

A model of selective perception: The effect of presenting alternatives before or after the stimulus

KENT GUMMERMAN†

The University of Texas, Austin, Tex. 78712

Selective perception, a concept supported by recent "before/after" studies, is usually thought to function by guiding the extraction of information from briefly viewed stimuli. However, recent findings from a related type of study are not in agreement, so a second interpretation of before/after data is developed here. This interpretation, based in part on a statistical decision model taken from signal detection theory, assumes that selective perception functions by guiding the perceptual process of comparing extracted stimulus information with internal representations of stimulus alternatives

Recent studies using Lawrence & Coles's (1954) before/after method have provided data that can be interpreted in terms of selective perceptual processing of brief visual stimuli (Egeth & Smith, 1967; Gummerman, 1971). This method may be briefly characterized as follows. On each trial, one member of a master stimulus set, *M*, containing *m* items, is presented. For every trial, a subset of *M* is selected by the *E*; this subset, *S*, contains *s* items ($s < m$) and always includes the presented item. The composition of *S* is made known to the *O*s either *before and after* the stimulus exposure (before condition) or only *after* the exposure (after condition). Thus, at the time the stimulus is being processed (i.e., during the exposure but not afterwards, since the masking pattern that should follow the stimulus presumably stops processing at the stimulus offset), the *O* operates either on the basis of the reduced stimulus set, *S* (on all before trials), or on the basis of the larger master set, *M* (on all after trials). In all cases, the *O* is told the composition of *S* before he actually responds, so his chance performance rate (assuming a forced-choice task) is always $1/s$.

Selective perceptual processing is indicated when performance is better on before trials than on after trials. Egeth & Smith (1967) suggested that their enhanced before scores were due to the fact that their *O*s could "determine which dimensions were likely to be effective for discriminating within a set of alternatives, and they also had the opportunity to use this knowledge in extracting information from the test picture. The observers in the After conditions could not know what kind of information to extract first; they got what they could from the flash, but there was an excellent

probability that ...it would not be adequate for discrimination [p. 549]." Similarly, Gummerman (1971) found a before vs after difference, with $s = 2$ and $m = 16$, and concluded that selective perception "appears to consist of 'instructions' as to which of the stimulus attributes are relevant and should be processed [p. 177]." In this latter study, no before enhancement occurred in a second condition, with $s = 2$ and $m = 4$. This outcome was taken to mean that processing on the basis of two stimuli (before condition) was not measurably more efficient than processing on the basis of four (after condition). The entire set of data from this experiment is shown in Table 1.

However reasonable these arguments may sound, it is now necessary to consider alternative interpretations of the data. New findings,¹ from an experiment similar in purpose to the before/after experiments, show that the process of extracting information from a stimulus is no more efficient (i.e., does not involve dimensions that are more likely to be task-relevant) when there are few stimulus alternatives than when there are many. In this experiment, the task was a type of "free response"—only one master set, *M*, was used in a session, and no trial-by-trial selection of subsets (before/after cueing) was used. The size of *M* varied from $m = 2$ to $m = 16$ across sessions, and the *O*s always knew its composition. Of course, since the guessing rate is much higher when $m = 2$ than when $m = 16$, the raw scores (percent correct) were greater when *m* was small; but, after the raw scores were adjusted for differences in guessing rate (using several methods, including ones based on signal detection theory, the constant-ratio rule, and the all-or-none guessing model), no superiority for $m = 2$ remained.

Since the selective perception model mentioned above predicts superiority for $m = 2$, another model has been formulated—one which draws heavily upon a well-established statistical decision model (Birdsall & Peterson, 1954; Green & Birdsall, 1964). Developed in the context of signal detection theory (see Egan & Clarke, 1966; Green & Swets, 1966), this model assumes that processing efficiency in the "free-response" task is invariant over changes in the value of *m* and that poor performance for high values of *m* results solely from the

Table 1
Proportion Correct Identifications as a Function of Cueing Condition and *m* (from Gummerman, 1971)

<i>m</i>	Cueing Condition	
	Before	After
4	.846	.846
16	.883	.729

*This work was supported in part by the National Science Foundation's University Science Development Program, Grant GU-1598. Dennis McFadden and Cynthia R. Gray read drafts of an early manuscript.

†Author's address: Department of Psychology, Mezes Hall 328, The University of Texas at Austin, Austin, Texas 78712.

large number of comparisons that must be made between extracted stimulus information and internal representations of the m alternatives. This model not only describes the free-response data well but also, we will show, serves as the basis for an extended model which nicely accounts for the data in Table 1.

