Criterion change in continuous recognition memory: A sequential effect

DANIEL B. BERCH
Morehead State University, Morehead, Kentucky 40351

Data from a previous study were reanalyzed to determine the basis for the increasing false-alarm (FA) rate exhibited by subjects as they proceed through a continuous recognition memory task. The results indicated that: (a) the shifting response bias did not reflect an attempt to probability match; (b) the increase in the FA rate over trial blocks was exclusively a sequential effect, i.e., the increase occurred only for trials following the presentation of a new stimulus item; (c) this sequential effect was due to a shifting response bias. It was suggested that the shifting response bias may have resulted from a subjective probability judgment based on a belief similar to that of the "gambler's fallacy."

Shepard and Teghtsoonian (1961) developed a technique for studying memory, referred to as the continuous recognition task. One of the interesting results of their initial experiment using this procedure was that the false alarm (FA) rate increased as the subjects progressed through the task. This finding has since been replicated in a number of other studies (Donaldson & Murdock, 1968; Martin & Melton, 1970; Underwood, 1965; Wells, 1974). Shepard and Teghtsoonian first concluded that this phenomenon resulted from a shifting criterion or response bias. Moreover, they suggested that the subjects tended to bias their responses so as to match the a priori probability (the proportion of times that "old" items are presented, hereafter referred to as q). However, the findings of a subsequent experiment by Shepard and Chang (1963) led these writers to conclude that the increasing FA rate found in the earlier study resulted from an accumulation of diffused elements (proactive interference) rather than a bias toward saying "old."

This issue was partially resolved by Donaldson and Murdock (1968), who employed signal detection analyses to separate the retention-based component of recognition performance from the criterion-based component. Their findings showed quite clearly that the increasing FA rate reflected a shifting criterion rather than a build-up in proactive interference.

In a recent developmental study of continuous recognition memory (Berch & Evans, 1973), we similarly found an increase in the FA rate over trial blocks for both 5½- and 9-year-old children. A signal detection analysis revealed that, as with adults, this change reflected a shifting response bias. Furthermore, we speculated that the children may have changed their criterion in an attempt to probability match. However, as with previous investigators, we did not provide any decisive evidence in support of this explanation. Consequently, a major purpose of the present investigation was to reexamine the Berch and Evans data in order to determine whether the subjects were indeed attempting to probability match.

At this point, it should be noted that Berch and Evans' subjects received neither prior instructions nor informative feedback regarding the value of q.[1] This procedural aspect could prove to be important, however, considering the results of a recent study by Marken and Sandusky (1974). These investigators pointed out that, if a subject received no information about q, no standard would be present to allow acquisition and thus matching of q. Using a continuous recognition task and q values of .2 and .5, they found that the .2 group exhibited a more lenient criterion than the .5 group. In other words, their subjects showed a probability "contrast" type of response bias. Furthermore, they found that this overall contrast effect could at least partially be accounted for by the occurrence of first order, sequential stimulus contrast. That is, the subjects tended to be more likely to say "old" following a previous trial in which the stimulus item was new. Finally, it is apparent from Figure 1 of Marken and Sandusky's report that, as in the Berch and Evans study, the .5 group responded "old" approximately 50% of the time, thus matching their value of q. Nevertheless, these subjects did show evidence of sequential stimulus contrast. Since the children in the Berch and Evans study were similarly uninformed of the value of q, it is certainly possible that their response biases were also in part a function of sequential stimulus contrast. To the extent that this was in fact the case, reanalysis of the Berch and Evans data could reveal whether there was any relationship between first order sequential dependencies and the criterion change that occurred as subjects progressed through the task. The present paper reports the results of a series of such analyses, which were carried out in order to clarify the mechanisms underlying the increasing FA rate exhibited in recognition memory performance.

