & Kraemer, 1982; Santi, 1984). Thus, it would not be surprising to find that variables influencing those stimulus functions also effect the ability of a stimulus to exert discriminative control across a temporal distance. Indeed, the intimate association between conditioned reinforcement and stimulus control in general was recognized long ago (Keller & Schoenfeld, 1950; Skinner, 1938). Keller and Schoenfeld, for example, proposed that "In order to act as a conditioned reinforcer for any response, a stimulus must have status as a discriminative stimulus for some response" (p. 236). Although this statement has proven to be too strong, many findings support the general idea that discriminative and conditioned reinforcer strengths covary (e.g., Dinsmoor, 1950; Fantino & Logan, 1979).

Although still an area of active research, one variable consistently found to determine the strength of a conditioned reinforcer is the average reduction in delay to reinforcement correlated with the onset of the stimulus (Fantino, 1977). That is, if the average delay between successive reinforcements is $\bar{T}$, and the average delay to reinforcement associated with the onset of a stimulus is $\bar{d}$, then the quantity $\bar{T} - \bar{d}$ represents the reduction in delay to reinforcement associated with onset of the stimulus. According to the delay-reduction hypothesis, the strength of a conditioned reinforcer varies directly with the quantity $\bar{T} - \bar{d}$ (Squires & Fantino, 1971).

Unlike standard conditioned reinforcement arrangements such as the concurrent chains procedure, the onset of the sample stimulus in DMTS predicts the imminent presentation of stimuli intermittently associated with reinforcement (i.e., the comparison stimuli) rather than the imminent availability of the reinforcer itself (which depends on whether or not the response is correct). Nevertheless, a large number of studies to be considered later appear to be consistent with the hypothesis that the discriminative strength of the DMTS sample stimulus is determined by its delay-reduction properties measured with respect to the presentation of the trial outcome stimuli. That is, $\bar{T}$ is now the average delay between successive reinforcement occasions (i.e., between successive DMTS choice phases), and $\bar{d}$ is the average delay from the onset of the DMTS sample stimulus to the presentation of the choice phase. As long as the comparison stimuli are at least occasionally associated with reinforcement, the delay reduction quantity, $\bar{T} - \bar{d}$, appears to offer an adequate measure of the discriminative strength of the DMTS sample stimulus.

Several researchers, especially Roberts and Kraemer (1982), have recognized the importance of delay-reduction considerations in DMTS, albeit in the context of Gibbon's (1977) scalar expectancy theory. However, no

previous attempts to develop an integrated account of the empirical literature based on such concepts can be identified. Because of the now extensive DMTS data base, it is possible to formulate a tentative quantitative account of discriminative strength involving delay-reduction variables. As defined above, the delay-reduction quantity, $\bar{T} - \bar{d}$, happens to be equivalent to the intertrial interval (ITI) in a standard DMTS arrangement. Several early experiments testing the notion of intertrial interference established that DMTS performance generally improves as the size of the ITI increases (Grant, 1975; Maki, Moe, & Bierley, 1977; Nelson & Wasserman, 1978). More recently, the interaction between the size of the retention interval, t, and the size of the ITI has been clarified as well in that DMTS performance is usually found to vary as a function of the ratio of those two measures (Roberts & Kraemer, 1982; Santi, 1984; Wilkie, 1984). That is, the larger the ratio of ITI to t, the better the performance.

With some modification, the empirically derived ratio, ITI/t, was assumed to provide a quantitative estimate of the discriminative strength of a DMTS sample stimulus. Although performance generally improves as that ratio increases, absolute sensitivity to variations in the size of the ITI appears to vary considerably across experiments (Edhouse & White, 1988; Roberts & Kraemer, 1982; Santi, 1984). Santi (1984), for example, found that ITI effects are greatly attenuated when the houselight remains on throughout the experimental session. To allow for relative differences in sensitivity to these temporal variables across experiments, the discriminative strength of the DMTS sample stimulus was assumed to vary according to the ratio (ITI + γ)/(t + α), where α and γ are scaled in temporal units. The larger the value of these parameters, the less sensitive the ratio is to variations in t and ITI. Thus, for example, the proportional change in the value of the ratio resulting from a 5-s increase in the ITI would be greater when γ equals 5 s than when it equals 50 s.

Although in most experiments the quantity $\bar{T} - \bar{d}$ and the ITI are equivalent measures, in some of the studies considered below they differ. Thus, the discriminative strength of the sample may be more appropriately represented as (d_r + γ)/(t + α), where d_r is the delay-reduction quantity, $\bar{T} - \bar{d}$. According to this measure, when the delay reduction, d_r, is very small or the retention interval, t, is very large, the sample stimulus should exert very little discriminative control. Under those conditions, other stimuli present in the environment, notably the comparison stimuli themselves, presumably will govern responding independent of the sample. If the discriminative strength of the sample stimulus varies according to (d_r + γ)/(t + α) and the discriminative strength of all other stimuli present at the time of the response remains a constant, β, across changes in d_r and t, then the relative strength of the sample, p, may be expressed as

$$p = \frac{\dfrac{d_r + \gamma}{t + \alpha}}{\dfrac{d_r + \gamma}{t + \alpha} + \beta}, \qquad (1)$$

where the denominator represents a cumulative measure of the control exerted by all past and present stimuli. As arranged in Equation 1, the parameters α and γ capture sensitivity to changes in t and d_r, but they do not correctly scale the absolute magnitudes of the quantities t + α and d_r + γ. Thus, for example, Equation 1 requires that performance improve as d_r increases, but not when γ increases; its only function is to calibrate sensitivity to d_r. Therefore, when γ changes, the scale on the entire quantity d_r + γ changes as well. The practical implication for Equation 1 is that changes in sensitivity to d_r and/or t will be reflected by changes in β due to the adjustment in scale.

As indicated earlier, the delay-reduction properties of the sample stimulus are assumed to be relatively insensitive to the end-of-trial reinforcement probabilities and are instead determined by delays to the presentation of stimuli intermittently associated with reinforcement (i.e., the comparison stimuli). Nevertheless, responding during the choice phase of a DMTS trial may be biased in one direction or the other when asymmetrical reinforcement probabilities are introduced. For example, if correct responses to one comparison stimulus are reinforced at a much higher probability than correct responses to the other, a bias toward the more frequently reinforced alternative may be observed independent of the discriminative strength of the sample stimuli. Some consideration of the relative probability

of reinforcement associated with each choice alternative would thus seem to be in order (cf. Squires & Fantino, 1971).

In DMTS, the prevailing reinforcement contingencies during the choice phase can be viewed from two perspectives. Considered in light of the preceding sample, the probability of reinforcement associated with choosing the matching alternative is usually (but not necessarily) one, and the probability of reinforcement for choosing the nonmatching alternative is usually zero. Considered independent of the preceding sample, the probability of reinforcement associated with each comparison stimulus is the same. Presumably, when the retention interval is short the conditional reinforcement probabilities (i.e., those defined by the sample stimulus) should describe performance, and when the retention interval is long the unconditional reinforcement probabilities (i.e., those associated with the comparison stimuli independent of the sample) should provide a better description. At intermediate delays, both may contribute to varying degrees. One way to represent this conceptualization mathematically for a trial involving Sample 1 (S_1) is:

$$\frac{R_{1m}}{R_{1m} + R_{2n}} = p\left\{\frac{r_{1m}}{r_{1m} + r_{2n}}\right\}$$

$$+ (1 - p)\left\{\frac{r_1}{r_1 + r_2}\right\}, \quad (2a)$$

where R_{1m} denotes responding to Comparison Stimulus 1 (the matching alternative on S_1 trials), R_{2n} denotes responding to Comparison Stimulus 2 (the nonmatching alternative on S_1 trials), r_{1m} and r_{2n} refer to the conditional probabilities of reinforcement for matching (Alternative 1) and nonmatching (Alternative 2) responses, respectively, and r_1 and r_2 refer to the unconditional reinforcement probabilities associated with Alternatives 1 and 2 (cf. Herrnstein, 1961). The measures r_1 and r_2 are simply the obtained reinforcement frequencies averaged across all trials independent of the sample stimulus. The value of p in Equation 2a ranges from zero to one as defined by Equation 1. Note that the quantity on the left side of Equation 2a represents the familiar dependent measure "proportion correct" on trials involving