First consider the decision process in a free-response situation. If M is completely specified for the O , there exists for each stimulus in M a corresponding representation in the O 's memory. The information obtained from a stimulus is assumed to be cross-correlated with every representation, producing m correlations on each trial. Ideally, one correlation would be high, and $m-1$ correlations would be very low. The process is complicated, however, by random error, which leads to the establishment of two *distributions* of correlations which, when appropriately transformed, will be normal with unit variance and mean equal to zero (for the $m-1$ correlations between the stimulus and noncorresponding representations) or to d' (for the correlation between the stimulus and its own representation). Then, assuming the O 's response is always based on the item in memory that correlated most highly with the stimulus, the probability of a correct response is equal to the probability that the correlation belonging to the distribution with mean equal to d' is larger than the largest of the $m-1$ correlations belonging to the distribution with mean equal to zero. This probability is equivalent to the probability that one drawing from a normal distribution with mean d' and unit variance is larger than the largest of $m-1$ drawings from a normal distribution with mean zero and unit variance. Birdsall & Peterson (1954) computed this probability for selected values of d' and m ; Elliott (1964) published more extensive tables. Green & Birdsall (1964) also discussed this type of model.

To apply the model to raw scores in a free-response task, one need only convert them to values of d' with Elliott's tables. If the model is correct, d' will not vary as a function of m , and this was the outcome of the recent study mentioned previously (see Note 1).

Now consider the task of an O in a before/after experiment. On all before trials, the subset S is identified before the stimulus exposure. Consequently, the O is free to respond with any of the items in S —he is essentially working in a free-response situation: for any given trial, he is concerned with a new master set of stimuli, M' , in which $m' = s$. This situation is unaffected by the fact that the composition of S changes from trial to trial. Thus, in the Gummerman (1971) study, all O s in before conditions were working in a free-response situation with $m' = s = 2$, both when $m = 4$ and when $m = 16$. Taking the mean of the two before scores in Table 1 (for $m = 4$ and $m = 16$) as the best estimate of the probability of a correct choice in a free-response task with $m' = 2$, we obtain .865. This value, according to Elliott's tables, corresponds to a d' of about 1.565.

The after condition is somewhat less

straightforward. Since S has not yet been revealed by the time stimulus processing is complete, the O must operate, at least temporarily, on the basis of the true master set, M . Assume, therefore, that the m correlations that would be made in a free-response condition also are made in the after condition, so that the O would be capable of making any one of m responses. It is reasonable, then, to construct the $m \times m$ stimulus-response confusion matrix that would be produced in this situation, assuming that the O 's sensitivity remains the same as in the before condition ($d' = 1.565$). From Elliott's tables, we find that the symmetrical 4×4 matrix corresponding to $d' = 1.565$ has .735 in all cells on the major diagonal and .088 in all other cells. The symmetrical 16×16 matrix corresponding to $d' = 1.565$ has .440 on the major diagonal and .037 elsewhere. Thus the O who is correct on .865 of the before trials (where $m' = 2$) would be correct, in a free-response task, on .735 of the trials when $m = 4$ and on .440 of the trials when $m = 16$. These lower scores follow directly from the increasing number of samples that are taken from the zero-mean distribution as m grows.

What happens, then, in an actual after task when knowledge of S follows the stimulus? Assume that the O makes the response corresponding to the largest correlation whenever that item appears in S ; but, whenever the item with the highest correlation is *not* represented in S , the O cannot give that response—he must guess at random between the items in S . Therefore, the three events listed below can occur in the after condition.

(a) The item with the highest correlation is contained in S and is, in fact, the stimulus presented. This event will occur on .735 of the trials when $m = 4$ and on .440 of the trials when $m = 16$. On all of these trials, the response will be correct.

(b) The item with the highest correlation is contained in S and is *not* the stimulus presented. This will occur on .088 of the trials when $m = 4$ and on .037 of the trials when $m = 16$. On all of these trials, the response will *not* be correct.

(c) The item with the highest correlation is *not* in S . This event will occur on $(2)(.088) = .176$ of the trials when $m = 4$ and on $(14)(.037) = .518$ of the trials when $m = 16$. Presumably, the O guesses on these trials and will be correct with probability .50.