ASSESSMENT OF PROBABILITY MATCHING

In order to determine whether the subjects were attempting to probability match, the proportion of

"old" responses was computed for each subject on each trial block. Next each score was subtracted from .50 (q) and the absolute value determined. The means (and standard deviations) of these absolute deviations for the three trial blocks were .12 (.07), .08 (.04), .09 (.06), and .06 (.05), .09 (.07), .13 (.06) for the kindergartners and third graders, respectively. If the subjects were indeed shifting criterion in order to probability match, then these absolute deviation scores should decrease over trial blocks. To test this hypothesis, the arcsin transformations of the absolute deviation scores were subjected to a 2 by 3 mixed analysis of variance (grade level by trial blocks). Although the main effect of trial blocks was nonsignificant [$F(1,28) < 1$, $p > .05$], the Grade Level by Trial Blocks interaction was significant [$F(1,28) = 5.94$, $p < .025$]. [A nonadditive model was assumed, necessitating the use of a conservative F test for the within-subjects main effect of trial blocks (Myers, 1966, p. 210). Conservative tests were also necessary for all subsequent analyses of the trial blocks main effect, except for that involving response bias values (RI).] The trends of the means show that, while the third graders' scores were increasing, the kindergartners' scores were decreasing. Thus, the third graders could not have been attempting to probability match. To determine whether there was a significant decreasing effect for the kindergartners, a one-factor repeated-measures analysis of variance was performed on the arcsin transformations of the absolute deviation scores. The main effect of trial blocks was nonsignificant [$F(2,28) = 1.462$, $p > .05$], suggesting that the kindergartners were not attempting to probability match.

Although the kindergartners and third graders showed opposite trends in their absolute deviation scores, the proportion of "old" responses increased over trial blocks for both groups; the means (and standard deviations) were .43 (.12), .51 (.09), .53 (.10), and .47 (.07), .57 (.08), .61 (.10) for the kindergartners and third graders, respectively. Thus, since the third graders averaged .47 "old" responses for Block 1 as compared to .43 for the kindergartners, their increasing proportion of "old" responses was greater than .50, resulting in increasing deviation scores. On the basis of these analyses, then, it appears that the shifting criterion cannot be explained as resulting from an attempt to probability match.

ANALYSIS OF SEQUENTIAL DEPENDENCIES

First order sequential dependencies were analyzed by computing the following formula for each subject:

$$\text{"Old" bias} = P(\text{Hit}/O_{n-1}) + P(\text{FA}/O_{n-1})$$

$$- P(\text{Hit}/N_{n-1}) - P(\text{FA}/N_{n-1}) \quad (1)$$

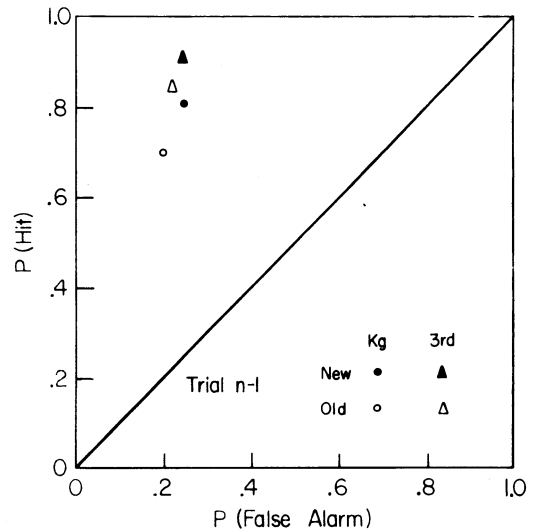where O_{n-1} and N_{n-1} stand for an "old" and a "new"

Figure 1. MOC plot showing conditional probabilities of the response "old" on Trial n as a function of the stimulus value (new or old) on Trial n − 1.

stimulus on the previous trial, respectively. Negative scores resulted for 13 out of 15 kindergartners and 10 out of 15 third graders, indicating sequential stimulus contrast. This effect can be seen in Figure 1, which displays the average first order sequential dependencies. The memory operating characteristic (MOC) plot shows that, for both age groups, the bias for saying "old" is greater when the prior stimulus is new.