S_1. A similar expression describes proportion correct on trials involving S_2:

$$\frac{R_{2m}}{R_{1n} + R_{2m}} = p\left\{\frac{r_{2m}}{r_{1n} + r_{2m}}\right\}$$

$$+ (1 - p)\left\{\frac{r_2}{r_1 + r_2}\right\}. \quad (2b)$$

In their reformulation of the delay-reduction account of the concurrent chains procedure, Fantino and Davison (1983) found that the biasing effect of asymmetrical reinforcement was most accurately predicted by the square roots of the obtained reinforcement probabilities. A similar strategy (or some other means of accommodating undermatching) might be needed here as well, but, with the exception of a few studies considered later, the issue is usually irrelevant. Typically, r_{1m} and r_{2m} in Equations 2a and 2b (the reinforcement probabilities for matching responses on S_1 and S_2 trials, respectively) are both equal to one, and r_{2n} and r_{1n} (the reinforcement probabilities for nonmatching responses on S_1 and S_2 trials, respectively) are both equal to zero. Furthermore, r_1 equals r_2 such that, in the standard case, Equations 2a and 2b both reduce to the general expression:

$$\frac{R_{1m}}{R_{1m} + R_{2n}} = \frac{R_{2m}}{R_{2m} + R_{1n}} = 0.5p + 0.5, \quad (3)$$

which, after replacing p with the right hand side of Equation 1 and rearranging, becomes

$$\frac{R_m}{R_m + R_n} = 0.5\left\{\frac{1/\beta}{1/\beta + \dfrac{t + \alpha}{d_r + \gamma}}\right\} + 0.5, \quad (4)$$

where R_m and R_n denote responding to the matching and nonmatching alternatives, respectively.

Equation 4 represents a tentative quantitative formulation of the present functional account of DMTS performance. Its essential feature is the predicted relationship between delay reduction and proportion correct. One previous functional analysis of DMTS performance approached the subject from a different point of view. White and McKenzie (1982) expanded upon the signal detection model of

Davison and Tustin (1978) to derive a bias-free measure of initial and delayed stimulus discriminability. The strength of their analysis is that it permits a clear empirical dissociation of the variables affecting initial stimulus discriminability, decay rate of discriminative control over the retention interval, or both (McCarthy & White, 1987; White, 1985). On the other hand, the signal detection model makes few predictions about how variables such as delay reduction ought to affect performance. In what follows, the framework specified by Equation 4 will be evaluated in light of the large cognitive literature on nonhuman short-term memory.

Temporal Analyses

The most elementary finding from the DMTS literature is the negatively accelerated relationship between the size of the retention interval and proportion correct. Figures 1 and 2 illustrate this effect using representative pigeon group data taken from Roberts (1972) and Grant (1975). The ITI was constant in both of these experiments, and under such conditions, Equation 4 reduces the two-parameter equation,

$$\frac{R_m}{R_m + R_n} = 0.5 \left\{ \frac{1/\beta'}{1/\beta' + (t + \alpha)} \right\} + 0.5, \quad (5)$$

where β' is equal to $\beta/(d_r + \gamma)$. Equation 5 was fitted to the data shown in Figures 1 and 2 using a nonlinear least squares regression program (Wilkinson, 1988). Clearly, the function provides an adequate description of these basic results, though it should be acknowledged that many two-parameter functions (e.g., a negative exponential) would do as well.

Another basic finding from the DMTS literature is that performance improves as the ITI (and, hence, d_r) increases. White (1985), for example, exposed 5 pigeons to a DMTS task involving either a 5-s or 20-s ITI, and the retention interval ranged from 0.5 to 20 s. A given ITI in this experiment was in effect for at least 15 sessions. Figure 3 illustrates the ITI effects for the group data and shows the best fit of Equation 4, setting d_r equal to the scheduled ITI. The negative value of γ implies that performance should reach chance levels when the ITI is reduced to 1.43 s (Equation 4 would obviously be undefined at lower ITI values).
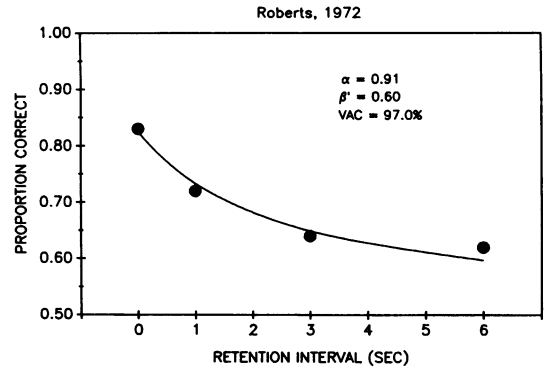

Fig. 1. Proportion of correct responses as a function of retention interval averaged over 10 pigeons from Roberts (1972; copyright 1972 by the American Psychological Association). The data were taken from the FR 5 (sample response requirement) condition and were estimated from his Figure 1. The solid curve represents the least squares fit of Equation 5. Adapted by permission.

In the most comprehensive study of the effects of temporal variables on DMTS performance in pigeons, Roberts and Kraemer (1982) factorially manipulated the retention interval (0.5, 1, 2, and 4 s) and ITI (4, 8, 16, and 32 s) across sessions. The results of this study are presented in Table 1 along with the predicted values obtained from a least squares fit of Equation 4. Note that in both the obtained and predicted functions, performance remains essentially constant across different values of ITI within a fixed ITI/t ratio. Furthermore, matching accuracy generally improves as the size of that ratio increases.

Figure 4 shows the obtained data as well as the predicted function obtained by averaging across multiple values for each ITI/delay ratio in Table 1. The figure clearly reveals that performance improves in a linear fashion with the log of the ratio ITI/t and that the present model closely conforms to that result. It is interesting to note that the value of γ is rather high in this experiment (24.39), suggesting relatively little sensitivity to variations in the ITI. One possible explanation for its high value is that the ITI was changed on a daily basis such that behavior may not have become fully stable before a new one was introduced. In that case, the scheduled ITIs in the extreme conditions (4 and 32 s) would overestimate the functional range of ITIs, generating a flatter function than would otherwise exist.
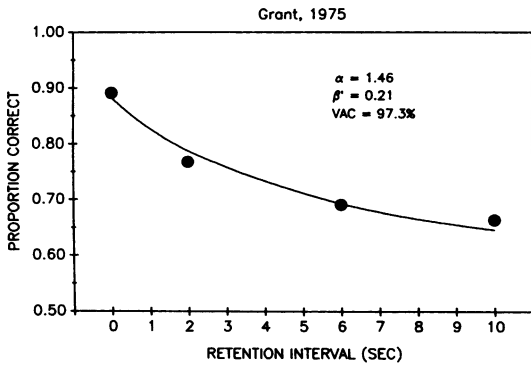
Two other experiments reported by Roberts

Fig. 2. Proportion of correct responses as a function of retention interval averaged over 10 pigeons from Grant (1975; copyright 1975 by the American Psychological Association). The data were taken from the last 5 days of baseline training in his Experiment 1. The solid curve represents the least squares fit of Equation 5. Adapted by permission.
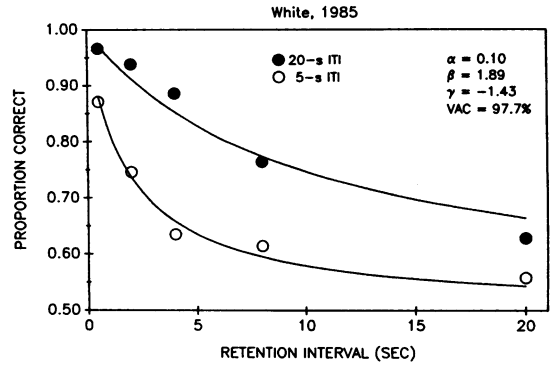
Fig. 3. Proportion of correct responses as a function of the retention interval for 5- and 20-s intertrial intervals. The data, which are averaged over 5 pigeons, were taken from Table 4 of White (1985; copyright 1985 by the Society for the Experimental Analysis of Behavior, Inc.). The solid curve represents the least squares fit of Equation 4. Adapted by permission.

and Kraemer (1982), conducted under the framework of scalar expectancy theory, bear on the present analysis. As indicated above, d_r is an average value that is established across a number of trials (i.e., the average reduction in delay to reinforcement associated with the onset of the sample). Therefore, although variations in the size of t from trial to trial may have profound effects on performance, transient variations in the ITI should not. In agreement with this notion, Roberts and Kraemer (1982) found that the usual performance advantage resulting from the use of long ITIs can be eliminated by varying the duration of the ITI within a session. If, however, the average of multiple ITI durations in one condition exceeds the average of multiple ITI durations in a second condition, response accuracy in the former condition should exceed that of the latter. A second experiment reported by Roberts and Kraemer (1982) supported this prediction.