Now, we can sum across these events to predict after performance. The O will be correct on $(.735)(1.00) + (.088)(.00) + (.176)(.50) = .823$ of the trials when $m = 4$; when $m = 16$, he will be correct on $(.440)(1.00) + (.037)(.00) + (.518)(.50) = .699$ of the trials. Comparing these two figures with the after data of Table 1, we find discrepancies of only .023 for $m = 4$ and .030 for $m = 16$.

It seems clear that the statistical decision model of Birdsall and Peterson, combined with restriction of the O 's responses according to the stimuli contained in S ,

can predict the after scores in Gummerman (1971) very nicely from the before scores. Presumably, this analysis would also be applicable to the data of Egeth & Smith (1967); however, these authors did not define the master set, M, for their Os, so it is not possible to make the calculations.

REFERENCES

- Birdsall, T. G., & Peterson, W. W. Probability of a correct decision in a choice among m alternatives. Electronic Defense Group, University of Michigan, Quarterly Progress Report No. 10, 1954.
- Egan, J. P., & Clarke, F. R. Psychophysics and signal detection. In J. B. Sidowski (Ed.), *Experimental methods and instrumentation in psychology*. New York: McGraw-Hill, 1966. Pp. 211-246.
- Egeth, H., & Smith, E. E. Perceptual selectivity in a visual recognition task. *Journal of Experimental Psychology*, 1967, 74, 543-549.
- Elliott, P. B. Tables of d' . In J. A. Swets (Ed.), *Signal detection and recognition by human observers*. New York: Wiley, 1964. Pp. 651-684.
- Green, D. M., & Birdsall, T. G. The effect of vocabulary size on articulation score. In J. A. Swets (Ed.), *Signal detection and recognition by human observers*. New York: Wiley, 1964. Pp. 609-619.
- Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- Gummerman, K. Selective perception and the number of alternatives. *The American Journal of Psychology*, 1971, 84, 173-179.
- Lawrence, D. H., & Coles, G. R. Accuracy of recognition with alternatives before and after the stimulus. *Journal of Experimental Psychology*, 1954, 47, 208-214.

NOTE

1. Gummerman, K., and Dykes, J., in preparation.

(Received for publication July 24, 1973.)

Formation of extinction sets in monkeys

ISAAC BEHAR*

U. S. Army Medical Research Laboratory, Fort Knox, Ky. 40121

Discrimination-sophisticated Old World monkeys received 108 two-choice problems, each containing 4 to 8 discrimination trials followed by 4 extinction trials for four animals and by 10 extinction trials for four animals. Following this, they received 108 additional problems with extinction trial length reversed. With 10 extinction trials per problem, an orderly increase in the efficiency of extinction occurred, with one-trial extinction occurring in a large proportion of problems. Virtually no tendency to extinguish occurred with four extinction trials per problem. There was a marked increase in responses to the formerly negative stimulus on the second extinction trials.

According to Harlow, the defining characteristic of a learning set is the change in *efficiency* with which learning takes place. It is a change from trial-and-error learning to "adapting to a changing environment . . . by seeming hypothesis and insight [Harlow, 1949, p. 51]." Learning sets typically are acquired when the animals are trained on an extended series of problems, all of comparable difficulty level, and each for a small number

of trials instead of to some criterion of mastery. Using these general operations, the formation of learning sets has been demonstrated for a wide variety of learning tasks, including object discrimination, reversal, oddity, matching to sample, and object alternation.

The present study represents an application of the principles of learning set formation to another aspect of learning performance, namely, that of extinction. While most previous studies of extinction have been concerned with variables that influence the *resistance* to extinction, this study is concerned with the *efficiency* or ease with which extinction can occur. Specifically, the purpose was to determine whether monkeys, given a long series of different extinction problems analogous to the paradigm used in the study of learning set formation, would transfer across extinction problems such that they would eventually extinguish in a single trial.

METHOD

Eight mature monkeys, four rhesus and four sooty manglebeys, all with excellent learning ability as demonstrated in a wide variety of learning tasks, were Ss. Although the specific test histories varied among animals, all had been trained on object-discrimination learning sets; only one had experience with reversal problems. Just prior to the present study, they were given 60 six-trial object-discrimination problems and, based upon that performance, were assigned to either of two groups equated in learning ability, each consisting of two rhesus and two manglebeys.

*Reprint requests should be sent to Isaac Behar, U.S. Army Aeromedical Research Laboratory, Fort Rucker, Ala. 36360.