SEQUENTIAL ANALYSIS OF FALSE ALARM RATE DATA

Figure 2a shows the basic finding of an increase in the FA rate over trial blocks from the Berch and Evans (1973) study. Since there is evidence of sequential stimulus contrast, it is possible that this increase represents a sequential effect. To determine this, FA rate values were recalculated for each subject by conditionalizing on the type of stimulus item occurring on the preceding trial. As shown in Figure 2b, the FA rate for trials following the presentation of new items increases over trial blocks, just as in Figure 2a. A 2 by 3 (grade level by trial blocks) mixed analysis of variance performed on the arcsin transformations of these data revealed a significant main effect of trial blocks [$F(1,28) = 42.54$, $p < .001$], thus providing statistical support for this observation. Interestingly, Figure 2c seems to show that the finding of an increasing FA rate disappears when this measure is calculated for trials following the presentation of old items only. Indeed, a 2 by 3 mixed analysis of variance performed on the arcsin transformations of these data resulted in a nonsignificant main effect of trial blocks [$F(1,28) = 1.632$, $p > .05$].

Of course, an additional analysis is necessary to determine whether the increasing FA rate reflects a
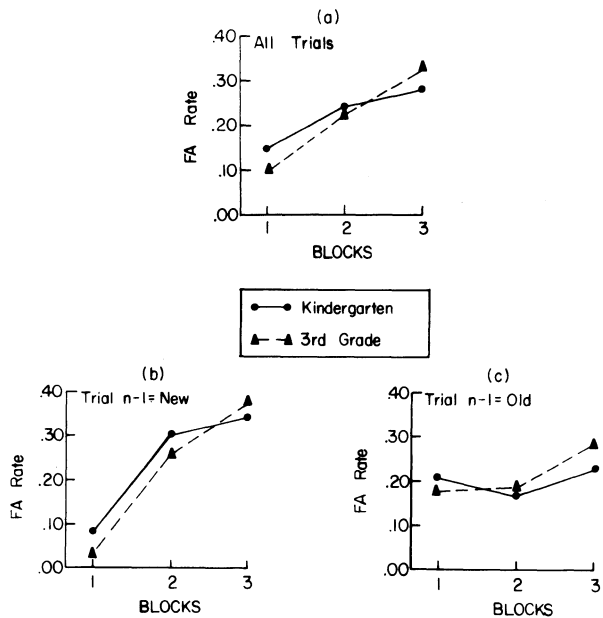
Figure 2. False alarm (FA) rates as a function of blocks of items for: (a) all trials, (b) trials following the presentation of new items only, and (c) trials following the presentation of old items only.

change in response bias. However, a comparison of sequential stimulus contrast scores for each block using Formula 1 would be inappropriate. This is because the sequential score is a relative measure of response bias. Instead, a measure of responsivity, preferably nonparametric, is needed that can be calculated for only those trials where the prior stimulus was new. Frey and Colliver (1973) have devised such a measure (RI), which, although used primarily with discrimination learning data from animals, does have applicability to studies of human memory. Individual conditionalized hit rates and FA rates were substituted into the appropriate equation to obtain RI values for each subject. The means (and standard deviations) of the RI values for the three trial blocks were $-.35$ (.32), .19 (.37), .27 (.34), and $-.21$ (.41), .32 (.38), .38 (.46) for kindergartners and third graders, respectively. A 2 by 3 (grade level by trial blocks) mixed analysis of variance was then performed on the square root transformations ($\sqrt{x + 1.5}$) of these data, resulting in a highly significant main effect of trial blocks [$F(2,56) = 23.42$, $p < .001$]. This finding suggests that, as subjects progressed through the task, they exhibited an increasing bias toward saying "old" when the stimulus on the preceding trial was new.

