PRODUCTION NOTE

University of Illinois at Urbana-Champaign Library
Technical Report No. 60
SUPERFICIAL PROCESSING OF EXPLICIT INFERENCES IN TEXT
Rand J. Spiro and Joseph Esposito
University of Illinois at Urbana-Champaign
December 1977

Center for the Study of Reading
Technical Report No. 60

SUPERFICIAL PROCESSING OF EXPLICIT INFERENCES IN TEXT

Rand J. Spiro and Joseph Esposito

University of Illinois at Urbana-Champaign

December 1977

This research was supported by the National Institute of Education under Contract No. US-NIE-C-400-76-0116 and by a grant from the University of Illinois Research Board to the first author. Helpful comments by Richard C. Anderson, Allan Collins, and Andrew Ortony are gratefully acknowledged.
Abstract

The hypothesis that pragmatic inferences presented in text are taken for granted, superficially processed, and not stably or enduringly represented in memory was investigated. Stories were read which in some conditions contained information vitiating the implicational force of explicit inferences. The vitiating information was presented either before or after the inferences. In Experiment I, errors in memory for the inferences were prevalent in the "after" but not the "before" condition. Two kinds of errors were made: saying the inference had not been presented in the story; or, if it was remembered as having been presented, altering the specific content of the inference to produce the opposite of what was actually presented. The latter errors produced coherence with the vitiating information, and subjects were not able to differentiate these errors from correct responses. In Experiment II, the results of Experiment I were replicated, and a "spontaneous correction" interpretation was rejected. The results of both experiments combine to support the hypothesis of superficial processing and unstable representation of explicit inferences. The results provide a link between processes occurring at comprehension and recall in the State of Schema model of accommodative reconstruction.
Language is characterized by considerable semantic predictability. Explicitly presented information may be logically or probabilistically implied by earlier parts of a discourse. How is such information processed and stored? The view typically espoused in discourse processing and representation theories is that all presented information, including inferences, receives sufficient processing to be encoded in long-term memory (e.g., Anderson & Bower, 1973; Fredericksen, 1975; Kintsch, 1974; Meyer, 1974). For example, Kintsch assumes "that subjects process and store [an inference] whether or not it is presented explicitly" (p. 154). This view will be referred to as the "storage of presented inferences" (SPI) hypothesis.

An alternative hypothesis is that predictable information, however central to a discourse, is taken for granted, processed only superficially and receives an attenuated cognitive representation or no enduring representation at all. This view will be referred to as the "superficial processing of presented inferences" (SPPI) hypothesis. It is important to determine which of these hypotheses is correct for several reasons. Obviously, the issue is basic for any discourse processing or representation model. Additionally, the SPPI hypothesis is a crucial link in a theory of accommodative reconstruction processes in memory for discourse (Spiro, 1977).

Spiro (Note 1) found a pervasive tendency for subjects to produce predictable meaning-changing distortions and importations in text recall under certain conditions. In general, when subsequently encountered information contradicted continuation expectations derived from a target story, the story frequently was reconstructed in such a way as to reconcile or cohere with the continuation information. This process of inferring
the past based on the present was termed accommodative reconstruction. After a long retention interval, subjects tended to be more confident that their accommodative recall errors had actually been included in the story than they were confident about the accurate aspects of their recall. Why should such gross errors occur and then be assigned such high confidence? Part of the answer surely involves their function in producing coherence. Still, it is somewhat surprising that subjects should be so sure they read information that bore not even a distant inferential relationship to what they actually did read.

Spiro (1977) suggested that the basis for such an effect may be in the way inferentially related information is treated at the time of comprehension. If the SPPI hypothesis is correct, individuals should know (at least tacitly) that considerable amounts of predictable or derivable information they have encountered will not be available in memory. In that case, recall would typically involve deriving a lot of missing information. Accordingly, it would not be surprising that subjects faced with memories that lack coherence would assume that missing reconciling information was presented but only superficially processed at comprehension. The information could then be derived at recall with high confidence. Hence the capacity for restructuring the past based on the present.