Roberts (1980) examined DMTS performance using either a 1-s ITI or a 20-s ITI when the same sample (a red keylight) was used repeatedly within a session (following baseline training using red and green keylights as samples). The purpose of the experiment was to test a trace strength interpretation of the ITI effect. According to trace strength theory, DMTS performance is poor when the ITI is short because the memory trace of the sample from the preceding trial competes with the

memory trace of the sample on the current trial. Increasing the ITI allows the preceding trace to degrade more completely, thereby reducing competition. When the same sample is used repeatedly in a session, however, the residue of the sample memory trace from the preceding trial should, if anything, strengthen the memory trace of the sample on the current trial. In that case, performance at the 1-s ITI should exceed performance at the 20-s ITI. From a delay-reduction perspective, however, the opposite result would be anticipated. That is, in the 20-s ITI condition, the sample stimulus is associated with a much greater reduction in delay to reinforcement than in the 1-s ITI condition. In agreement with a delay-reduction interpretation, Roberts found that performance was better with the 20-s ITI.

Differential Sample Delays

The studies reviewed above, which manipulated the size of the ITI, support the notion that delayed stimulus control is at least partially determined by the sample's delay-reduction properties. Another way to manipulate delay reduction is to vary the average delay, $\bar{d}$, associated with individual sample stimuli while holding $\bar{T}$ constant. Honig (1987) arranged such a procedure using successive DMTS. In a successive DMTS procedure, a sample is presented, followed by a retention interval, followed by the presentation of a single test stimulus. If the test stimulus matches

Table 1

Observed and predicted performance as a function of the intertrial interval (ITI) and retention interval (*t*).

ITI/*t*	ITI (s)	*t* (s)	Observed	Predicted
1	4	4	.740	.728
2	4	2	.772	.783
2	8	4	.740	.745
4	4	1	.806	.821
4	8	2	.796	.799
4	16	4	.788	.772
8	4	0.5	.848	.844
8	8	1	.838	.835
8	16	2	.818	.824
8	32	4	.810	.813
16	8	0.5	.861	.858
16	16	1	.864	.859
16	32	2	.850	.860
32	16	0.5	.894	.879
32	32	1	.890	.890
64	32	0.5	.901	.907

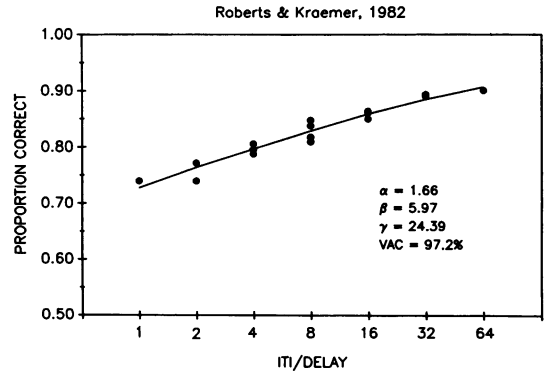*Note.* The data are from Roberts and Kraemer (1982).

Fig. 4. Proportion of correct responses as a function of the ratio ITI/*t*. The data were taken from Table 2 of Roberts and Kraemer (1982; copyright 1982 by the American Psychological Association). The solid curve represents the least squares fit of Equation 4. Adapted by permission.

the sample, responding is reinforced, otherwise responses have no scheduled consequences. The dependent measure on such a procedure is referred to as the discrimination ratio, which is equal to the number of responses made to the matching stimulus divided by the number of responses made to both the matching and nonmatching stimuli across trials. Honig (1987) compared performance maintained by two sample stimuli, one of which, S1, was followed routinely by a 1- or 5-s retention interval and the other of which, S2, was followed routinely by a 5- or 10-s retention interval. Because S1 is followed by a shorter average retention interval than S2, the reduction in delay to reinforcement associated with the onset of S1, $\bar{T} - \bar{d}_1$ (about 12 s in this experiment), is greater than that associated with the onset of S2, $\bar{T} - \bar{d}_2$ (about 7.5 s). Therefore, when both are tested at the same retention interval (viz., 5 s), performance following S1 should be more accurate than that following S2.

Although Equation 4 was developed within a choice framework, for purposes of the present analysis it was assumed to apply to successive DMTS as well. Honig's (1987) results are depicted in Figure 5 along with the least squares fit of Equation 4. Because only four data points were fitted, the high proportion of data variance accounted for is to be expected. Nevertheless, Figure 5 graphically illustrates

the important point: Performance is higher at the 5-s retention interval for the stimulus associated with the greater reduction in delay to reinforcement.

Several conceptually similar experiments have presented cues in compound with a sample stimulus to signal the size of the upcoming retention interval (MacDonald & Grant, 1987; Wasserman, Grosch, & Nevin, 1982). Thus, for example, a red sample in compound with horizontal lines (R/H) might predict a short retention interval, and a red sample in compound with vertical lines (R/V) might instead predict a long retention interval. Under these
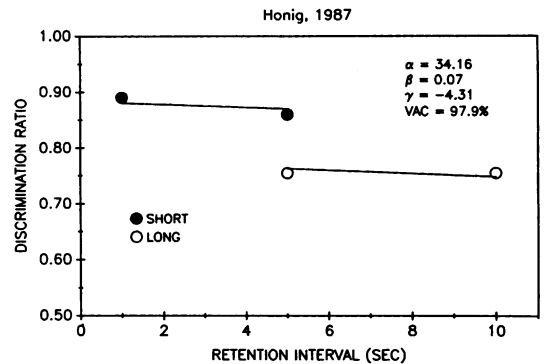

Fig. 5. Proportion of correct responses as a function of the retention interval for the long and short average retention interval conditions. The data were estimated from Figure 3 in Honig (1987) and averaged across separate and combined sessions (copyright 1987 by the Psychonomic Society, Inc.). The solid curves represent the least squares fit of Equation 4. Adapted by permission.

conditions, the R/H compound would be associated with a greater reduction in delay to reinforcement than the R/V compound. Furthermore, compared to red alone (which is sometimes followed by a short delay, sometimes by a long delay), one would expect better performance with the R/H compound sample and worse performance with the R/V compound sample. A series of experiments reported by Wasserman et al. (1982) were consistent with this prediction, although attempts to fit Equation 4 to their data failed to yield unique parameter estimates.

MacDonald and Grant (1987) examined the effects of miscuing the size of the retention interval. In baseline training, color sample stimuli were presented in compound with horizontal or vertical lines. The horizontal lines signaled a short (1-s) retention interval, and the vertical lines signaled a long (5-s) retention interval. On test trials, these retention intervals were reversed unexpectedly. In general, the delay-reduction analysis predicts that, for any retention interval tested, the compound usually associated with the shorter delay to reinforcement (and therefore the greater delay reduction) should produce a higher level of accuracy than the compound usually associated with the longer delay to reinforcement. With one exception (their Experiment 4) the results reported by McDonald and Grant were consistent with this prediction. However, one finding not anticipated by the present analysis was that performance following the long-cue compound was especially poor at the (miscued) short delay and actually improved as the length of the delay increased. This interesting finding suggests that the strength of a discriminative stimulus may be delay specific when only one retention interval is employed. That is, a generalization gradient of discriminative strength may be conditioned around a particular delay such that it is strongest at the baseline delay and weaker at other delays (longer or shorter).

Serial Probe Recognition

Several investigators have modified the standard DMTS arrangement by presenting several sample stimuli in succession (i.e., a list of stimuli) on each trial. The question of interest is whether species other than humans exhibit better retention for items at the beginnings and ends of lists relative to the middle items. These effects are commonly termed *primacy* and *recency* effects, respectively. The most widely used technique in the study of list retention in nonhumans is serial probe recognition (SPR). In a typical SPR arrangement, a trial consists of the successive presentation of several discriminative stimuli followed by the retention interval. In the choice phase, a single probe item is presented that may or may not match one of the stimuli in the most recently presented list. The subject must correctly classify the probe as a matching or a nonmatching item in order to procure a reinforcer.