CONCLUSIONS

The findings of the present investigation indicate that with respect to the Berch and Evans (1973) data: (a) the shifting response bias did not reflect an attempt to probability match; (b) the increase in the FA rate over trial blocks was exclusively a sequential effect; (c) this sequential effect was due to a shifting response bias. At this time, however, it cannot be determined

whether these results have any bearing on the interpretation of the increasing FA rate found in studies that have used adult subjects. For one thing, it is quite possible that the matching hypothesis can account for the shifting criterion exhibited in studies where subjects were explicitly informed of the value of q prior to testing (Donaldson & Murdock, 1968; Wells, 1974). Moreover, it is not clear whether the subjects in some of these studies were explicitly informed about q (Martin & Melton, 1970; Shepard & Teghtsoonian, 1961; Underwood, 1965). In any event, it is hoped that this paper will stimulate reanalysis of the relevant data from these studies so that the generality of the findings described above can be assessed.

There is still the question, however, of why Berch and Evans' subjects shifted to a more lenient criterion as they proceeded through the task. One possible explanation is that an individual's subjective expectation that an old item will appear (Trial n) following the presentation of a new item (Trial n – 1) increases directly with the number of new items previously encountered in the experimental session. This "gambler's fallacy" hypothesis is based on an assumption that, when the subject is uncertain about an item, his task becomes similar to that of predicting the next outcome of a binary series, as in a probability learning situation. In other words, he has to make a "prediction" of "old" or "new" based in part on a subjective, qualitative probability judgment.

Regardless of the validity of this hypothesis, it should be noted that no extant model of the decision component in recognition memory can accommodate the results of either the present investigation or of the Marken and Sandusky (1974) study. There are, of course, a number of psychophysical models that have been proposed specifically to handle the sequential and probability contrast effects found in studies of signal recognition, where the subject is not given any information regarding the value of q (Sandusky, 1971). Presumably, some parameters of these models are applicable to the types of response biases that emerge in recognition memory performance. Nevertheless, as Lockhart and Murdock (1970) have cautioned, "an adequate account of the decision process for a given test of memory performance cannot be made without a commitment to certain aspects of a theory of the memory system for that situation" (p. 100).

REFERENCES

BERCH, D. B., & EVANS, R. C. Decision processes in children's recognition memory. *Journal of Experimental Child Psychology*. 1973, **16**, 148-164.

DONALDSON, W., MURDOCK, B. B., JR. Criterion change in continuous recognition memory. *Journal of Experimental Psychology*, 1968, **76**, 325-330.

FREY, P. W., & COLLIVER, J. A. Sensitivity and responsivity measures for discrimination learning. *Learning and Motivation*, 1973, **4**, 327-342.

LOCKHART, R. S., & MURDOCK, B. B., JR. Memory and the theory of signal detection. *Psychological Bulletin*, 1970, **74**, 100-109.

MARKEN, R. S., & SANDUSKY, A. J. Stimulus probability and sequential effect in recognition memory. *Bulletin of the Psychonomic Society*, 1974, **4**, 49-51.

MARTIN, E., & MELTON, A. W. Meaningfulness and trigram recognition. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 126-135.

MYERS, J. L. *Fundamentals of experimental design*. Boston: Allyn and Bacon, 1966.

PARKS, T. E. Signal-detectability theory of recognition-memory performance. *Psychological Review*, 1966, **73**, 44-58.

SANDUSKY, A. Signal recognition models compared for random and Markov presentation sequences. *Perception & Psychophysics*, 1971, **10**, 339-347.

SHEPARD, R. N., & CHANG, J. J. Forced-choice tests of recognition memory under steady-state conditions. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 93-101.

SHEPARD, R. N., & TEGHTSOONIAN, M. Retention of information under conditions approaching a steady state. *Journal of Experimental Psychology*, 1961, **62**, 302-309.

THOMAS, E. A. C., & LEGGE, D. Probability matching as a basis for detection and recognition decisions. *Psychological Review*, 1970, **77**, 65-72.

UNDERWOOD, B. J. False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, 1965, **70**, 122-129.

WELLS, J. E. Strength theory and judgments of recency and frequency. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 378-392.

NOTE

1. The boundary conditions of the formal model of the matching hypothesis require that the subject must possess a value of q, either through instructions or feedback (Parks, 1966; Thomas & Legge, 1970). Thus, even if Berch and Evans' subjects were not attempting to probability match, such a finding would not detract from the validity of this model.