Recency discrimination predicted from absolute lag judgments¹

ROBERT S. LOCKHART²
THE UNIVERSITY OF TORONTO

An experiment is described which tests the plausibility of the assumption that two-alternative forced-choice (2AFC) recency discrimination is based on an independent lag estimate of each item. 39 Ss made absolute recency judgments for items (common nouns) at varying lag intervals while an independent group of 74 Ss made 2AFC judgments for pairs of items at comparable lag combinations. It was found that the absolute judgment data generated accurate predictions of 2AFC performance.

In a two-alternative forced-choice (2AFC) recency discrimination task, two previously presented items are shown and S is required to state which of these appeared more recently. Such judgments were obtained in an experiment reported by Yntema and Trask (1963). Using nouns as stimuli and a self-paced study procedure, they found the probability of a correct recency discrimination to be an increasing function of the difference between the lags of the two test items and a decreasing function of the lag of the more recently presented item. Lag was measured in terms of the number of items intervening between the item's initial presentation and its subsequent appearance in a test pair.

There are many theories that might account for this relationship between lag and the accuracy of discrimination. A subset of such theories would be those based on the assumption that 2AFC discrimination involves an independent lag estimate for each of the two items; S is assumed to make an absolute lag judgment independently for each of the 2AFC items and then to respond "more recent" to that item for which the judged lag is less. The feasibility of this type of theory is subject to direct experimental investigation and it is clearly a sensible research strategy to perform such a test before attempting to formulate a more precise quantitative theory.

The experimental test involves a comparison with performance on a 2AFC task of the order information contained in absolute judgments of recency for single items. In an absolute judgment (AJ) task, a sequence of items is presented and when a test item appears S estimates how far back in the list that item last appeared. Such judgments were obtained in an experiment by Peterson (1967). Let λ_i denote a lag condition of i items, a condition in which i items have intervened between the initial and test presentation of an item. The AJ of a given S for this lag condition may be considered a random variable, X_i , and for a series of lag conditions, $\lambda_i, \lambda_j, \dots$ there will be for each S a corresponding series of random variables, $X_i, X_j, \dots$. These will be assumed independent. If $\lambda_i > \lambda_j$ then the relevant comparison is between prob ($X_i > X_j$) and the probability of a correct 2AFC recency discrimination between items at lag λ_i and λ_j . If a theory of 2AFC recency discrimination based on the assumption of independent lag estimation is a reasonable one, then 2AFC probabilities should be well approximated by the proportion of times S assigns to an item at lag λ_i and AJ which is larger than that assigned to an item at lag λ_j .

A slight complication arises when the same AJ is assigned to items at different lags. For the purpose of estimating the corresponding performance in a 2AFC task, it will be assumed that in such cases the probability of a correct ordering is .5. If

f_{is} and f_{jt} denote the frequency with which items at lags λ_i and λ_j are assigned respectively an AJ of s and t , then the relevant estimate to be compared with 2AFC performance is

$$P(C) = \frac{1}{n_i n_j} \sum_{s=1}^t f_{is} f_{jt} + \frac{1}{2 n_i n_j} \sum_{s=t}^t f_{is} f_{jt} \quad (1)$$

where n_i and n_j are the total number of judgments for the respective lag conditions.

METHOD

AJ Task

Lists were constructed from a pool of common nouns with a Thorndike-Lorge (1944) frequency of 10 or more and from which obviously abstract nouns were excluded. Each item was typed on a separate card; S turned over the cards one by one, putting each card face down before looking at the next. When S came to a test card, distinguished by a large question mark, he recorded on a separate answer sheet his judgment of how many items had been presented since he last saw that word. The lag conditions were: 4, 8, 12, 16, 20, 24, 32, 36, 40, 48, and 64 items. The sequence of items and tests for each list was generated by computer; the program initially randomized the order of the lag conditions, then, subject to this ordering, inserted items and tests at appropriate locations in such a fashion as to minimize the total list length. Several further constraints were imposed on the sequencing of items; the first 10 items in each list were filler items (items not subsequently tested), the word following a test card was always a filler item, and, except for its appearance on a test card, no word was repeated. Since the interval constituting an item's lag typically contained some test cards, S was instructed that in estimating the lag, an intervening test card should be counted as just another item.

One practice and three experimental lists were constructed. The practice list contained a total of 50 cards and included one test for each of the lag conditions. The experimental lists were approximately 100 cards in length and included two tests for each lag condition. For each of the three experimental lists there were 10 different randomizations of the order of lag conditions.

The procedure was self-paced. S worked through the practice deck first and then through one version of each of the experimental decks.

2AFC Judgments

Lists were constructed from the same pool of items and the presentation procedure was the same as for the AJ condition except that the test cards now contained two previously presented items and S wrote down that word of the pair he judged to have been presented more recently. (S also gave a confidence rating, but this data will not be considered in the present paper.)

There were 16 test conditions representing different combinations of the lags used in the absolute judgments condition. These conditions were: (8,4); (12,4); (20,4); (36,4); (12,8); (16,8); (24,8); (40,8); (20,16); (24,16); (32,16); (48,16); (36,32); (40,32); (48,32); (64,32) where the first

Fig. 1. Median judged recency as a function of lag. The straight line represents perfect accuracy.

number in each pair represents the lag of the item further back and the second number the lag of the more recent item. These 16 conditions represent four lag intervals of the more recent item combined orthogonally with four lag differences (4, 8, 16, 32 items). The practice deck contained eight test cards and the three experimental decks contained one test for each lag combination. The sequencing constraints imposed on the AJ lists were also imposed on these lists. There were four parallel versions of each of the experimental decks, formed by reversing the nearer and farther item for each lag combination and for each of these, counterbalancing the order in which the two words appeared on the test card. S worked through the practice deck and then through one version of each of the experimental decks. Again, the procedure was self-paced.

