Memory & Cognition
1974, Vol. 2, No. 1A, 95-100

Memory in reading and listening to discourse*

JACQUELINE S. SACHS†
University of Connecticut, Storrs, Connecticut 06268

After reading or listening to short passages, Ss attempted to recognize semantically changed sentences and paraphrases (syntactically and lexically changed sentences). The intervals between the original presentation and test ranged from 1 to 23 sec. In general, paraphrases were poorly detected after a brief time, supporting earlier findings that the exact wording of sentences is not stored in long-term memory. An exception was the high recognition of active-passive changes with the visual presentation. Recognition at the first test interval was significantly better after listening than after reading, although the eventual level of recognition memory was not different in the two modes. This result, consistent with other studies of modality effects in short-term memory, suggests that acoustic-phonetic memory played a role in the storage of the auditorally presented material.

An early technique for investigating memory for language was to have someone read or listen to some sentences, and then ask him to try to recall them. In 1908, Karl Bühler reported that the most striking observation from this method was that the exact words were not recalled, although the general meaning was. Bühler concluded that the process of recall involved a reconstruction from the remembered meaning, rather than a simple expression of something stored in memory.

In 1967, Sachs reported a study in which Ss listened to passages and then attempted to recognize changes in sentences they had heard. With this method, Ss try to remember as much about the original form of the sentences as they can, but Ss are also forced to perform their usual sort of semantic processing to succeed at recognizing semantic changes in the sentences. By comparing recognition for semantically changed sentences and for various sorts of paraphrases of the original sentences, conclusions can be reached about the duration of the exact wording in memory and about the characteristics of the encoding of the linguistic material at different points in time. Sachs found that paraphrases of sentences were poorly identified as different from the original sentences after 80 syllables of interpolated linguistic material had been heard. Changes in the meaning of the sentence, on the other hand, were well recognized at that and longer intervals. These findings were taken as support of the idea that linguistic material becomes encoded in terms of meaning.

In that study, however, the shortest time interval between the original presentation and recognition test (other than immediate repetition) was about 27 sec, and, therefore, no conclusions could be reached about the characteristics of the encoding between these two points. Other subsequent studies using the recognition method have not investigated the effect of amount of interpolated material systematically. In the present study, a method similar to that reported by Sachs was used to extend the investigation of memory for discourse to time intervals shorter than 27 sec. Shorter time intervals are of special interest because it seems as if some representation fairly close to the original stimulus may be stored for a short time after semantic encoding. For example, Lashley (1951) noted that one experiences a feeling of surprise when one hears the sentence "Rapid righting with his uninjured hand saved from loss the contents of the capsized canoe." The surprise may be taken as evidence that the semantic encoding of "righting" as "writing" had already occurred and that the later material in the sentence forced the listener to reject that interpretation. The fact that reencoding was possible is evidence that some representation of the sound of the word, and not just its semantic interpretation, was still in memory when "capsized canoe" was heard.

The present study also broadens the generality of the earlier study by including both auditory and visual presentations of the linguistic material. The question is whether the pattern of retention of the wording and the meaning is the same during reading as during listening.

METHOD

Auditory Presentation

Subjects

The Ss were 120 Fairleigh Dickenson University students (60 female, 60 male) who were paid for participation in the experiment.

Materials

Twenty-four passages of connected discourse were written.

*Data were collected and partially analyzed at Bell Telephone Laboratories, Murray Hill, New Jersey. Further analyses were supported in part by Grant No. OEG-0-9-160440-4144(010) for basic research in education from the U.S. Department of Health, Education, and Welfare. The author is indebted to M. L. Goodman and C. L. Mallows for assistance with statistical analyses and to R. Speelman, P. S. Vivekananthan, and M. Friedman for assistance in data collection. Thanks are also due M. Turvey and A. M. Liberman for helpful comments on an earlier version of this paper.

†Requests for reprints should be sent to Jacqueline Sachs, U-85, University of Connecticut, Storrs, Connecticut 06268.

