JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR, 13, 209–216 (1974)

Negative Recency In Delayed Recognition[1]

RANDALL W. ENGLE

King College

There is some controversy as to whether negative recency is found on a delayed final recognition test following immediate free recall. The present experiment demonstrated how presentation rate of input and the method of analysis of the data interact to yield the differential results.

One theoretical explanation of the free recall serial position function is that the terminal items, being output first, are recalled from a transitory short-term store. The memorial traces of the items in this store decay over time unless rehearsal continually re-enters them into the store or simultaneously builds up a trace or traces in the long-term store (Atkinson & Shiffrin, 1968).

Tulving (1968) has proposed an alternative explanation to the two-store concept. He offers the notion that all input information is stored in the same unitary storage system and that differential recall of primacy, middle, and terminal items reflect differences in the retrievability of these items and not storage strength. The end items are thus argued to be retrieved more easily because they have stored with them at input certain kinds of supplementary information not available to the other items. This extra information serves as a retrieval cue to the end items and this causes higher recall performance for these end items.

Tulving (1968) argues that one possible kind of auxiliary information that is available to these end items is an acoustic trace for items that have been presented auditorily or for visual items that the subject has recoded into an acoustic representation. Another possible type of the hypothesized extra retrieval information is the temporal dating or time tagging of items. This temporal information, like the other ancillary retrieval information, is proposed to decay rapidly over time. This leaves the early and middle items with, at the time of recall, only the more permanent retrieval cues. With delayed recall, the terminal items are proposed to lose this added edge in retrieval.

Craik (1970) has argued that the one-store model (Tulving, 1968) and two-store model (Atkinson & Shiffrin, 1968) make different predictions regarding the final storage of those items making up the beginning and middle of the list and those items making up the terminal section of the list. Tulving (1967) found that recall of an individual item in free recall improves its recall performance on later trials. Since probability of recall in immediate free recall is highest for the terminal items, they should, according to Tulving (1967), show the greatest improvement in performance on subsequent trials. This should be true if recall of the terminal items is from the same store as recall of the earlier items, which is what Tulving proposes.

The two-store model, at least the one proposed by Atkinson and Shiffrin (1968), predicts differently for subsequent attempted recall of the terminal items. This model predicts that items output from the short-term store, such as the terminal items, are

[1] This research was conducted while the author was a National Science Foundation Trainee at the Ohio State University. I wish to thank Dr. D. D. Wickens for his advice and criticism.

dumped from the store and thus would have little or no permanent trace built up in the long-term store.

Craik (1970) tested the predictions of the two models against each other. He presented 10 lists of 15-words each and asked for immediate recall after each list. After the recall of the last list was completed the subject was given a sheet of paper and instructed to recall as many words from *all* the lists as he could remember. The final recall performance of the individual items was plotted as a function of their input positions in the original list. The serial position curve exhibited a negative recency effect. That is, the terminal items, which showed very high performance on immediate recall, were recalled less well than either primacy or middle items at final free recall. This supported the two-store model and seriously called into question the credibility of Tulving's (1968) one-store theory

The one-store model might say that the terminal items are not retrieved as well at final free recall because permanent retrieval cues have not been developed. In other words, terminal items are stored at least as well as primacy and middle items but they just cannot be retrieved as well.

The test of this suggestion would be to look at final recognition instead of recall. If a positive or no recency effect is found this could lend support to Tulving's (1968) theory. Craik, Gardiner, and Watkins (1970) performed two experiments concerned with performance on a final recognition test following a series of immediate free recall lists. Both experiments resulted in a negative recency effect on the final recognition serial position curves.

Cohen (1970), however, has obtained contradictory results using final recognition. In one study Cohen presented for immediate free recall ten 10-word lists at a 2-second rate. After the ten lists were completed the subjects had final free recall and then final recognition. The final recognition data exhibited no negative recency effect and even a slight positive recency. In a second experiment Cohen presented 27 9-word lists at a 1-second rate with final recall and final recognition after successive groups of nine lists. This time there resulted a clear positive recency effect with the final recognition data. Cohen concluded that the negative recency with final recall did indeed represent a failure to retrieve and not a failure to store the terminal items. In other words, he opted for the Tulving (1968) theory.