Subjects

Subjects in the AJ condition were 39 student volunteers of both sexes taking introductory psychology courses at the Pennsylvania State University. Ss in the 2AFC condition were an independent group of 74 volunteers from the same pool. All Ss received extra grade points for their participation in the experiment.

RESULTS

The median judged lag for the AJ group is shown in Fig. 1. Each point is based on 234 observations (six judgments from each of 39 Ss). It can be seen from the figure that Ss tended to overestimate short, and underestimate long lags, a trend found also by Brelsford, Freund, and Rundus (1967) and by Peterson (1967). The cross-over point from under- to overestimation appears to be a function of the range of lags employed.

The numerical accuracy of the judgments, however, is not of primary interest. For the purpose of predicting 2AFC performance, the important aspect of the data is the extent to which the judgments of individual Ss preserve the correct ordering of the lag magnitudes. To make the comparison Formula 1 was applied to the AJ data for individual Ss and these probabilities were then averaged over the 39 Ss. The results of this analysis for those lag combinations relevant to a comparison with the 2AFC data are shown in Fig. 2, together with the obtained 2AFC probabilities themselves. The data have been plotted with the lag of the more recent item as the independent variable and the lag difference as the parameter. Unfortunately it is difficult to conduct a rigorous statistical test of the difference between the observed 2AFC performance and the probabilities derived from the AJ data because the latter are not independent. For example, the derived probabilities for lag combinations (8,4) and (12,8) are

both based on the same set of AJs for Lag 8. It would, of course, be possible to use independent subsets of the AJs for the common lag condition but in the present case this seems unnecessary. The hypothesis that 2AFC performance is superior to that derived from the data can be immediately rejected; collapsing over the 16 lag combinations for which there are corresponding 2AFC probabilities gives a mean of .675 for the AJ-derived probabilities while the mean 2AFC probability is .663. Moreover, if each 2AFC is assumed to be an independent Bernoulli trial then all but one of the AJ-derived probabilities fall within the 95% confidence interval for the 2AFC probabilities.

DISCUSSION

The results make it quite clear that the assumption of independent lag estimation is a tenable basis for a theory of 2AFC recency discrimination. Of course they do not establish that the 2AFC task is, in fact, performed in such a fashion, but the data are certainly compatible with, and indeed well predicted by, such an assumption. The major problem in understanding recency discrimination is that of identifying the form in which recency information is encoded. Unlike such dimensions as pitch or loudness, the answer to this question in the case of recency is not at all obvious. To say that items contain "time tags" is simply to restate the fact that S is capable of discriminating recency; nothing has been gained by way of explanation until the notion of a "time tag" is given some further characterization. The results of the present experiment suggest that one possible characterization is S's ability to estimate the amount of material between the initial and test presentation of an item. According to this view, 2AFC recency discrimination would be not unlike comparative judgments of numerosity in which stimuli are presented in temporal sequence.

A somewhat different view, and one that has received some experimental support (Fozard & Yntema, 1966; Morton, 1968), is that time tags are a direct reading of the strength of a continuously decaying trace, so that in the 2AFC task, S would judge as more recent that item judged to have the greater strength. The present results do not contradict such a theory, although if it is assumed that absolute judgments are also based on an item's strength in memory, it does become necessary to make the additional assumption that the transformation that maps trace strength into a lag estimate is executed without further loss of order information.

Fig. 2. Observed 2AFC and derived AJ probabilities as a function of the lag of the more recent item. Δ is the lag difference.

Alternatively, the 2AFC and the AJ tasks might involve quite different processes and the close correspondence between them be quite fortuitous. But at present, this alternative would seem to be an unnecessary retreat from parsimony.

REFERENCES

- BRELSFORD, J., JR., FREUND, R., & RUNDUS, D. Recency judgments in a short-term memory task. *Psychonomic Science*, 1967, 8, 247-248.
- FOZARD, J. L., & YNTEMA, D. B. The effect of repetition on the apparent recency of pictures. Paper presented at the meeting of Eastern Psychological Association, New York City, April, 1966.
- MORTON, J. Repeated items and decay in memory. *Psychonomic Science*, 1968, 10, 219-220.
- PETERSON, L. R. Search and judgment in memory. In B. Kleinmuntz (Ed.), *Concepts and the structure of memory*. New York: John Wiley & Sons, 1967.

THORNDIKE, E. L., & LORGE, I. *The teacher's wordbook of 30,000 words*. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

YNTEMA, D. B., & TRASK, F. P. Recall as a search process. *Journal of Verbal Learning & Verbal Behavior*, 1963, 2, 65-74.

NOTES

1. This research was conducted at the Pennsylvania State University and supported by Grant No. HD-00151 from the National Institute of Child Health and Human Development and Grant No. 86 from the P.S.U. Central Fund for Research. The manuscript was prepared while the author held a postdoctoral fellowship at the University of Toronto, supported by Grant APA 146 from the National Research Council of Canada.

2. Address: Department of Psychology, University of Toronto, Toronto 5, Canada.

(Accepted for publication December 26, 1968.)