

# Matching behavior: Some methodological problems<sup>1</sup>

M. R. D'AMATO and WARD O'NEILL,  
RUTGERS UNIVERSITY, New  
Brunswick, New Jersey 08903

*Two classes of problems associated with specifying the controlling stimuli in matching experiments were discussed. The first class—control of matching behavior by cues other than the standard stimuli—was demonstrated by the results of a pseudomatching experiment. The second class related to the "effective" sample or standard stimulus set and to changes in the S's matching strategy that may occur in a matching experiment.*

The purpose of this note is to call attention to certain methodological problems that, while not unique to investigations of matching behavior, are especially prone to arise within that context. The number of publications relating to matching behavior in animals and in humans has been on the rise during the last several years, and examination of many of these reports does not leave one with the feeling that the methodological issues discussed below are sufficiently appreciated by all contemporary investigators in this area.

The assumption commonly made in matching studies is that if a S performs successfully it is because of the contingencies imposed by the E with respect to the "standard" and comparison stimuli. In matching-to-sample tasks, for example, above-chance performance is assumed to be due to the control exerted by the sample stimuli. Rarely is this assumption tested directly by evaluating the controlling function of stimuli other than the sample. However, because of the restrictions on randomization usually imposed in the programming of stimuli in most experiments, there are a number of potential cues that can gain control over S's behavior and lead to significant criterial performance. Among these restrictions are the presentation of equal numbers of the stimuli that comprise the sample set, the balancing of stimuli that serve as comparison stimuli, and restrictions on the number of times that a specific sample occurs in succession. Moreover, the particular sequence of trials employed, if used more than once, can provide discriminative cues.

In many situations the configuration of stimuli that appears at the time of choice can also serve to control S's performance. For example, in our experimental apparatus there are four projectors upon which comparison stimuli can appear. In a


matching-to-sample task employing two comparison stimuli, S might very well learn that when a triangle appears on the upper-left projector and a square on the upper-right projector, the square is correct, but when the square appears on the lower-left projector and the triangle on the upper-right projector, the triangle is correct. In the present experiment, we evaluated whether or not the potential discriminative cues just described are sufficient to lead to significant "matching" performance in a pseudo matching-to-sample task.

## METHOD

### Subjects

The three Ss of this experiment, Daisy, Pee Wee, and Babe, were adult capuchin monkeys (*Cebus apella* and *albifrons*) with previous experience but not with matching to sample.

### Apparatus

Two Lehigh Valley monkey chambers (Model 1317) were modified by adding five Industrial Electronic Engineers, Inc., inline stimulus projectors arranged in the pattern . Each projector was faced with a plastic key, which served as the response mechanism. To be counted as a response or to be recorded, a key press had to be maintained for a minimum of 0.4 sec. A microswitch, with which S could initiate a trial, was centered below the five projectors. CIBA 190-mg banana pellets served as rewards and were delivered to a small cup located on the right wall 11 cm from the stimulus panel.

Programming of stimuli was accomplished by a block tape reader described previously (D'Amato, 1965). The form stimuli appeared as white figures on black backgrounds. All forms but the dot were composed of lines approximately 1.5 mm thick; their overall height (and width) was 7 mm in diam. The five stimuli used were a square, a triangle, a vertical line, a dot, and a red field that illuminated a circular area approximately 25 mm in diam.

### Procedure

All Ss were trained on a pseudo matching-to-sample task. On each trial the "sample" was the dot stimulus, and the "comparison" stimuli consisted of two of the remaining four stimuli. One of the two comparison stimuli was designated as "correct" and the other as "incorrect." Two sequences of 24 trials each were patterned after comparable sequences of trials used in actual matching-to-sample experiments, the only difference being that

in the present sequences the dot was always substituted for the sample stimulus. One of the two sequences was quasirandomly chosen for each of the 24 experimental sessions and each sequence served for 12 of the 24 sessions. After completion of the 24 sessions, four new sequences of 24 trials each (employing the same stimuli as in the first two sequences) were constructed and eight additional daily sessions were given, with two sessions on each sequence.