SAMPLE SETS OF RELATED SENTENCES:

BASE:	THE FOUNDING FATHERS CONSIDERED OWNING SLAVES TO BE IMMORAL
SEMANTIC:	THE FOUNDING FATHERS DIDN'T CONSIDER OWNING SLAVES TO BE IMMORAL.
PASSIVE/ACTIVE:	OWNING SLAVES WAS CONSIDERED TO BE IMMORAL BY THE FOUNDING FATHERS.
FORMAL:	THE FOUNDING FATHERS CONSIDERED OWNING SLAVES IMMORAL.
LEXICAL:	THE FOUNDING FATHERS THOUGHT OWNING SLAVES TO BE IMMORAL.
BASE:	A WEALTHY MANUFACTURER, MATTHEW BOULTON, SOUGHT OUT THE YOUNG INVENTOR.
SEMANTIC:	THE YOUNG INVENTOR SOUGHT OUT A WEALTHY MANUFACTURER, MATTHEW BOULTON.
PASSIVE/ACTIVE:	THE YOUNG INVENTOR WAS SOUGHT OUT BY A WEALTHY MANUFACTURER, MATTHEW BOULTON.
FORMAL:	A WEALTHY MANUFACTURER, MATTHEW BOULTON, SOUGHT THE YOUNG INVENTOR OUT.
LEXICAL:	A RICH MANUFACTURER, MATTHEW BOULTON, SOUGHT OUT THE YOUNG INVENTOR.

Fig. 1. Sample sets of related sentences.

Four of these were used as warm-up passages. The other 20 were constructed as follows: For each passage there was a set of related sentences, any one of which could appear in the passage. A "test sentence" tested the S's recognition memory of the "original sentence" presented. The passages were similar to those used in Sachs (1967), except that a new test condition, the substitution of a synonym, was added and the intervals between the original sentence and the test sentence were different.

The passage length preceding the original sentence ranged from 27 to 180 syllables, with a mean of 98 syllables. Three aspects of the construction of the passage and test sentence were varied systematically: (1) the amount of interpolated material between the original and test sentence, (2) the relationship between the original and test sentence, and (3) the order of presentation of the two related sentences.

(1) Amount of Intervening Material. Four amounts of intervening material (IM) occurred between the original and test sentences: 0, 20, 40, and 80 syllables. Based on the average duration between the end of the original sentence and the beginning of the test sentence on the stimulus tapes, these correspond in time to 3, 7.5, 12.5, and 23 sec, respectively. The IM was a continuation of the passage, with no break of any kind at the original sentence, so S never knew what sentence would be heard in the recognition test. At the end of the passage, a bell rang softly to signal the onset of the test sentence.

(2) Relationship Between the Original and Test Sentences. Figure 1 shows sample sets of related interchangeable sentences. The first sentence in each set will be referred to as the "base" sentence. The others were derived by changing the base sentence in some respect. Therefore, each base sentence has a corresponding set of sentences:

Semantic–The meaning of the base sentence was altered by interchanging subject and object, by negation, or by substitution of a word that occurred elsewhere in the passage. The passage was constructed to be neutral with respect to the semantic content in the original sentence.

Active/Passive–The base sentence was changed from active to passive voice. In many cases the change from active to passive was accompanied by dropping the agent in order to preserve the ordinary usage of the passive.

Formal–The form of the base sentence was changed without changing its meaning.

Lexical–A word that was synonymous in the context was substituted, such as "ship" for "boat" or "great" for "famous."

Identical–Another reading of the sentence, rather than a replay of the base sentence, so that the acoustic cues were not identical, though words and word order were.

(3) Order of Presentation of the Two Related Sentences. It is possible that one form from any of the pairs of related sentences might be a more common form or that it might be preferred for some reason. Therefore, the base sentence was used in the passage as the original half the time and a changed version was used as the original half the time.

All stimulus materials were tape recorded on a master tape by a male professional announcer at Cue Recording Studio, New York City. To achieve balancing of the design, 40 separate tapes of 20 passages each were made from the master tape.

Design

The experimental design was a 1/40 replicate of a complete factorial design, with each S tested in half the 40 possible conditions. There were five relationships between original and test sentence, two orders of presentation of the related sentences, and four amount of interpolated material. These combined to form 40 possible conditions. Forty tape recordings of the stimulus materials were prepared, with half of all conditions represented in 20 passages per tape. The order of passages on a tape was randomzied by experimental block, so that several instances of the same type of condition never occurred together. Each tape was heard by three Ss at separate times; therefore, there are three responses to each passage in a particular condition. Ss were assigned to the tapes in a randomized order.