The positive or negative direction of the recency function seems to be an important and valid problem worthy of an attempted solution. Cohen (1970) analyzed his data in terms of recognition *unconditional* on immediate free recall while Craik *et al.* (1970) analyzed recognition of only those words recalled on immediate free recall, that is, recognition *conditional* upon recall. Cohen, in his first study, presented his materials at a 2-second rate and found a nearly flat function, that is, little or no recency. But in the second experiment he used a 1-second rate and found a notably positive recency effect for final recognition.

This paper is an investigation of whether these two variables, mode of analysis and presentation rate, are responsible for these contradictory findings.

Method

An experiment was conducted to investigate the potential effects of the two variables described above. The method consisted of presenting subjects with lists of words at one of several rates of presentation for immediate free recall. After the last list was recalled subjects received either a recall or recognition test of all the words in the lists.

Design

The independent variables were presentation rate, serial input position, and type of final retention test. Each subject was presented

the to-be-remembered words at either a 1, 2, or 4-second rate and received either a final free recall (FFR) or a final recognition (FRN) test for all list words. Rate of presentation and type of delayed test were both between subject variables and were varied factorally. List words were partially counterbalanced across input positions.

Subjects and Materials

The subjects were 192 introductory students at The Ohio State University. They were tested in groups of 1–4 individuals and were alternately assigned to the groups according to their arrival.

Each subject received ten lists of 16 words each. The words were all high frequency nouns from Kučera and Francis (1967). The final recognition test was composed of the 160 list words and 160 lures of the same characteristics as the list words. The list and lure words were randomly assigned positions in the recognition test.

Procedure

The subjects were told that the experiment would consist of two parts but not that the second part would be a test of all the words. The immediate recall followed usual procedure. The words were presented visually at one of the three rates and 2 minutes were allowed for the written recall. After the recall was complete for the last list, subjects waited 2 minutes while the experimenter ostensibly worked with the apparatus. This was simply a delay before the final retention task.

The FFR group was given a sheet of paper and instructed to write down as many of the 160 words as possible. Five minutes were allowed for FFR. The FRN group was given test sheets and told to rate their confidence that each word was old or new. These subjects were instructed to give serious consideration to each item on the test and to rate the item 1, very sure old; 2, fairly sure old; 3, think it is old; 4, think it is new; 2, fairly sure new; or 6, very sure new. The recognition task took 15–20 minutes on the average.

Results

Analysis of the immediate free recall data did not yield anything that was unexpected, as can be seen in Figure 1. The 4-second group performed best with a systematic decline in performance for the 2-second and 1-second groups, $F(2, 189) = 73.85$, $p < .001$. The serial position data were typical for all three groups.

The FFR and FRN data were analyzed by two methods. *Unconditional* analysis is final performance on all words presented in the

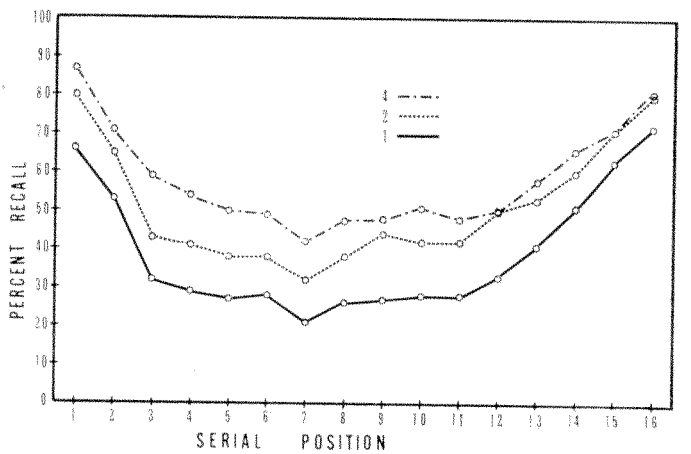

FIG. 1. Serial position curves for immediate free recall.