A typical trial proceeded as follows. After completing a fixed-ratio 15 on the microswitch, the dot stimulus appeared on the center key and remained there until S pressed the key for a minimum duration of 0.3 sec. The dot then disappeared and two comparison stimuli immediately appeared on two of the four outer keys. (We refer to this condition as 0-delay matching to sample.) The comparison stimuli remained present until S responded to one of them. If the "correct" stimulus was responded to, one reward pellet was delivered, and S was free to initiate the next trial at any time. "Incorrect" responses were followed by a 1-min timeout, signaled by the dimming of the overhead houselight. At the end of the timeout, S was free to initiate the next trial. During the experiment Ss were maintained at approximately 80% of their full-ration body weight.

## RESULTS

Figure 1 presents the accuracy function for each S, expressed as percentage of correct responses per experimental session. Although there is considerable variability in day-to-day performance, the data of the first 24 sessions reveal unmistakable evidence of above-chance performance in the case of two of the three Ss. Pee Wee averaged 61% correct responses over the entire first 24 sessions, which is considerably in excess of chance expectation ( $t = 6.42$ ,  $df = 23$ ,  $p < .001$ ). Although Daisy's performance over the first 24 sessions is not different from chance expectation, she showed definite signs of learning during the latter sessions. She averaged 65% correct responses over Sessions 18-24, which is considerably in excess of chance expectation ( $t = 2.68$ ,  $df = 6$ ,  $t < .05$ ). Babe, the one *albifrons* in the study, showed no sign of "acquiring" the pseudo matching-to-sample task, averaging 51% correct responses over the entire 24 sessions and 53% over the last seven sessions.

Analysis of the performance achieved with each of the two trial sequences revealed only small and insignificant

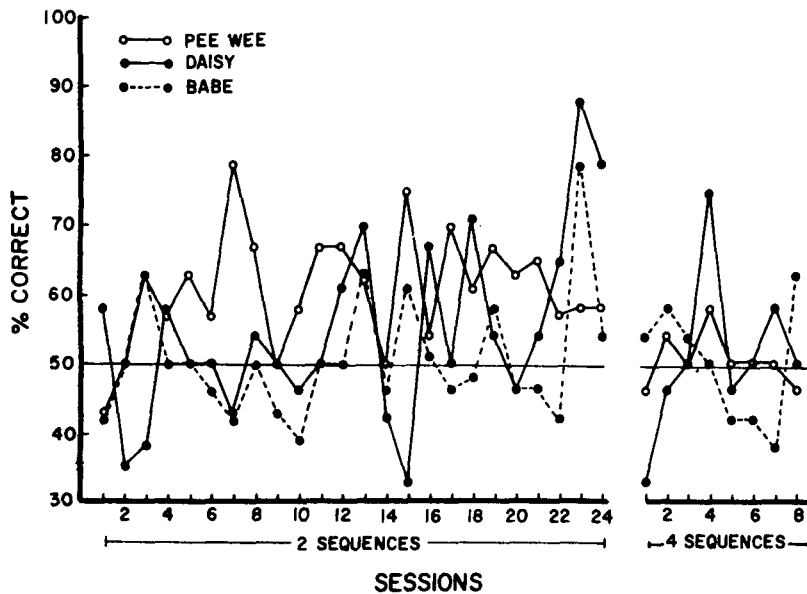


Fig. 1. Accuracy functions of the three Ss over the 32 sessions (24 trials/session) of the pseudomatching experiment. During Sessions 1-24, two different sequences of trials were employed; both Pee Wee and Daisy showed significant above-chance "matching" performance during this phase. Four new sequences were introduced in the subsequent eight session, during which the performance of all Ss remained at chance levels.

differences, suggesting that whatever the controlling cues might have been they occurred equally in both trial sequences.

It is evident from Fig. 1 that when four new trial sequences were administered, the performance of all Ss returned to, or remained at, chance levels. The average percentage of correct responses during this phase of the experiment was 51, 51, and 49 for Pee Wee, Daisy, and Babe, in that order.