Procedure

Each S was tested individually in one 45-min session. The S heard four warmup and 20 experimental passages on tape. Each passage was interrupted by a bell, and a test sentence was then heard. If S thought the test sentence was identical to one in the preceding passage, he marked "identical" in an answer booklet. If he perceived any change at all, he marked "changed," and then attempted to classify the type of change as being in the meaning or in the form of the sentence. He also rated his confidence in his judgment of "identical" or "changed" on a five-point scale. Ss were instructed to listen for changes in both meaning and form.

Visual Presentation

The experimental method was the same as for the auditory presentation, with the following exceptions: The Ss were 120 Fairleigh Dickenson students, none of whom had served in the auditory condition. The passages and test sentences were the same as those described above, except that Ss read the passages and the test sentences silently rather than listening to a tape recording. In the identical sentence condition, the words appearing at the ends of lines were shifted so that, again, the physical stimulus was not identical, though the words and word

Table 1
Percentage of Sentences of Each Test Type That Were Judged to be "Changed"

		Amount of Intervening Material			
Condition	Test Type	0	20	40	80
Auditory	Semantic	92	78	87	80
	Active/Passive	90	73	61	58
	Formal	80	55	59	57
	Lexical	69	53	45	50
	Identical	14	37	46	42
Visual	Semantic	84	88	82	87
	Active/Passive	88	80	79	78
	Formal	79	62	58	61
	Lexical	57	56	61	59
	Identical	25	44	43	42

order were. The materials were bound in a loose-leaf notebook, one passage per page, and as soon as the S came to the end of the passage he turned the page and read the test sentence. Responses, as in the auditory condition, were "identical" or "changed," a confidence rating from 1 to 5, and a classification of the kind of change as in the meaning or in the form of the sentence. Instructions were given not to reread anything in the passage, and E sat across from S during the whole session, watching that he did not reread or turn back pages. Thus, there was pressure to read at a normal speed rather than memorize. The design was identical to that used for the auditory presentation.

RESULTS

Table 1 shows the results for each type of change at the four levels of intervening material. Given in this table is the percentage of sentences of particular test type that was judged to be "changed," not the percentage correct. For the top four rows, saying "changed" is the correct response, but, for the bottom one, "changed" is incorrect. The results are given in this manner in order to consider the identical sentence as a control against which to test the hypothesis that Ss were not performing at a chance level. It would be incorrect to assume that chance performance would result in 50% correct, because there may be a bias to say "changed" or "same" in the absence of information. However, if responses to identical sentences and changed sentences differ, we have reason to think that the changes can be recognized.

Figure 2 is based on the data in Table 1 and shows the results for each type of change, plotted as the percentage of correct responses minus the percentage of "changed" responses to identical sentences at each level of IM. This figure may be interpreted as the "amount above chance" performance.

The "semantic change" condition actually involved both a change in meaning and a change in the form of the sentence. The data on Ss' ability to identify the type of change was used to determine whether Ss made their judgments on the basis of meaning or form. For semantic changes at IM 0, the percentage of correct responses that also included the correct identification of type of change was 79% in the auditory condition and 80% in the visual condition. These high percentages indicate that the semantic changes were being identified on the basis of the change in meaning rather than form. Formal changes were also identified very accurately as to type of change at IM 0. The percentage of correct responses that also included the correct identification of change was 95% in the auditory condition and 82% in the visual condition.

Ss had rated their confidence in each judgment. A score for each response was based on the correctness and the confidence rating. Thus, a score could be from +5 to +1 when correct and from −1 to −5 when incorrect. Table 2 shows the mean score and standard error for each type of change at each amount of interpolated material. For the auditory condition, analysis of variance