The present experiments were designed to test the SPPI hypothesis. Stories were presented which contained information A, B, and C such that B was strongly implied by A except in the presence of C. For example, the A, B, and C elements in one story (about a demonstration by a karate
champion) could be paraphrased as follows:

A: The karate champion hit the block
B: The block broke
C: He had had a fight with his wife earlier. It was impairing his concentration. He doesn't perform as well when he can't concentrate.

C is either presented prior to A and B (C-Before), after A and B (C-After), or not at all (No-C). When C is not included in the story, if SPPI is correct, the B element would be taken for granted, processed only superficially, and not stably represented. It would be derivable if needed. However, if C is presented after A and B, memory for B should be impaired since B was not stored and C will block its derivation from A at the time of test. On the other hand, if C occurs in the text prior to A and B, then B is not strongly implied by A. B cannot be taken for granted with the assumption that it can be generated later if needed. Here B should be stably represented and memory for B should not be impaired.

However, if SPI is correct, memory for B should not be affected by whether C is before or after A and B, since B is stored whether it is implied by A (C-After) or not implied by A (C-Before). Two objections to this argument can be made. The information might be stored, but remembering C might lead to a decision that the memory for B must be mistaken (a kind of output interference). However, C is present whether it occurs before or after A and B, so such an explanation would not account for differential effects of C-placement. The other possibility is that B is represented in C-After, but
the representation is altered or corrected when the C information is encountered. This possibility will be investigated in Experiment II.

In the first experiment, the following predictions of the SPPI hypothesis were tested. More errors in response to questions about the presented implied information (B) should be made in the C-After than in the C-Before conditions. Errors can be erroneous judgments that nothing about the implied information was presented, called B-Mention errors (e.g., the story did not mention whether the block was broken), or, when the subject believes that something about B was mentioned, remembering incorrectly what was specifically said in the direction of conforming with the C information, called B-Incorrect errors (e.g., it said in the story that the block did not break when he hit it). Confidence in errors of the latter kind will also be analyzed. If subjects are as confident about these errors as they are about their accurate responses, it would be even more difficult to maintain the hypothesis that the implied information was represented.

In the No-C condition, B-Mention errors may occur since B would not be represented according to the SPPI hypothesis. The more important prediction regarding the No-C condition is that B-Incorrect errors should not occur more often than in the C-Before condition. Otherwise, the differences between C-Before and C-After might be attributable to heightened accuracy due to greater salience of the implied information in the former condition rather than greater inaccuracy due to a failure to store the implied information in the latter condition.
Experiment I

Method

Subjects. Twenty students in an introductory educational psychology class at the University of Illinois served as subjects. Participation in the experiment partially fulfilled a course requirement.

Materials. A 24 page booklet entitled, "What Became of the Old Classmates?" was constructed. The story featured eleven character vignettes of approximately 250 words each. Three of the vignettes were used as unanalyzed buffers, one at the beginning of the book and two at the end. Also, there was an introductory page which said the vignettes would describe events involving several old classmates who had gone their individual ways and were now being observed several years after graduation. The other eight vignettes contained the target information for the study. Each vignette contained information, A and B, such that A pragmatically implied B. The vignettes also contained information, C, which lessened the extent to which B was implied by A (see the karate champion example presented earlier). In one condition, C was presented after A and B (C-After condition). In another, C was presented before A and B (C-Before condition). Each vignette was presented on two pages with C always on a separate page from A and B. In a third condition, there was no C information (No-C condition). Booklets were constructed in such a way that there were two random orderings of the eight experimental passages.

A test booklet was constructed which contained 7 questions for each of the vignettes. The order of the sets of questions in the booklets was the same as the order of the vignettes which each subject received. The set
of questions for each vignette was introduced by the character's name and a brief description (e.g., "Steve: the passage about the karate champion"). The first question in each set asked whether any B information had been mentioned (e.g., "Did it say in the story whether or not the block broke when he hit it?"). These will be referred to as B-Mention questions. Remember that B was always mentioned. The second question had the following form: "If you answered the last question 'no' do not answer this question. If you answered 'yes', did it say X in the story?" For the karate champion example, X was "The block did not break." The questions were phrased so that "yes" and "no" responses were correct equally often. The subjects answered and rated the certainty of their response on a six-point scale. Errors on the second question will be called B-Incorrect responses. The next two questions had the same format as the two just described. These dealt with some other arbitrarily chosen information given in the vignette. The fifth and sixth questions called for subjects to make a judgment about the character's popularity and academic success in high school. These questions were intended to help mask the true intentions of the experiment. The seventh question elicited recall of the C information (e.g., "What was Steve's problem the day of the demonstration?").