Several experiments using SPR or similar procedures have found evidence for a recency effect, but not a primacy effect, in dolphins (Thompson & Herman, 1977), pigeons (MacPhail, 1980; Shimp, 1976; Shimp & Moffitt, 1974), rats (Roberts & Smythe, 1979), and monkeys (Gaffan, 1977; Roberts & Kraemer, 1984). A number of other studies, however, have found clear evidence of both primacy and recency effects in pigeons, monkeys, and humans (Buchanan, Gill, & Braggio, 1981; Roberts & Kraemer, 1981; Sands & Wright, 1980a, 1980b; Santiago & Wright, 1984; Wright, Santiago, & Sands, 1984).

Santiago and Wright (1984) conducted the most comprehensive analysis of SPR performance in pigeons. They trained 4 pigeons on an SPR procedure in which four samples were presented for 2 s each and were separated by an interstimulus interval (ISI) of 1 s. The delay following the last item ranged from 0 to 10 s. At the 0-s delay, all subjects exhibited a strong recency effect, but a primary effect was not apparent. At intermediate delays (1 to 2 s), the pigeons exhibited both primacy and recency effects. When the delay following the last stimulus was increased to 10 s, the recency effect disappeared completely for all subjects, and, for 2 pigeons, a primacy effect was still evident. Wright et al. (1984) conducted a similar study with monkeys and obtained comparable results.

These findings are especially interesting because, for the first time, reliable serial position effects that have facilitated memory research in humans have been demonstrated in nonhumans. Wright et al. (1984) did not offer a comprehensive theory to account for their results, but they did appeal tentatively to interference principles. According to this view, when the retention interval is short, retroactive in-

terference (i.e., interference caused by subsequent items) is high and proactive interference (i.e., interference caused by preceding items) is low, thus accounting for the recency effect. At longer delays, the situation is gradually reversed (i.e., retroactive interference is low and proactive interference is high) such that a primacy effect is favored.

The delay-reduction analysis of these findings is based on the assumption that, when a series of stimuli is employed on a single trial, the delay-reduction characteristics of each stimulus must be measured relative to the delay to reinforcement signaled by the preceding stimulus. For the first stimulus in the series, S_1, the reduction in delay to reinforcement is equal to the ITI, as usual. For the second stimulus, S_2, the delay reduction is equal to the delay signaled by S_1 (d_1) minus the delay signaled by the presentation of S_2 (d_2). In general, the reduction in delay to reinforcement associated with S_i is equal to $d_i - d_{i-1}$.

According to this analysis, the first stimulus in a series typically will be associated with the largest reduction in delay to reinforcement, whereas that associated with subsequent stimuli will be much smaller. However, although delay-reduction considerations may favor the first stimulus in a series, the retention interval invariably favors the last. When the delay following the last stimulus is very short, a pronounced recency effect may be observed despite unfavorable delay-reduction characteristics. When the delay following the last stimulus is increased, the relative delay differences between the individual stimuli in the series begin to diminish, and performance should be determined primarily by the delay-reduction characteristics of each. In that case, the serial position function should be characterized primarily by a primacy effect.

Figure 6 shows the serial position functions from Santiago and Wright (1984) averaged over the 4 subjects at delays of 0, 2, and 10 s. The figure also depicts quantitative predictions derived from the present model. An assumption required for this fit was that the functional ITI increased along with the size of the retention interval. This assumption was necessary to capture an unusual characteristic of the data in Figure 6, namely, the percentage of correct responses associated with the stimulus in Serial Position 1 actually increased as the size of the delay increased. Although the
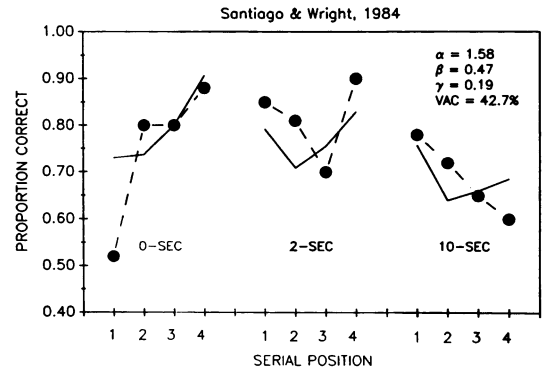

Fig. 6. Observed and predicted serial position functions for retention intervals of 0, 2, and 10 s. The data were averaged over 4 pigeons and were estimated from Figure 4 of Santiago and Wright (1984; copyright 1984 by the American Psychological Association). The solid curves represent the least squares fit of Equation 4. Adapted by permission.

scheduled ITI in this experiment was 3 s, pigeons initiated each trial with a peck in response to an auditory stimulus. Other investigators have observed that, following a long retention interval, pigeons tend to pause before initiating the next trial with a response to the sample stimulus (Zentall, Hogan, Howard, & Moore, 1978). For the present set of data, actual ITIs of 4, 8, and 10 s were assumed for retention intervals of 0, 2, and 10 s, respectively.

It is evident from the figure that the present model can accommodate some features of the data better than others (variance accounted for was only 43%). At the 0-s delay, only a recency effect is obtained, but the absolute level of performance associated with the first stimulus is considerably overestimated. At the 2-s delay, clear primacy and recency effects are apparent for both the obtained and predicted functions, but, again, some inaccuracies are apparent. At the longest delay, the model predicts a serial position function characterized primarily by a primacy effect, although a recency effect is still predicted (though none was obtained).

The data shown in Figure 6 suggest that there is surely more to SPR performance than is suggested by the present analysis. Nevertheless, the delay-reduction approach can account for some important aspects of the data and makes several predictions that may serve to facilitate further inquiry into the observed serial position effects. First, the model predicts

that the magnitude of primacy should increase as a function of the size of the ITI. As yet, the effects of this temporal parameter on SPR performance have not been studied. Second, with regard to the recency effect, the model predicts that increasing the ISI should offset the retention interval effects observed in the experiment by Santiago and Wright (1984). That is, as the ISI increases, the reduction in delay to reinforcement associated with the onset of the last stimulus in the series increases such that it should be able to exert discriminative control over a longer delay. Incidentally, this prediction is exactly the same as that made by a popular law of human short-term memory termed the *ratio rule* (Crowder, 1976). According to the ratio rule, as the ISI between items (e.g., words in a list) increases, the recency effect will be maintained over a longer delay. The present model suggests that the ratio rule may apply to other species as well.

An interesting unresolved question is why so many researchers have failed to detect a primary effect using SPR procedures. Santiago and Wright (1984) account for this on the basis of intertrial interference resulting from the use of a small sample stimulus pool. In contrast to other investigators, they selected sample stimuli for each trial from a pool of hundreds of items rather than from a pool of 10 or 12 items as is usually the case. When the sample pool is small, the subject may become confused and regard a test item as having been presented on the current list when in fact it was presented on an earlier trial. This problem is circumvented when a large stimulus pool is used and thus may allow the emergence of serial position effects that would otherwise be obscured by poor overall performance.

A delay-reduction interpretation suggests a different possibility. When the sample stimulus pool is very small, every sample appears equally often in each serial position. In that case, the average reduction in delay to reinforcement associated with each stimulus is the same and independent of its serial position on a given trial. Under those conditions, performance on a given trial should be determined solely by the retention interval associated with each stimulus. Because the retention interval is always shortest for the last stimulus and longest for the first, only a recency effect should be obtained. On the other hand, when a large number of stimuli are employed, each sample appears essentially once such that the temporal position of a stimulus in a series differentially signals the delay to reinforcement. In that case, the first stimulus in the series will generally be associated with the largest reduction in delay to reinforcement and a primacy effect may be observed. According to this analysis, it should be possible to generate a primacy effect using a smaller sample stimulus pool as long as each stimulus always appears in only one serial position.

Differential Reinforcement Probabilities

As indicated earlier, the majority of DMTS experiments arrange a reinforcement probability of one for a correct response and zero for an incorrect response. Under those conditions, performance should be described by Equation 4 (or Equation 5 for a constant ITI). However, a few studies have arranged asymmetrical reinforcement probabilities for a correct response to the comparison stimuli (DeLong & Wasserman, 1981; Harnett, McCarthy, & Davison, 1984; Santi & Roberts, 1985). Under these conditions, performance should be described more accurately by Equations 2a and 2b. In a study by Santi and Roberts (1985), the probability of reinforcement for a correct response following one sample, S_1, was 1.0, whereas a correct response following the other sample stimulus, S_2, was only .20. For a control group of pigeons, the probability of reinforcement for a correct response following either sample was .60. Note that, in both cases, the overall probability of reinforcement is the same. Nevertheless, pigeons in the differential outcome group (averaged over S_1 and S_2 trials) performed significantly better than those in the nondifferential outcome group. This result essentially replicated an earlier finding of DeLong and Wasserman (1981), who used a successive DMTS procedure.