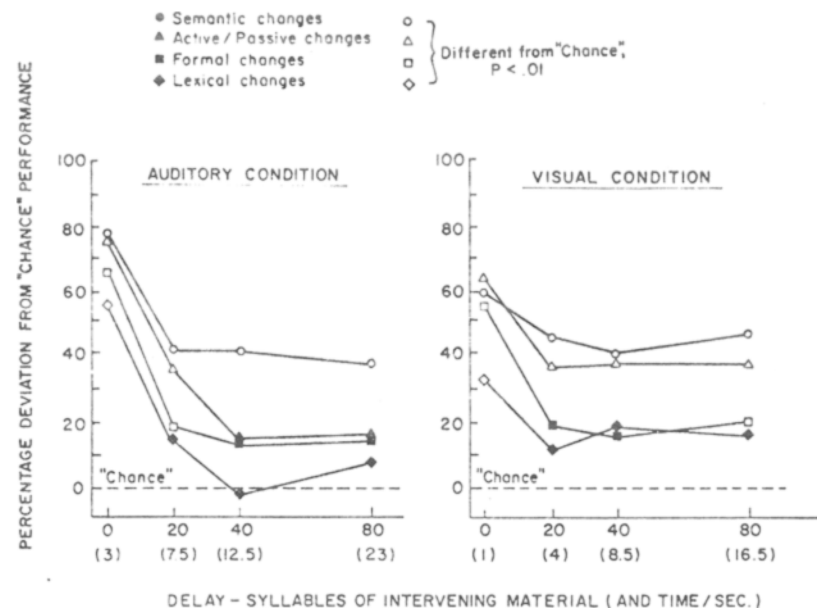

Fig. 2. Accuracy in judging changes, expressed as percentage of deviation from "chance" performance at the four levels of intervening material for auditory and visual conditions.

Table 2
Mean Score and Standard Error for Each Type of Change at Each Amount of Interpolated Material

Presentation Mode	Type of Change	0 Syllables Mean	0 Syllables SE	20 Syllables Mean	20 Syllables SE	40 Syllables Mean	40 Syllables SE	80 Syllables Mean	80 Syllables SE
Auditory	Semantic	4.16	.21	2.55	.31	3.14	.26	2.68	.29
	Active/Passive	3.83	.25	1.65	.32	1.00	.33	0.58	.32
	Formal	2.98	.33	0.43	.36	0.59	.34	0.36	.35
	Lexical	1.97	.37	0.22	.36	−0.33	.34	0.21	.33
	Identical	3.48	.26	1.38	.34	0.43	.35	0.85	.33
Visual	Semantic	3.23	.31	3.41	.27	2.85	.29	3.40	.25
	Active/Passive	3.55	.25	2.75	.30	2.53	.27	2.18	.29
	Formal	2.65	.33	1.08	.35	0.87	.34	0.87	.34
	Lexical	0.72	.40	0.42	.36	0.88	.35	0.87	.33
	Identical	2.37	.34	0.82	.36	0.52	.35	0.79	.33

of these scores revealed significant effects for both types of change and amount of interpolated material [$F(4,76) = 13.65$ and $F(3,57) = 40.50$, in both cases $p < .01$]. From Fig. 2 (auditory) we can see that semantic changes were easier to recognize than the others. Memory for nonsemantic changes in the sentences was measured by McNemar tests for the significance of changes, in which recall of identical sentences was compared with recall of changed sentences. At 20 syllables of intervening material, the identification of active/passive and formal changes was above chance, with χ^2 = 29.82 and 8.17, respectively ($p < .01$). Identification of lexical changes was poor but also significantly better than chance, with $\chi^2 = 6.02$, $p < .05$. At 40 and 80 syllables of IM, none of the changes that did not affect meaning were recognized significantly better than chance, except for the active/passive change at 80 syllables of IM ($\chi^2 = 6.56$, $p < .05$). Conditions in which recognition memory exceeded chance levels ($p < .01$) are indicated by open symbols in Fig. 2.

For the visually presented material, analysis of variance of the scores based on correctness and confidence revealed significant effects both for type of change and for amount of intervening material [$F(4,76) = 19.39$ and $F(3,57) = 6.57$, $p < .01$]. (See Table 2 for means and standard errors.) At all levels of IM, the semantic and active/passive changes were recognized well above chance levels, based on McNemar tests. At IM 0, all types of sentence changes were very well identified. At IM 20 formal changes were recognized above chance ($\chi^2 = 5.79$, $p < .05$). At longer delays, performances continued to be somewhat above chance. At IM 40 recognition of lexical changes was significantly different from chance ($\chi^2 = 6.35$, $p < .05$). At IM 80 both lexical and formal changes were identified above chance ($\chi^2 = 6.35$, $p < .05$ and $\chi^2 = 7.14$, $p < .01$, respectively). In Fig. 2 differences significant at $p < .01$ are indicated by open symbols.