Procedure

Subjects were randomly assigned to conditions and were run in groups of four or less. There were four subjects in the C-Before and C-After conditions and eight subjects in the No-C condition. The same investigator tested all subjects. An instruction sheet was passed out, and subjects were told to
read it silently while the investigator read it aloud to them. The instructions were as follows:

You will read a story. Following the completion of the story, you will be asked some questions about it. One kind of question you will be asked will involve your personal reactions to the characters in the story as they are described in various vignettes.

A crucial aspect of this study is that we are interested only in your first reactions, rather than your reactions after you've taken time to think about what you've read. Accordingly, we will ask that you read as quickly as possible and that once you turn a page you never look back at the preceding pages. However, be sure to understand all that you read, or you will be unable to answer some of the questions. Read and understand. Simply do it as quickly as is reasonable.

Also, one again, to ensure rapid progress and to ensure that your first reactions aren't re-evaluated, never turn back to previously completed pages!

Subjects were allowed nine minutes to read the story. To ensure that subjects had sufficient time to complete the target passages, subjects were instructed to circle the last word read when time was called. All subjects finished the target passages. The question booklet was then distributed and the experimenter worked through with them the first set of questions (which dealt with a buffer passage). Subjects were instructed to work through the questions in order, to answer every question, and not to change any answers.
after they were written. They were allowed sufficient time to complete the questions.

Results

There were no significant effects involving list order. Accordingly, all analyses are collapsed across lists.

Two kinds of errors could be made: saying B was not mentioned (B-Mention errors) and saying B was mentioned but responding with the incorrect specifics (B-Incorrect errors). Combining both types of errors, significantly more errors were made in the C-After than in the C-Before condition, \( t(18) = 2.37, p < .015 \) (means and standard deviations are presented in Table 1).

In this and all other analyses where significant differences are reported, the trends were in the same direction for each of the eight passages. Group differences accounted for 23.7% of the variance. More B-Mention errors were made in C-After than in C-Before (see Table 1), but the difference was non-significant. Although the incidence of B-Mention errors in the No-C condition was not of particular interest, the high proportion of such errors provides an indication that the implied information was not directly stored.

In the most important analysis, the conditional probability of a B-Incorrect error given that B-Mention was correct was significantly higher for C-After than C-Before or No-C, \( F(2,25) = 4.48, p < .025 \). The probability of this type of error was significantly higher \( (p < .05 \) in each Newman-Keuls
Processing of Inferences

comparison) for C-After than for the other two conditions, which did not differ from each other. Group differences accounted for 26.4% of the variance.

B-Incorrect errors for C-After indicate that the predictable information is superficially processed and not stably represented. The question remains whether these errors are guesses or the result of a natural and undetectable accommodative process of reconstruction of the past based on incomplete memorial data (Spiro, Note 1, 1977). If subjects are able to detect their errors, as indicated by their confidence ratings, the guessing interpretation would be supported. If not, that would suggest the operation of accommodative reconstruction. On the filler questions unrelated to the experimental manipulation, subjects were able to detect their errors. For subjects who made the equivalent of B-Incorrect errors on the fillers, confidence was significantly higher for correct responses ($M = 5.19; SD = .823$) than for incorrect responses ($M = 4.49; SD = .996$), $t(16) = 2.38, p < .015$. However, for the B information in the C-After condition, confidence when correct ($M = 5.37; SD = .586$), did not differ significantly from confidence in B-Incorrect errors ($M = 5.06; SD = 1.32$), $t(7) < 1$. For the eight C-After subjects who made B-Incorrect errors, four had higher mean confidence when B was correct, three had higher mean confidence when B was incorrect, and there was one tie. For these eight subjects, only one was better able to detect their B-Incorrect errors than their errors on the filler questions, as indicated by comparison of average confidence when correct minus confidence when incorrect for the two kinds of questions ($p < .04$ in a one-tailed sign
test). The confidence analyses support the accommodative reconstruction interpretation of B-Incorrect errors over the guessing interpretation.