Equations 2a and 2b apparently predict that, under the conditions employed by Santi and Roberts (1985), responding in the differential and nondifferential conditions should be equal. Because the delay-reduction parameters are assumed to be relatively insensitive to obtained reinforcement probabilities (as long as responding to the comparison stimuli is occasionally reinforced), the discriminative strength of the sample stimuli in the two conditions (p) should be the same. Any differences in per-

formance should therefore be captured by the reinforcement probability measures. Because r_{2n} and r_{1n} (the conditional reinforcement probabilities for nonmatching responses on S_1 and S_2 trials, respectively) in Equations 2a and 2b both equal zero, those equations may be rewritten as

$$\frac{R_{1m}}{R_{1m} + R_{2n}} = p + (1 - p)\left\{\frac{r_1}{r_1 + r_2}\right\} \quad (6a)$$

and

$$\frac{R_{2m}}{R_{1n} + R_{2m}} = p + (1 - p)\left\{\frac{r_2}{r_1 + r_2}\right\}. \quad (6b)$$

According to the above equations, performance on S_1 trials should exceed performance on S_2 trials because r_1 is greater than r_2. Furthermore, in both cases performance should deviate equally (but in opposite directions) from the nondifferential case in which r_1 equals r_2. Thus, averaged across S_1 and S_2 trials, performance in the differential and nondifferential conditions should be the same. In contrast to this prediction, Santi and Roberts (1985) found a clear advantage for the differential outcome condition. The authors interpreted their findings as evidence for the operation of reinforcement probability expectancies that served to augment delayed discriminative performance (cf. Brodigan & Peterson, 1976; Edwards, Jagielo, Zentall, & Hogan, 1982; Peterson, Wheeler, & Armstrong, 1978)

A possible alternative explanation for the different performance levels in the differential and nondifferential outcome conditions is suggested by examining the rate of responding to the sample stimuli in Santi and Roberts' (1985) experiment. Although no response requirement was in effect, the rate of responding to S_1 was two to three times higher than the rate of responding to S_2. Response rates to the two samples associated with an intermediate probability of reinforcement (i.e., .60) were not reported, but one can probably safely assume that they were essentially equal to each other. Because sample-specific responding has been shown to facilitate matching-to-sample performance (Urcuioli & Honig, 1980; Zentall et al., 1978), it might be argued that the differential outcome effect observed in this experiment was an indirect result of the differential sample stimulus response rates. Indeed, DeLong and Wasserman (1981) showed that

eliminating sample-specific response rates greatly attenuates (albeit does not completely eliminate) the differential outcome effect using successive DMTS.

Harnett et al. (1984) conducted another differential reinforcement probability experiment in which sample-specific responding was much less likely to occur. In that experiment, a single response to the sample stimulus turned off the stimulus and initiated the retention interval. When sample-specific responding is eliminated, the improvement in performance associated with the high probability of reinforcement sample stimulus should more or less offset the deterioration in performance associated with the low probability of reinforcement sample stimulus. Furthermore, if the retention interval is increased to a large value (such that p in Equations 6a and 6b is essentially zero), responding on S_1 trials should stabilize above .50, and responding on S_2 trials should stabilize below .50.

Harnett et al. (1984) exposed 6 pigeons to a range of reinforcement probabilities for correct response across 14 conditions. In the most extreme case, matching responses following S_1 were almost 10 times as likely to be reinforced as matching responses following S_2. Figure 7 shows the averaged data from this condition (VI 17 s/VI 135 s) as well as data from conditions in which the reinforcement probabilities were equated (VI 30 s/VI 30 s). In order to fit these data, Equations 6a and 6b were combined into the general expression:

$$\frac{R_m}{R_m + R_n} = p' + (1 - p')\left\{\frac{\sqrt{r_m}}{\sqrt{r_m} + \sqrt{r_n}}\right\}, \quad (7)$$

where p' represents the two-parameter (i.e., constant ITI) version of Equation 1. The square roots of the obtained reinforcement probabilities were employed to accommodate undermatching (cf. Fantino & Davison, 1983). Indeed, when the exponent on obtained reinforcement frequency was allowed to vary freely, Harnett et al. found it to be almost exactly .50, averaged across pigeons and retention intervals. The pattern of results depicted in Figure 7 is in accordance with the predictions of the present analysis—performance following S_1 (high) exceeds performance following S_2
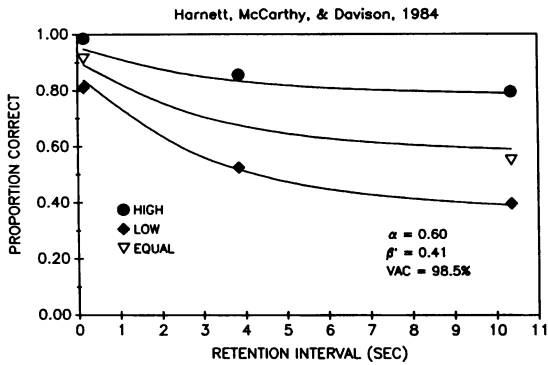
Fig. 7. Proportion of correct responses as a function of the retention interval for high, low, and equal reinforcement conditions. The data, which are averaged over 6 pigeons, were taken from Appendix C of Harnett, McCarthy, and Davison (1984; copyright 1984 by the Society for the Experimental Analysis of Behavior, Inc.). The solid curves represent the least squares fit of Equation 7 using the square roots of obtained reinforcement probabilities to accommodate undermatching. Adapted by permission.

(low), and performance in the equal condition falls approximately midway between them.

It may be that, as in the Delong and Wasserman (1981) experiment, some effect of differential outcomes over and above that predicted by the present analysis would have been observed in the experiment by Harnett et al. (1984) had they acquired a measurement at the 3.85-s delay for the equal condition (e.g., that data point might have fallen somewhere between the predicted value of .676 and the obtained 10.36-s delay value of .551). Nevertheless, in the absence of sample-specific responding, the present analysis appears to offer a reasonably good description of the results.

Directed Forgetting

A relatively new line of research has been concerned with establishing discriminative control over rehearsal processes in pigeons. The basic procedure, termed *directed forgetting*, involves the occasional presentation of a "forget cue" (F cue) during the retention interval that signals the cancellation of the choice phase. On an F-cued trial, the presentation of the sample (e.g., a red light) is followed by a brief presentation of an F cue (e.g., vertical lines). At the end of the retention interval, the comparison stimuli are simply omitted and the ITI commences. Other trials, which either contain a "remember cue" (R cue) or no rehearsal cue

at all during the retention interval, terminate normally with the presentation of the comparison stimuli. Because the F cue indicates that the comparison stimuli will not be presented on that trial, rehearsal of the sample's memorial representation is no longer necessary and should cease. If so, then the unexpected presentation of the comparison stimuli on F-cued probe trials should result in very poor performance because the memory trace of the sample will have dissipated entirely. Indeed, at least five studies have confirmed this prediction (Grant, 1981b; Maki & Hegvik, 1980; Maki, Olson, & Rego, 1981; Stonebraker & Rilling, 1981; Stonebraker, Rilling, & Kendrick, 1981). On the other hand, if the presentation of the F cue is withheld until the end of the retention interval, performance on probe trials should be much better. That is, despite the presumed cessation of rehearsal, the memory trace will not have fully disintegrated by the time the comparison stimuli are presented. Again, the results of several studies agree with this analysis (Grant, 1981b; Stonebraker & Rilling, 1981).

An interpretation of these findings within a delay-reduction framework is possible if it is assumed that a stimulus series (e.g., sample followed by rehearsal cue) will, under certain conditions, operate as a unified serial stimulus compound. A compound discriminative stimulus is usually defined to be two or more stimulus elements from different dimensions that are presented simultaneously (Reynolds, 1961). If a compound stimulus is presented prior to the opportunity to respond for reinforcement, however, it becomes possible to present the elements of the compound successively rather than simultaneously. The hypothesis entertained here is that a stimulus series may acquire the properties of a unified or "configural" stimulus compound (cf. Rescorla, 1973; Rescorla, Grau, & Durlach, 1985), depending upon the temporal parameters in effect.