Although the passages were identical in the two modes of presentation and, therefore, the number of syllables of interpolated material were identical, the time between original and test sentence necessarily differed with the two presentation methods. In the auditory presentation, the time interval was predetermined by the amount of time on the tape recording between the original and test sentence, but in the visual presentation, the time varied according to the S's reading speed and page-turning speed. One can estimate an average time for the various amounts of intervening material, based on an average reading speed that had been computed for Fairleigh Dickenson Ss.[1] When Ss read the stimulus material silently, the times between the end of the original sentence and test are estimated to average, for the four amounts of intervening material, 1, 4, 8.5, and 16.5 sec. With auditory materials, these times were 3, 7.5, 12.5, and 23 sec. Both the amount of interpolated material in syllables and the time intervals involved have been indicated in Fig. 2.

Comparing the pattern of results obtained in the two modalities, we find that the order of recognition is the same across the various types of changes. Initial recognition was higher when listening ($t = 3.89$, $df = 4$, $p < .01$), but relative amount of loss was higher in that mode. After delays up to 80 syllables of IM between the original sentence and test sentence, performance was similar in the auditory and visual conditions with respect to semantic, formal, and lexical changes.

A striking difference appeared in the results for detection of active/passive changes (see Fig. 2). With auditory presentation, this type of sentence change grouped with the other formal, nonsemantic changes. However, with written materials, the active/passive changes were recognized almost as readily as semantic changes. At IM 40 and 80, Ss recognized active/passive changes better when reading ($\chi^2 = 9.60$, $df = 1$ and $\chi^2 = 10.26$, $df = 1$, respectively, $p < .01$ in both cases).

DISCUSSION

This experiment confirms the results of Sachs (1967) that within 80 sec of interpolated discourse, Ss cannot distinguish between paraphrases and original sentences. Their retention of semantic content, however, is high for at least that period. The similarity in the general pattern of retention of auditorally and visually presented materials broadens the generality of the earlier findings to include reading. (Changes from active to passive and vice versa in the visual condition were an exception and will be discussed below.) Testing at earlier intervals after the original sentence has now revealed that forgetting of the exact words is essentially complete within 40 syllables and is substantial within 20 syllables (4-7 sec, depending on modality). The combined outcomes of quick memory loss of the exact words of a sentence together with good retention of the meaning of those words is consistent with the view that the material in a sentence is encoded in an abstract representation, and the exact words are rapidly forgotten.

A possible alternative interpretation of these results is that the interpolated material interfered more with the syntactic structure of the sentences than with the meaning. This view could explain the poorer recognition of the paraphrases without requiring the inference that sentence wording is lost during an encoding process. In the context of the present study, two objections can be raised against the explanation: (1) The degree of interference typically is a function of similarity. There is, with the present stimulus materials, no independent criterion for measuring similarity of the syntactic structures or of the meanings. (2) The interference hypothesis cannot explain the difference in the pattern of retention found with lexical (synonym substitution) as compared to semantic changes. In the lexical change condition, which showed the lowest level of recognition in both modalities, there were typically no "similar" words in the interpolated material that would have served as interference. If words as a class are considered "similar" and, therefore, interference, one would not expect to find the sustained high level of recognition of a change in a word, if that word affected the meaning of the sentence.

One result of the present study conflicts with the view that, regardless of modality of input, only an abstract representation of sentences is retained after a few seconds. Changes from active to passive voice and vice versa were well recognized at all test intervals if the material had been read. The active/passive changes were responded to much as semantic changes in this modality. We have no confident interpretation of this phenomenon but can explore two hypotheses. (1) A change in voice can be viewed as a change in focus or emphasis. Thus, although the truth-value meaning is the same for the active and passive version of a sentence, the two forms have a different function. In reading, the only cues to the semantics available are the words and the linguistic form of those words, whereas when one hears a sentence, the stress and inflection may help to communicate the meaning. In this sense reading is a more demanding task, and Ss would need to pay more attention to the form to get the meaning than they would when hearing. Moreover, since reading is a self-paced activity as compared to listening, perhaps Ss in this condition had more opportunity to become aware of the implications of voice and, therefore, were more aware of the change in focus when a change was made in voice. (2) Since reading combines a linguistic and nonlinguistic visual mode of perception, perhaps the degree of spatial reorganization of a sentence was an important variable. We have probably all experienced the phenomenon of remembering the location on the page of something we have read. This phenomenon has been confirmed experimentally by Rothkopf (1971). In the present experiment, active/passive changes caused the greatest rearrangement of words and, thus, the greatest visual change.