#### DISCUSSION

The results show clearly that in matching experiments stimuli other than the sample, or standard stimulus, can gain control over S's behavior and lead to above-chance performance. In the present experiment these stimuli appear to be related to the configuration of the comparison stimuli and to trial sequence. To avoid the development of response biases, in our matching experiments we routinely equate the number of times each stimulus of the stimulus set serves as sample. We also pair all stimuli equally often, a control measure, which, when it cannot be performed in a single trial sequence, is accomplished over two or more sequences. Since these restrictions on randomization were imposed on all of the trial sequences employed in the experiment, and since all Ss responded at chance levels during the last eight sessions, in which four new trial sequences were substituted for the earlier two, it appears that the stimuli controlling S's behavior during Sessions 1-24 probably arose from a different source. The most likely candidates are configuration of comparison stimuli and trial sequence. We have already illustrated the nature of the first category of cues, though it ought to be added here

that Farrer (1967) has reported that chimpanzees can learn a series of 24 visual discriminations among four stimuli simply on the basis of stimulus configuration. It is also possible that S can learn about certain features of the trial sequence; perhaps that in one trial sequence, red is more likely to be correct at the beginning of a day's trials than during the middle, or that after a particularly distinctive comparison stimulus configuration, the correct sequence is triangle, square, red. Clearly these two sources of controlling stimuli may act together, as in the last illustration, to produce performance that is considerably in excess of chance expectation.

Even with only two response keys available during the choice period, which is the context in which very many matching experiments take place, trial sequence and stimulus configuration can provide controlling stimuli for S's behavior, particularly where a fairly large number of different standard stimuli are employed. The indicated conclusion is that the investigator of matching behavior should always be on guard to see that S's behavior is under the control of the standard stimuli rather than extraneous cues, a caveat that is especially important when variables are manipulated in an experiment that strain S's matching ability. In an easy matching task the standard stimulus provides a potent controlling stimulus for S's discrimination behavior, and consequently the latter is likely to be based on these stimuli rather than on extraneous cues. However, if the potency of the standard stimuli is reduced, for example, by using a relatively long delay in a delayed matching-to-sample task, it is conceivable that control of S's behavior will shift from

the standard stimuli to other cues present in the experimental situation (cf. Sidman, 1969). Such shifts in controlling stimuli are consistent with contemporary attention theory, in which it is assumed that the degree to which stimuli come to control behavior depends, among other factors, on their "dominance" or "saliency." Consequently, a decline in matching performance at long delay intervals may not reflect solely a decline in the control of sample stimuli, but also be indicative of a shift of control to extraneous stimuli. The investigator can evaluate this possibility either by introducing pseudomatching trials, or by introducing novel trial sequences and stimulus configurations.

#### "Nominal" vs "Effective" Sample Set

Problems of specifying the controlling stimuli in matching experiments arise in another way, as indicated by the following result. Two of our Ss who had considerable experience in delayed matching-to-sample in a two-choice situation involving the sample set of red, vertical line, triangle, and square, were ultimately capable of performance levels of about 75% correct with delay intervals of 60 sec. In the process of shifting these Ss from two to four comparison stimuli, we reduced the delay interval to 2 sec, and very shortly their performance increased to the 90% correct level. When the number of comparison stimuli was increased to four, however, their performance deteriorated markedly (to 55% correct). Moreover, it did not improve significantly even after more than 10 sessions of training. Analysis of their errors showed that for each animal one of the four sample stimuli exerted far greater control than the others over S's matching behavior. One S, for example, rarely made an error when red served as the sample stimulus and he seldom chose red when it had not been the sample. When the sample was nonred, S's performance was essentially at a chance level, insofar as the nonred comparison stimuli were concerned. It appears that this animal collapsed the sample set from four to two, namely, to red and nonred, responding appropriately to the redefined set. We may

Table 1  
Expected Percentage of Correct Responses in a Four-Sample Matching-to-Sample Task as a Function of Number of Comparison Stimuli and Number of Controlling Samples

Number of Controlling Samples	Number of Comparison Stimuli		
	2	3	4
0	50	33	25
1	75	58 1/3	50
2	91 2/3	83 1/3	75
3	100	100	100

say that his behavior was under the control of a single sample stimulus (red).