A simple rule governing whether serial stimuli will operate separately or as a unit can be derived from D'Amato's (1973) theory of temporal discrimination. According to that view, the ease with which two previously presented stimuli can be distinguished depends upon the ratio of the ISI to the delay since the stimuli were presented. Adapting this principle to the present case, it might be hypothesized that when the ratio ISI/t is small, a stimulus

series may eventually acquire the properties of a stimulus compound. When that ratio is large, the two stimuli will maintain independent functions despite appearing together over the course of many trials.

According to this analysis, when the ISI is small relative to t (as is usually the case), a sample followed by a rehearsal cue should acquire the properties of a compound discriminative stimulus. The serial compound S|R cue has been associated reliably with reinforced responding to the comparison stimulus matching S and should therefore function much like an ordinary sample stimulus. By contrast, the series S|F cue has not been associated with reinforced responding to either comparison stimulus. Thus, this series would not be expected to exert significant discriminative control across a delay on a DMTS probe trial. As the ISI between the sample stimulus and F cue increases, however, the two stimuli should operate increasingly as individual stimulus elements. Considered individually, the sample, S, has been associated repeatedly with reinforced responding to the matching comparison stimulus as part of R-cued trials or baseline trials containing no rehearsal cues. Therefore, under these conditions, F-cued performance should be (and is) more accurate.

A simple quantitative measure of the functional delay reduction associated with the sample on F-cued trials is $(\text{ISI}/t)(d_r + \gamma)$ instead of the usual $d_r + \gamma$. When the ISI is very small relative to t (i.e., the F cue is presented early in the retention interval), the ratio ISI/t approaches zero and effectively cancels the reduction in delay to reinforcement associated with the sample. When the ISI is large relative to t (i.e., the F cue is presented late in the retention interval), the ratio ISI/t approaches one and drops out of the equation.

Figure 8 shows the results of the F-cue experiment conducted by Stonebraker and Rilling (1981). R-cued and F-cued performances, averaged over 3 pigeons, are plotted as a function of the ISI. Because the ITI was constant in this experiment, Equation 5 (modified to include the serial compound ratio for the F-cued case) was fitted to the data. For the R-cued case, the serial compound ratio was dropped from the equation by setting ISI equal to t. Hence, the equation predicts the same level of performance regardless of the actual ISI in effect. The quality of the fit suggests
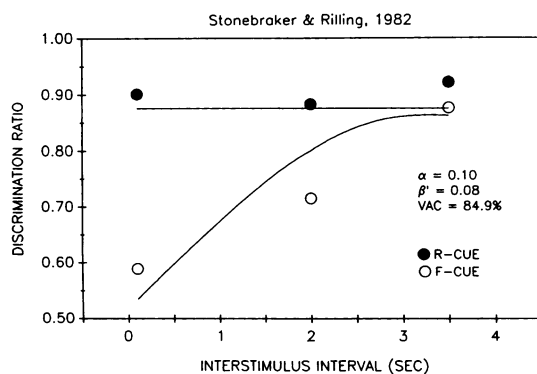

Fig. 8. Proportion of correct responses as a function of the interstimulus interval for the remember-cue (R-cue) and forget-cue (F-cue) conditions. The data are averaged over 3 pigeons and were estimated from Figure 2 of Stonebraker and Rilling (1981; copyright 1981 by the Psychonomic Society, Inc.). The solid curve represents the least squares fit of the modified version of Equation 5. Adapted by permission.

that the simple function, ISI/t, does a reasonably good job of capturing F-cued effects, although R-cued and F-cued performance does not converge as rapidly as predicted.

Several experiments have examined the effects of varying the size of the retention interval on F-cued probe trial performance when the ISI was held constant (Grant, 1981b; Maki & Hegvik, 1980; Maki et al., 1981). According to rehearsal theory, performance should deteriorate most rapidly on F-cued trials as the retention interval increases because the memorial representation is not being maintained actively by rehearsal. A similar prediction may be derived from the delay-reduction account. With a small delay, the ratio, ISI/t, is relatively large and the elements of the series S|F cue should, to a large extent, individually govern choice behavior. As the retention interval increases, however, the ratio decreases and the stimulus series should function increasingly as a serial stimulus compound (which is not associated with a reduction in delay to reinforcement). Therefore, the data should exhibit a detrimental effect on choice performance over and above that due to increasing the length of the retention interval alone. The results of all three experiments supported this prediction.

Figure 9 illustrates this phenomenon using the data reported by Maki et al. (1981) as well as the fit provided by Equation 5. The data from this experiment were rather variable
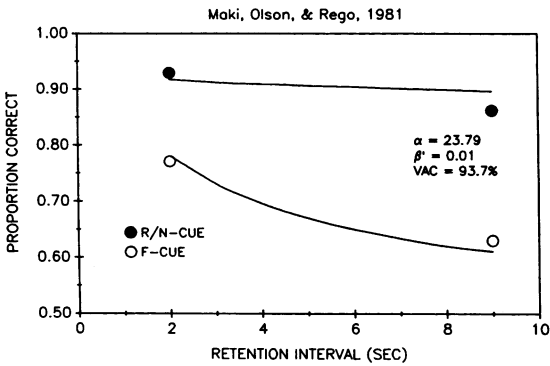
Fig. 9. Proportion of correct responses as a function of the retention interval for the combined remember-cue/no-cue (R/N-cue) and forget-cue (F-cue) conditions. The data are averaged over 6 pigeons and were taken from Table 2 of Maki, Olson, and Rego (1982; copyright 1981 by the Psychonomic Society, Inc.). The solid curve represents the least squares fit of the modified version of Equation 5. Adapted by permission.

when considered separately for different DMTS samples (which consisted either of brief food presentations or a brief blackout). Nevertheless, the results were averaged over the two sample types for purposes of this fit. Further, R-cued data and no-cue (N-cue) data were combined because performance on these two trial types did not differ. Although the empirical finding itself is not firmly established on the basis of these variable data, the figure clearly shows that Equation 5 predicts a more rapid decline for the F-cue case as the retention interval increases.

Surprisingness Procedures

Another line of rehearsal research has been concerned with determining the effects of "surprising" sample stimuli on the rehearsal process. The research is based on a theory proposed by Wagner (1976, 1978) known as *priming theory*, which maintains that unexpected, or surprising, sample stimuli will receive more rehearsal (and therefore be remembered longer) than expected, or unsurprising, stimuli. In a DMTS experiment, stimuli are rendered surprising by employing samples composed of two serial elements and occasionally arranging them in an unexpected way (Grant, Brewster, & Stierhoff, 1983; Maki, 1979). For example, one sample series may consist of an S+ followed by 3-s access to food, whereas the other sample series may consist of an S− followed by a 3-s blackout. Following

the presentation of the sample stimulus series, red and green choice stimuli are presented. A response to red is reinforced following samples of S+|food, and a response to green is reinforced following samples of S−|no food (blackout).

The question is how to make the samples of food and no food surprising in order to test priming theory. One way might be simply to switch the first stimulus in the sample series (e.g., S−|food instead of S+|food). Although this manipulation may indeed render the presentation of the food sample rather surprising, it also sets up a competition between S− and food for control over choice responding. That is, S− (usually followed by no food) has been associated with reinforced responding to the green choice stimulus, whereas food (usually preceded by S+) has been associated with reinforced responding to the red choice stimulus. The competition for discriminative control might overshadow any effects of surprisingness.

An alternative approach is to use a different set of stimuli (CS+ and CS−) to render the presentation of food or no-food samples surprising. These stimuli can be associated reliably with food or no food on separate discrimination trials that do not involve the red and green choice stimuli in any way. Thus, a surprising probe trial might be initiated by CS−|food and an expected probe trial by CS+|food. Following some delay, the red and green comparison stimuli would be presented. Presumably, a sample of food following CS− would be rather surprising and thus receive considerable rehearsal, but the same sample following C+ would be in accordance with expectations and thus not receive much retention interval processing. In agreement with the predictions of priming theory, Grant et al. (1983) found that performance on surprising probe trials was significantly better than performance on expected probe trials.

It might be possible to make sense of these findings without relying on the notion of rehearsal. Indeed, when the data are examined from a serial compound perspective, they appear to resemble the directed forgetting data rather closely. During baseline trials in Grant et al.'s (1983) experiment, for example, S+|food and S−|no food were always followed, after some delay, by the presentation of the comparison stimuli. By contrast,

CS+ | food and CS− | no food (i.e., the discrimination trials) were never followed by the presentation of the comparison stimuli. Thus, from one perspective, S+ and S− served as R cues, and CS+ and CS− served as F cues. The main difference between this procedure and an ordinary directed forgetting arrangement is that the rehearsal cues precede the presentation of the sample.