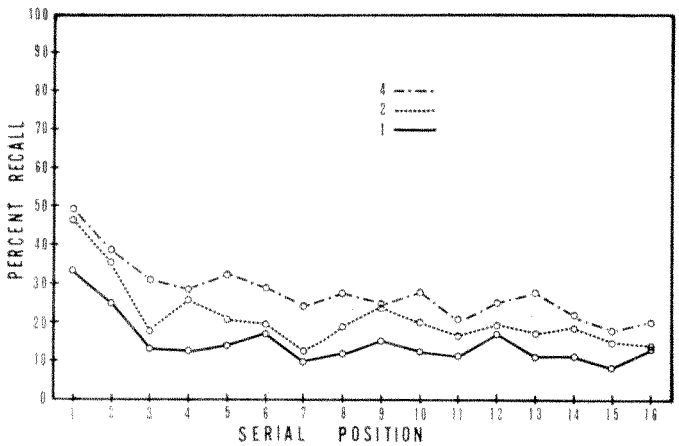

FIG. 2. Serial position curves for final free recall of all words presented at input.

lists. *Conditional* analysis is the performance on the final test for only those items recalled on immediate free recall. As was mentioned above, this is one of the major differences between the Craik *et al.* (1970) and the Cohen (1970) experiments.

The FFR performance for all items input, unconditional on immediate free recall, is shown in Figure 2. Analysis of variance of these data resulted in significant differences for rates, $F(2, 84) = 17.36$, $p < .001$, serial position, $F(15, 1290) = 35.55$, $p < .001$, and a small but significant interaction effect, ($F(30, 1260) = 1.58$, $p < .05$. Looking at Figure 2 there seems to be little if any negative recency but very prominent primacy effects. A Spearman's rho was performed correlating performance with positions 1–7 and positions 9–16. The primacy effect was significant for the 2- and 4-second groups with rhos of $-.79$ and $-.82$, respectively, $p < .05$, $n = 7$. While the trend was very clear for the 1-second group, it was nonsignificant, $\rho = -.67$. For the terminal positions, only the 2-second group exhibited a significant negative recency effect with $\rho = -.96$, $n = 8$, $p < .01$.

The conditional analysis on the FFR data is shown in Figure 3. Analysis of variance

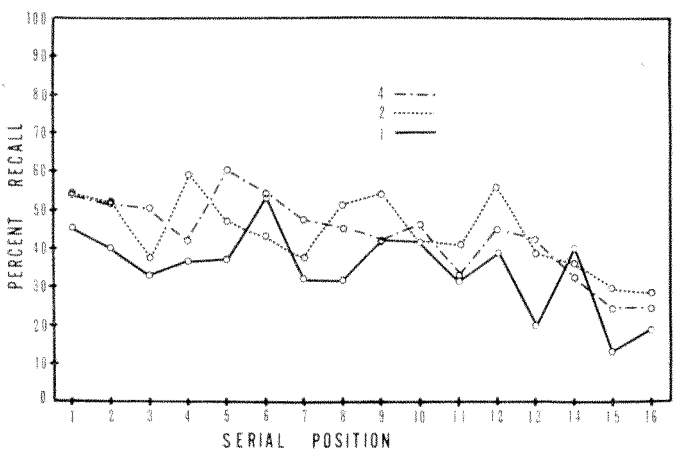

FIG. 3. Serial position curves for final free recall conditional upon recall in immediate free recall.

of these data resulted only in significant effects of rate, $F(2, 93) = 4.05$, $p < .05$, and position, $F(15, 1395) = 8.32$, $p < .001$. This method of analyzing the data resulted in essentially no primacy effects but the appearance of unmistakable negative recency effects for all three groups. The absence of primacy effects was confirmed by non-significant rhos of −.32, −.46, −.02 for positions 1–7 for the 1, 2-, and 4-second groups, respectively. The Spearman analysis on positions 9–16 resulted in rhos equal to −.81, −.86, and −.81, $p < .05$, for the 1-second, 2-second, and 4-second groups, respectively. The negative recency in FFR is obviously much more pronounced with a conditional analysis than with an unconditional analysis.