Memory for the exact wording of a sentence was not lost immediately in either modality. At the first test interval, Ss had already understood the sentence, as was indicated by their ability to recognize semantically changed sentences and to identify them as such. Yet, at that interval they were also able to recognize the nonsemantic changes very accurately and to identify them as changes in form. The brief capacity for correctly identifying nonsemantic changes indicates that a short-term memory operates during reading or listening. The failure to identify nonsemantic changes after brief intervals suggests that formal aspects of discourse do not normally enter long-term memory. Furthermore, the fact that at the first test interval semantic and formal changes were both correctly identified indicates that the understanding of sentences does not involve a sequential process of perception, storage in short-term memory and then encoding with simultaneous loss from STM. As was indicated by Lashley's example, the memory for the exact wording remains for a time after comprehension has occurred.

At the shortest delay interval tested, auditorally presented materials were remembered better than visually presented materials. This result may appear surprising, since (1) the time between the original presentation and test was longer for listening (3 sec for auditory as compared to 1 sec for visual), and (2) the self-pacing during reading would seem to allow Ss to adjust their reading speed for optimal learning. The superiority of memory for the auditorally presented material may be due to the availability of a special store for acoustic-phonetic input. The existence of such a store has been inferred from many studies, using word and digit lists, that have also revealed better retention after auditory than after visual presentation (e.g., Murray, 1966; Corballis, 1966; Crowder, 1970). In the present study, the fact that memory for the visually presented stimuli was above chance at the immediate test interval may reflect a visual short-term memory, a memory for internally generated phonetic images, or both.

It may be useful to view the special acoustic-phonetic store in terms of its function in linguistic processing. The primary linguistic mode is, of course, speech, in which acoustic signals arrive sequentially and disappear immediately. An acoustic-phonetic store would allow the hearer to retain the information, but this in itself is not a sufficient reason for a special store. Visual information arrives sequentially, too, and disappears. We suggest that it is the nature of syntactic encoding that makes such a store necessary. Linguistic processing is characterized by the recoding of sequential information into abstract meanings that do not typically reflect the temporal order in the input. That is, the linguistic mode of information transmittal is not an iconic one-to-one representation of experiences (as a pantomime of an event would be) but a complex recoding of the event, specifying an event or state and the participants in the event or state. With syntactic devices, information early in a sentence or discourse may depend on later information for interpretation, ideas can be embedded within other ideas, and so on. An acoustic-phonetic store would allow input arriving temporally to be processed by units rather than sequentially, would reduce the probability that errors of interpretation would occur, and would allow for "corrections" of misinterpretations on the basis of later information. In the present experiment, the fact that recognition of changes did not decrease substantially after about 20 syllables of interpolated material suggests that the loss of material from the store usually occurred some time during the next sentence. Conflict, or lack of conflict, of new information with assigned semantic encodings would typically become evident in this interval and reduce the need for holding material in storage. Forgetting the contents of the store would free the store's capacity for further incoming material.

REFERENCES

Bühler, K. Tatsachen und Probleme zu einer Psychologie der Denkvorgänge. Archiv für die gesamte Psychologie, 1908, 12, 1-92.

Corballis, M. C. Rehearsal and decay in immediate recall of visually and aurally presented items. Canadian Journal of Psychology, 1966, 20, 43-51.

Crowder, R. G. The role of one's own voice in immediate memory. Cognitive Psychology, 1970, 1, 157-178.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffries (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium.* New York: Wiley, 1951.

Murray, D. J. Vocalization-at-presentation and immediate recall, with varying recall methods. Quarterly Journal of Experimental Psychology, 1966, 18, 9-18.

Rothkopf, E. Incidental memory for localization of information in text. Journal of Verbal Learning & Verbal Behavior, 1971, 10, 608-613.

Sachs, J. S. Recognition memory for syntactic and semantic aspects of connected discourse. Perception & Psychophysics, 1967, 2, 437-442.

NOTE

1. E. Rothkopf, personal communication.

(Received for publication April 12, 1973;
revision received July 2, 1973.)