From this point of view, expected probe trials (initiated by CS+ | food or CS− | no food) are exactly analogous to F-cued probe trials from the directed forgetting literature. In both cases, the serial compounds are followed by the presentation of red and green comparison stimuli on probe trials only. On the other hand, baseline trials (initiated by S+ | food or S− | no food) are exactly analogous to R-cued trials from the directed forgetting literature in that the serial samples are always followed by the presentation of the choice stimuli. Based on the same reasoning as before, one would expect that performance on trials initiated by CS+ or CS− (F-cued trials) should be less accurate than performance on trials initiated by S+ or S− (R-cued trials). Indeed, this prediction conforms to the experimental results. Incidentally, on both kinds of trial, the sample of food or no food should be fully expected and should, according to a strong interpretation of priming theory, produce an equal level of (poor) performance.

Surprising probe trials (initiated by CS+ | no food or CS− | food) do not correspond directly to either F-cued or R-cued trials. Prior to the probe test, the stimuli were never paired (and thus should not function as serial compounds). As individual elements, CS+ and CS− have never been followed by reinforced responding to either choice stimulus. The food and no-food stimuli, on the other hand, have often been followed by reinforced responding to one of the choice stimuli. Therefore, one might expect a surprising stimulus series to result in reasonably accurate responding, with perhaps some decrement in performance relative to baseline due to the novelty of the arrangement. Grant et al. (1983) found that performance on surprising probes was somewhat less accurate than performance on baseline trials at delays of 0 and 5 s but about equal when the retention interval was 10 s.

The present serial compound analysis of the rehearsal literature suggests one way of integrating data from the directed forgetting and priming literatures. Perhaps more importantly, the analysis suggests connections to other experimental procedures not easily interpreted in terms of rehearsal. For example, White (1974) exposed pigeons to a DMTS procedure in which two samples were presented successively prior to the presentation of the choice stimuli. If the two samples were the same (S1 | S1 or S2 | S2), a response to the left choice stimulus was reinforced, but if they were different (S1 | S2 or S2 | S1) a response to the right choice stimulus was reinforced. Note that in this procedure the individual stimulus elements do not define the correct choice response, but the stimulus compounds do. Therefore, if the ISI is increased (such that the serial stimuli no longer operate in compound), performance should deteriorate. Indeed, that was the main finding of the experiment.

It may be possible to modify White's (1974) experiment in such a way that a counterintuitive prediction of the serial compound hypothesis may be tested. If a reasonably high level of performance could be achieved with a moderate delay interval in effect, then performance should actually deteriorate as the second stimulus is moved farther into the retention interval. In other words, as the ISI increases, control should increasingly be determined by the individual elements of the series. Under those conditions, performance should decline toward chance.

CONCLUSIONS

One reason for pursuing a descriptive analysis of the delayed response was to facilitate a connection between DMTS performance and other areas of behavioral research, such as choice and conditioned reinforcement. With regard to conditioned reinforcement, the present analysis suggests (quantitative consideration aside) a natural and intuitive principle: A stimulus that signals a shorter wait to reinforcing events is likely to govern responding across a longer temporal distance. An unanswered empirical question is whether the trial outcome stimuli, which are intermittently associated with reinforcement, constitute the critical reinforcing event (as assumed in this article) or whether end-of-trial reinforcement probabilities are the critical determinant. A similar question has received some attention

in the literature on autoshaping (e.g., Gibbon, Farrell, Locurto, Duncan, & Terrace, 1980).

The serial compound analysis attempts to bring an established concept to bear on procedures involving serial samples. Typically, stimuli are considered to operate in compound only when the elements are presented simultaneously. The data from the rehearsal literature suggest that it may be possible to extend this conception of stimulus compounds to the serial case. The relationship between the present account of serial stimulus compounds and the extensive literature on compound versus element samples in pigeons represents a potentially productive area of future research (Grant & MacDonald, 1986; Maki, Riley, & Leith, 1976; Roberts & Grant, 1978; Riley, 1984). That literature has generally found that matching with element samples is superior to matching with compound samples. An obvious question to investigate is whether the principle holds true when the elements of a compound are presented successively rather than simultaneously.

A second reason for pursuing an empirical analysis of DMTS performance was to provide a concise summary of the facts to be explained by those who prefer to construct more theoretical models of short-term memory. Without a clear understanding of the relevant empirical principles, constraints on hypothetical models of memory are few. Indeed, as Grant (1981a) has acknowledged, some contemporary theories of nonhuman short-term memory have grown in complexity to the point where they are close to achieving immunity from empirical contradiction. Closer attention to the empirical rules that govern DMTS performance may help to avoid this predicament while preserving the heuristic properties of cognitive theories.

Although it seems important to recognize the mutual benefits of contrasting approaches, it seems equally important not to confuse the two philosophies (cf. Williams, 1986). A functional approach, which seeks to identify ever-broadening empirical laws of behavior, is clearly different from a cognitive approach, which seeks to elucidate underlying memorial processes. Recognizing this difference and maintaining distinct modes of investigation may be the most reasonable approach to the analysis of memory in all species of animal, including humans (Wixted & McDowell, 1989).

REFERENCES

Blough, D. S. (1959). Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior, 2*, 151–160.

Branch, M. N. (1977). On the role of "memory" in the analysis of behavior. *Journal of the Experimental Analysis of Behavior, 28*, 171–179.

Brodigan, D. L., & Peterson, G. B. (1976). Two-choice conditional discrimination performance of pigeons as a function of reward expectancy, prechoice delay, and domesticity. *Animal Learning & Behavior, 4*, 121–124.

Buchanan, J. P., Gill, T. V., & Braggio, J. T. (1981). Serial position and clustering effects in a chimpanzee's "free recall." *Memory & Cognition, 9*, 651–660.

Catania, A. C. (1984). *Learning* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

D'Amato, M. R. (1973). Delayed matching and short-term memory in monkeys. In G. H. Bower (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Vol. 7, pp. 227–269). New York: Academic Press.

Davison, M. C., & Tustin, R. D. (1978). The relation between the generalized matching law and signal-detection theory. *Journal of Experimental Analysis of Behavior, 29*, 331–336.

DeLong, R. E., & Wasserman, E. A. (1981). Effects of differential reinforcement expectancies on successive matching-to-sample performance in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 7*, 394–412.

Dinsmoor, J. A. (1950). A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *Journal of Experimental Psychology, 40*, 458–472.

Edhouse, W. V., & White, K. G. (1988). Sources of proactive interference in animal memory. *Journal of Experimental Psychology: Animal Behavior Processes, 14*, 56–70.

Edwards, C. A., Jagielo, J. A., Zentall, T. R., & Hogan, D. E. (1982). Acquired equivalence and distinctiveness in matching to sample by pigeons: Mediation by reinforcer-specific expectancies. *Journal of Experimental Psychology: Animal Behavior Processes, 8*, 244–259.

Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313–339). Englewood Cliffs, NJ: Prentice-Hall.

Fantino, E., & Davison, M. (1983). Choice: Some quantitative relations. *Journal of the Experimental Analysis of Behavior, 40*, 1–13.

Fantino, E., & Logan, C. A. (1979). *The experimental analysis of behavior: A biological perspective.* San Francisco, CA: Freeman.

Gaffan, D. (1977). Recognition memory after short retention intervals in fornix-transectioned monkeys. *Quarterly Journal of Experimental Psychology, 29*, 577–588.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279–325.

Gibbon, J., Farrell, L., Locurto, C. M., Duncan, H. J., & Terrace, H. S. (1980). Partial reinforcement in autoshaping with pigeons. *Animal Learning & Behavior, 8*, 45–59.

Grant, D. S. (1975). Proactive interference in pigeon short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes, 1,* 207–220.

Grant, D. S. (1981a). Short-term memory in the pigeon. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 227–256). Hillsdale, NJ: Erlbaum.

Grant, D. S. (1981b). Stimulus control of information processing in pigeon short-term memory. *Learning and Motivation, 12,* 19–39.

Grant, D. S., Brewster, R. G., & Stierhoff, K. A. (1983). "Surprisingness" and short-term retention in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 63–79.

Grant, D. S., & MacDonald, S. E. (1986). Matching to element and compound samples in pigeons: The role of sample coding. *Journal of Experimental Psychology: Animal Behavior Processes, 12,* 160–171.

Harnett, P., McCarthy, D., & Davison, M. (1984). Delayed signal detection, differential reinforcement, and short-term memory in the pigeon. *Journal of the Experimental Analysis of Behavior, 42,* 87–111.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4,* 267–272.