The recognition data also were analyzed by both methods since this was one of the major differences between Craik *et al.* (1970) and Cohen (1970). Figure 4 shows mean confidence rating (CR) unconditional upon recall in IFR and the mean CR for the lure words for the three presentation rates. As with the unconditional recall data there are prominent primacy effects for all three groups. This was confirmed by rhos of 1.00, .86, and .80, $p < .05$, with $n = 7$, for positions 1–7 on the 1-, 2-, and 4-second groups, respectively. Spearman analysis of the 1-second group for positions 9–16 resulted in $\rho = -.81$, $p < .05$. In other words, a significant positive recency effect as in the Cohen (1970) study. The same analysis on the 2-second and 4-second groups resulted in $\rho = +.26$ and $\rho = -.20$, both non-significant. This means that the two slower rates resulted in no observable recency, positive or negative.

The conditional analysis of the FRN data is depicted in Figure 5. The top family of curves is the mean CR for all those words which were recalled in immediate free recall. The bottom family of curves is for those words *not* recalled in immediate free recall. Because the number of data points is different for each serial position and each group, the curves have been smoothed by averaging points $n - 1$, n, and $n + 1$. The first and last serial positions are unsmoothed data and, of course, the statistical analysis was done on the raw unsmoothed data. The Spearman analysis on positions 1–7 demonstrated that, again, conditional analysis had eliminated the primacy effects for all three groups. All three resulted in non-significant rhos for the early positions. The FRN curves for the words recalled in immediate free recall exhibited a very obvious negative recency for all three rates. This was shown by $\rho = 1.00$, 1.00 and 1.00, $p < .05$, for the last five positions of the

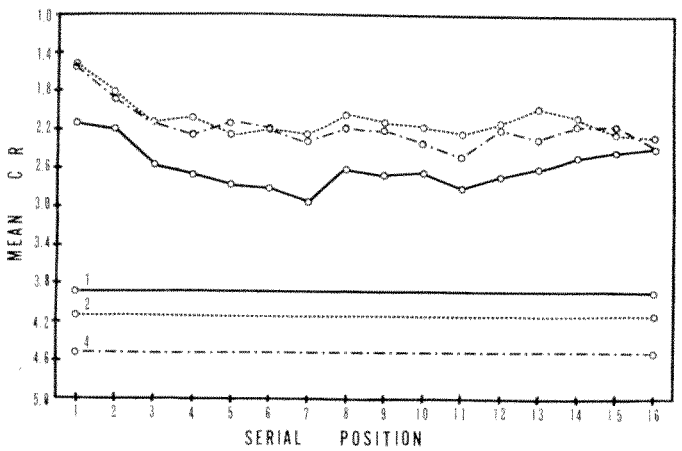

FIG. 4. Serial position curves for mean confidence rating for all words presented at input (top lines) and for all lure words in the FRN test (bottom lines).

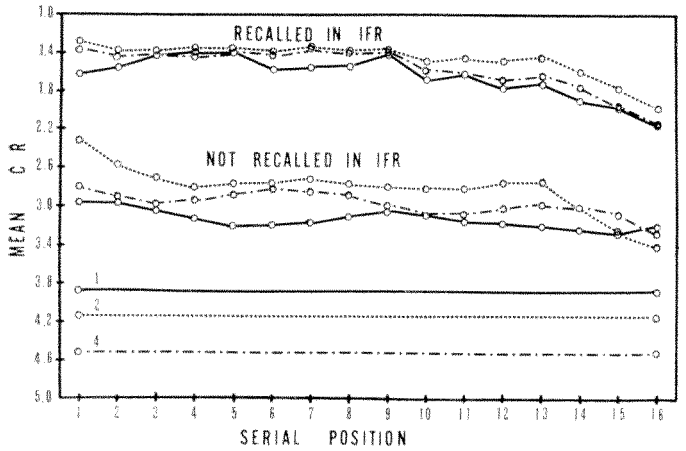

FIG. 5. Serial position curves for mean confidence ratings conditional on recall (top lines) or non-recall (middle lines) in immediate free recall and for all lures in the FRN test (bottom lines).