Simple calculation shows that if S responds consistently in the manner described (viz., when the controlling stimulus appears as a comparison stimulus it is *always* chosen if it is the sample of that trial, *never* if it is not), his expected percentage of correct responses with four samples and only two comparison stimuli is 75. If S partitions the sample set into three categories and responds consistently on that basis (i.e., his behavior is controlled by two of the sample stimuli), his expected percentage of correct responses reaches 91.67. With four comparison stimuli, on the other hand, knowledge of and consistent control by one sample stimulus in a set of four results in an expected performance level of only 50%; control by two samples raises this to only 75% (see Table 1).

One may also analyze the present situation in terms of the information that is available in the sample and in the comparison stimuli. When all four samples are presented as comparison stimuli, there is no information present at the time of choice and 2 bits are required to choose the correct comparison stimulus. In each presentation of a sample there are potentially 2 bits of information available, which S can extract and store over the

delay interval. When the number of comparison stimuli are reduced to two, only 1 bit of information is required for a correct response on any given trial. The effect of collapsing the sample set into two—e.g., red and nonred—is to reduce the information available in each sample presentation from 2 bits to 1. Irrespective of how S dichotomizes the sample set, on 50% of the trials only 1 bit is necessary for a correct response (in the previous illustration, for example, red occurs as a comparison stimulus on 50% of the trials). On the other 50% of the trials the bit of information possessed by S is irrelevant, and his performance will be at a chance level. His overall performance, however, will average 75% correct. Clearly, this strategy does not work as well with four comparison stimuli, since if S possesses on the average only 1 bit of information at the time of choice, his performance will be 50%. The suggestion is that reducing the information available at the time of choice—i.e., increasing the number of comparison stimuli—should serve to increase control of matching behavior by the sample stimuli.

Indication of the control exerted by each sample (or standard) stimulus can be gotten from S's "confusion" matrix, i.e., the sample-response matrix, which is also of value in detecting changes in S's matching strategy. Such changes are not unlikely when variables are investigated that strain S's matching ability. For example, in the case of the two Ss just discussed, it appears that at short delay intervals their matching behavior was under the control of more than one sample stimulus (with short delays they were capable of performance levels of at least 95% correct). However, when the delay interval was shifted abruptly to long intervals, the sample stimuli lost control over their behavior. The Ss then adopted

the strategy of partitioning the sample set into two, which, in conjunction with the information available in the two-choice situation, allowed them a reasonable density of reinforcement. It is interesting, in view of this interpretation, that one S whom we shifted from two to four comparison stimuli in exactly the same fashion as we attempted to shift the first two, showed almost no decrement in his performance. This animal had had far less experience with matching at long delay intervals.

In conclusion, the investigator of matching behavior must be wary about accepted above-choice performance as indicative of control by the standard stimuli. Where conclusions must be based on the absolute level of S's performance, it is wise to use many different stimulus configurations and trial sequences or, in some cases to introduce pseudomatching trials. Where control by the sample (standard) stimuli has been demonstrated, it may not be in direct congruence with the samples actually employed; the sample-response matrix provides information as to the nature of the controlling stimuli in the sample set and changes in S's matching strategy.

#### REFERENCES

- D'AMATO, M. R. Direct programming of multiple stimuli—the tape block reader. *Journal of the Experimental Analysis of Behavior*, 1965, 8, 230.  
 FARRER, D. N. Picture memory in the chimpanzee. *Perceptual & Motor Skills*, 1967, 25, 305-315.  
 SIDMAN, M. Generalization gradients and stimulus control in delayed matching-to-sample. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 745-757.

#### NOTE

1. This research was supported by National Science Foundation Grant GB-7951.