Honig, W. K. (1987). Memory interval distribution effects in pigeons. *Animal Learning & Behavior, 15,* 6–14.

Jans, J. E., & Catania, A. C. (1980). Short-term remembering of discriminative stimuli in pigeons. *Journal of the Experimental Analysis of Behavior, 34,* 177–183.

Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology.* New York: Appleton-Century-Crofts.

Kendrick, D. F., Rilling, M. E., & Denny, M. R. (Eds.). (1986). *Theories of animal memory.* Hillsdale, NJ: Erlbaum.

MacDonald, S. E., & Grant, D. S. (1987). Effects of signaling retention interval length on delayed matching-to-sample in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 13,* 116–125.

MacPhail, E. M. (1980). Short-term visual recognition memory in pigeons. *Quarterly Journal of Experimental Psychology, 32,* 521–538.

Maki, W. S. (1979). Pigeons' short-term memories for surprising vs. expected reinforcement and nonreinforcement. *Animal Learning & Behavior, 7,* 31–37.

Maki, W. S., & Hegvik, D. K. (1980). Directed forgetting in pigeons. *Animal Learning & Behavior, 8,* 567–574.

Maki, W. S., Moe, J. C., & Bierley, C. M. (1977). Short-term memory for stimuli, responses, and reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 156–177.

Maki, W. S., Olson, D., & Rego, S. (1981). Directed forgetting in pigeons: Analysis of cue functions. *Animal Learning & Behavior, 9,* 189–195.

Maki, W. S., Riley, D. A., & Leith, C. R. (1976). The role of test stimuli in matching to compound samples by pigeons. *Animal Learning & Behavior, 4,* 13–21.

McCarthy, D., & White, K. G. (1987). Behavioral models of delayed detection and their application to the study of memory. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 29–54). Hillsdale, NJ: Erlbaum.

Nelson, K. R., & Wasserman, E. A. (1978). Temporal factors influencing the pigeon's successive matching-to-sample performance: Sample duration, intertrial interval, and retention interval. *Journal of the Experimental Analysis of Behavior, 30,* 153–162.

Peterson, G. B., Wheeler, R. L., & Armstrong, G. D. (1978). Expectancies as mediators in the differential-reward conditional discrimination performance of pigeons. *Animal Learning & Behavior, 6,* 279–285.

Rescorla, R. A. (1973). Evidence for "unique stimulus" account of configural conditioning. *Journal of Comparative and Physiological Psychology, 85,* 331–338.

Rescorla, R. A., Grau, J. W., & Durlach, P. J. (1985). Analysis of the unique cue in configural discriminations. *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 356–366.

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

Riley, D. A. (1984). Do pigeons decompose stimulus compounds? In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 333–350). Hillsdale, NJ: Erlbaum.

Roberts, W. A. (1972). Short-term memory in the pigeon: Effects of repetition and spacing. *Journal of Experimental Psychology, 94,* 74–83.

Roberts, W. A. (1980). Distribution of trials and intertrial retention in delayed matching to sample with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 6,* 217–237.

Roberts, W. A., & Grant, D. S. (1976). Studies of short-term memory in the pigeon using the delayed matching to sample procedure. In D. L. Medin, W. A. Roberts, & R. T. Davis (Eds.), *Processes of animal memory* (pp. 79–112). Hillsdale, NJ: Erlbaum.

Roberts, W. A., & Grant, D. S. (1978). Interaction of sample and comparison stimuli in delayed matching to sample with the pigeon. *Journal of Experimental Psychology: Animal Behavior Processes, 4,* 68–82.

Roberts, W. A., & Kraemer, P. J. (1981). Recognition memory for lists of visual stimuli in monkeys and humans. *Animal Learning & Behavior, 9,* 587–594.

Roberts, W. A., & Kraemer, P. J. (1982). Some observations of the effects of intertrial interval and delay on delayed matching to sample in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 8,* 342–353.

Roberts, W. A., & Kraemer, P. J. (1984). Picture memory in monkeys. *Canadian Journal of Psychology, 38,* 218–236.

Roberts, W. A., & Kraemer, P. J. (1984). Memory for lists of spatial events in the rat. *Learning and Motivation, 10,* 313–336.

Sands, S. F., & Wright, A. A. (1980a). Primate memory: Retention of serial list items by a rhesus monkey. *Science, 209,* 938–940.

Sands, S. F., & Wright, A. A. (1980b). Serial probe recognition performance by a rhesus monkey and a human with 10- and 20-item lists. *Journal of Experimental Psychology: Animal Behavior Processes, 6,* 386–396.

Santi, A. (1984). The trial spacing effect in delayed matching-to-sample by pigeons is dependent upon the illumination condition during the intertrial interval. *Canadian Journal of Psychology, 38,* 154–165.

Santi, A., & Roberts, W. A. (1985). Reinforcement expectancy and trial spacing effects in delayed matching-to-sample by pigeons. *Animal Learning & Behavior, 13,* 274–284.

Santiago, H. C., & Wright, A. A. (1984). Pigeon memory: *Same/different* concept learning, serial probe recognition acquisition, and probe delay effects on the serial-position function. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 498–512.

Shimp, C. P. (1976). Short-term memory in the pigeon: Relative recency. *Journal of the Experimental Analysis of Behavior, 25,* 55–61.

Shimp, C. P., & Moffitt, M. (1974). Short-term memory in the pigeon: Stimulus-response associations. *Journal of the Experimental Analysis of Behavior, 22,* 507–512.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century.

Squires, N., & Fantino, E. (1971). A model for choice in simple concurrent and concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior, 15,* 27–38.

Stonebraker, T. B., & Rilling, M. (1981). Control of delayed matching-to-sample performance using directed forgetting techniques. *Animal Learning & Behavior, 9,* 196–201.

Stonebraker, T. B., Rilling, M., & Kendrick, D. F. (1981). Time dependent effects of double cuing in directed forgetting. *Animal Learning & Behavior, 9,* 385–394.

Thompson, R. K. R., & Herman, L. M. (1977). Memory for lists of sounds by the bottle-nosed dolphin: Convergence of memory processes with humans? *Science, 195,* 501–503.

Urcuioli, P., & Honig, W. K. (1980). Control of choice in conditional discriminations by sample-specific behaviors. *Journal of Experimental Psychology: Animal Behavior Processes, 6,* 251–277.

Wagner, A. R. (1976). Priming in STM: An information-processing mechanism for self-generated or retrieval-generated depression in performance. In T. J. Tighe & R. N. Leaton (Eds.), *Habituation: Perspectives from child development, animal behavior, and neurophysiology* (pp. 95–128). Hillsdale, NJ: Erlbaum.

Wagner, A. R. (1978). Expectancies and the priming of STM. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 177–209). Hillsdale, NJ: Erlbaum.

Wasserman, E. A., Grosch, J., & Nevin, J. A. (1982). Effects of signaled retention intervals on pigeon short-term memory. *Animal Learning & Behavior, 10,* 330–338.

White, K. G. (1974). Temporal integration in the pigeon. *British Journal of Psychology, 65,* 437–444.

White, K. G. (1985). Characteristics of forgetting functions in delayed matching to sample. *Journal of the Experimental Analysis of Behavior, 44,* 15–34.

White, K. G., & McKenzie, J. (1982). Delayed stimulus control: Recall for single and relational stimuli. *Journal of the Experimental Analysis of Behavior, 38,* 305–312.

Wilkie, D. M. (1984). Pigeons' spatial memory: IV. Effects of intertrial manipulations on delayed matching of key location. *Canadian Journal of Psychology, 38,* 178–195.

Wilkinson, L. (1988). *SYSTAT: The system for statistics.* [Computer program]. Evanston, IL: SYSTAT.

Williams, B. A. (1986). On the role of theory in behavior analysis. *Behaviorism, 14,* 111–124.

Wixted, J. T., & McDowell, J. J. (1989). Contributions to the functional analysis of single-trial free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 685–697.

Wright, A. A., Santiago, H. C., & Sands, S. F. (1984). Monkey memory: *Same/different* concept learning, serial probe acquisition, and probe delay effects. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 513–529.

Zentall, T. R., Hogan, D. E., Howard, M. M., & Moore, B. S. (1978). Delayed matching in the pigeon: Effect on performance of sample-specific observing responses and differential delay behavior. *Learning and Motivation, 9,* 202–218.