1-, 2-, and 4-second groups. The reason for the inversion on points 12 and 13 in Figure 5 is that the figure is constructed from smoothed data. The raw data did not show an inversion for points 12 through 16 thus yielding rhos of 1.00.

The FRN data for those words *not* recalled in IFR showed no significant primacy effects and neither a significant positive nor negative recency effect ($\rho = +.71$, $+.62$, and $+.24$ for 1-, 2-, and 4-second groups). As mentioned earlier these curves are drawn from different numbers of data points for each group and have been smoothed so the statistical analysis may not correspond to what appears intuitive from the figure.

One final analysis that was made was final recognition performance as a function of the output position of the item in immediate free recall. This analysis is depicted in Figure 6. The 10th through 16th output items, being few in number, have been lumped together as end (E) items in the figure. As can be seen there is a nearly linear relationship between output position and mean CR, particularly for the first five output positions.

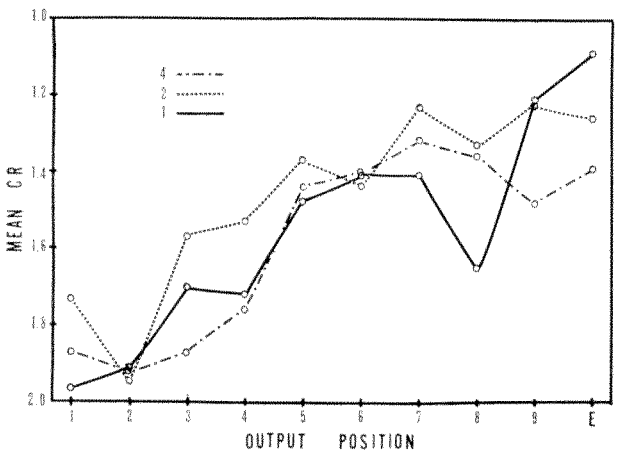

FIG. 6. Mean confidence rating as a function of the items output position on immediate free recall.

A Spearman analysis on all 10 output positions resulted in rhos of −.90, −.92, and −.77, $p < .02$, $n = 10$, for the 1-, 2-, and 4-second groups respectively.

Discussion

It is quite obvious from the results that the unconditional and conditional methods of analysis give different pictures of the recency effect in delayed recall and recognition. It is also clear that this difference is greater with a presentation rate as fast as one second than with slower rates. This is particularly relevant since Cohen (1970) used unconditional analysis and found a positive recency with FRN only after going from a 2-second to a 1-second rate of presentation. Cohen concluded that this proved the negative recency observed with FFR reflects a retrieval and not a storage difficulty. But it is clear from Figures 4 and 5 that the observation of the positive or negative recency with FRN depends on the rate of presentation and method of analysis, thus negating the basis for Cohen's (1970) conclusion. Neither those words recalled in immediate free recall nor those not recalled in immediate free recall show a positive recency.

The unconditional 1-second FRN data show a strong positive recency effect. When these data are broken down by conditional analysis they result in two curves, neither of which has a positive recency and one of which has a strong negative recency effect. The question may arise as to how this can be possible. In case the answer is not obvious, a demonstration is in order. Figure 7 presents some hypothetical FRN data with the solid lines representing the conditional analysis. The top line represents the mean confidence rating (CR) of all words recalled in immediate free recall and the bottom line represents the mean CR of those words not recalled in immediate free recall. The numbers in the lines represent the number of items on which the mean CR is based. This hypothetical case is similar to the 1-second group in the present

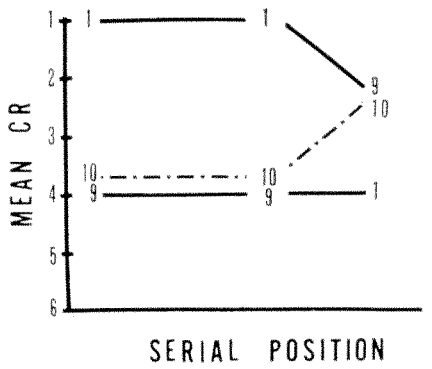

Fig. 7. A hypothetical set of curves demonstrating how two curves, neither of which has a positive recency, result in a curve with positive recency when combined if the number of observations per data point is extremely unequal for the two curves. The two solid lines represent conditional FRN analysis of words recalled (top line) and words not recalled (bottom line) in immediate free recall. The broken line represents the unconditional analysis.

experiment (and, indeed, other experiments as well) in that many words are recalled from the terminal positions (9 in this hypothetical case) but, due to the fast rate of presentation, few words are recalled from the earlier positions (1 in this case). The number of items contributing to a given data point for the items not recalled in immediate free recall (bottom line) is the complement of the number of items contributing to each point in the top line. This means that the number of items that contribute to the terminal positions is small but the number of items contributing to the early positions is large.

When the two types of data are combined, as is done with the unconditional analysis, the result is a curve that resembles the broken line. So a curve can be obtained with the unconditional analysis that shows a positive recency even though neither of its components has a positive recency. It therefore seems mandatory that delayed retention data be analyzed conditional upon recall on the earlier test.

This experiment also seems to deny any support for Tulving's (1968) theory, at least

from delayed retention data. The terminal items at the time of immediate free recall are assumed to be of equal strength to the earlier list items. So the theory, at least in the form espoused by Cohen (1970), should predict that the terminal items would do at least as well on delayed recognition as the items in the earlier part of the list. This was clearly not the case, as was observed on the conditional analyses.

A finding that seems pertinent to the Atkinson and Shiffrin (1968) model is the analysis of mean CR as a function of the items output position on immediate free recall. Figure 6 shows that final recognition of an item improves the later the item was retrieved. Actually most of the effect seems to be subsumed by the first five output positions, typically thought of as being items in short-term memory. Positions 5 through the end items show very little difference in performance, particularly for the 2- and 4-second groups. This would, at first blush, seem to lend credence to the idea that the longer an item remains in short-term memory the stronger that item becomes in long-term memory. However, if this were the case one should notice considerable differences over the first five output positions for the three rates of presentation. Presumably the second or third output item has been in short-term memory longer if input is at a 4-second rate than if input is at a 2- or 1-second rate. This should cause the 4-second group to perform better over the early output positions than the 2-second group. This is decidedly not the case and in fact the 4-second group is slightly lower.

The absence of primacy effects with the conditional analysis also presents at least one interesting possibility. It could be that items recalled from primacy positions in immediate free recall really have no greater strength at the time of delayed recognition than items recalled from the middle positions. This would seem to run contrary to the fact that primacy items receive more rehearsals than later items (Rundus & Atkinson, 1970). This leads to the possibility that the primacy effect was not observed with conditional FRN data because of a ceiling effect.

References

Atkinson, R. C., & Shiffrin, R. M. Human memory. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation*. New York Academic Press, 1968.

Cohen, R. L. Recency effects in long-term recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 672–678.

Craik, F. I. M. Fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 143–148.

Craik, F. I. M., Gardiner, J. M., & Watkins, M. J. Further evidence for a negative recency effect in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970. **9**, 554–560.

Kučera, H., & Francis, W. N. *Computational analysis of present day American English*. Provdence, R. I.: Brown University Press, 1967.

Rundus, D., & Atkinson, R. C. Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 99–105.

Tulving, E. The effects of presentation and recall of material in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 175–184.

Tulving, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton. (Eds.) *Verbal Behavior and General Behavior Theory*. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

(Received September 20, 1973)