Finally, there were no differences among the groups in combined error frequency on the filler items, $F < 1$ (means of 2.6, 2.7, and 2.5 for the C-After, C-Before, and No-C conditions, respectively).

Discussion

The results support the hypothesis that pragmatic inferences presented in text are superficially processed and do not receive a stable and enduring representation in memory. In the C-After condition, subjects tended either to report that the inference was not presented in the text or that the opposite of the inference was presented. It is difficult to retain the notion that inferences are deeply processed and stably encoded when the C-After manipulation can produce errors like remembering the block was not broken when the karate champion hit it. The results cannot be attributed to interference produced by the inference-vitiating C information at output, since the C-Before subjects would also be subject to such interference. Neither can the results be attributed to differential availability of C at output, perhaps due to primacy/recency effects related to the position of C in the text, since the information was almost always recalled. Also, unimportance of the B information is not a viable alternative since B tended to be central to the story (e.g., in a story about a karate champion's performance, information about his success in the demonstration is certainly important). If one were to argue that B is less important structurally because it is predictable, that is a novel conception of importance. This issue will be addressed in the General Discussion.
One interpretation that remains as a viable alternative is that subjects do deeply process and stably encode the presented inferences, but "correct" their representation when the inference-vitiating information is presented. The second experiment tested this hypothesis.

**Experiment II**

If subjects are storing B and then changing or correcting it at the time C is presented, errors on B should occur in the C-After condition no matter how soon the test is administered after reading. However, if the SPPI hypothesis is correct, when delay intervals are brief enough some surface memory for the superficially processed B information may remain, reducing the number of B errors. Accordingly, in this experiment subjects were tested either immediately after reading each story (Interspersed Questions condition) or, as in Experiment I, after the entire set of stories had been read (Questions-After condition). Again, the C-Before and C-After manipulations were employed.

**Method**

Twenty-four subjects from the same subject pool and participating for the same inducements as in Experiment I were randomly assigned to one of the eight between-subject cells determined by factorial combination of the two levels of C-placement, the two levels of question placement, and two list orders. In the Interspersed Questions condition, subjects received a single booklet in which each story was followed by the test on that story. In the Questions-After condition, subjects read the stories in one booklet and then went on to the test in a separate booklet. In all other details, the method was the same as in Experiment I.
Results and Discussion

Once again, no effects involving list order were found and the analyses were collapsed across that factor. Two-way (C-After and C-Before X Questions-After and Interspersed Questions) between-subjects analyses of variance were carried out with combined frequency of errors (B-Mention plus B-Incorrect), frequency of B-Mention errors, and the conditional probability of B-Incorrect errors given a correct B-Mention response as dependent variables.

In the total error analysis, all effects were significant. More errors were made in C-After than C-Before, $F(1,20) = 14.31, p < .001$ (means and standard deviations for all the dependent variables are found in Table 2). More errors were made in the Questions After than in the Interspersed Questions conditions, $F(1,20) = 5.91, p < .025$. Most important, the interaction was significant, $F(1,20) = 4.67, p < .05$. Looking at simple effects, significantly more errors were made in the C-After than the C-Before conditions only in the Questions-After condition ($F(1,20) = 17.67, p < .001$ for Questions-After; $F(1,20) = 1.31, p > .25$ for Interspersed Questions). Also, significantly more errors were made in the Questions-After than in the Interspersed Questions condition for C-After subjects but not for C-Before subjects ($F(1,20) = 10.53, p < .005$ for C-After; $F < 1$ for C-Before). These results clearly militate against the corrected representations hypothesis.
With frequency of B-Mention errors as the dependent variable, the C-Position effect was significant and the other two effects were marginally significant (see Table 2). More errors were made in the C-After than in the C-Before conditions, $F(1,20) = 13.19$, $p < .005$. In Experiment I, the same trend occurred but the difference was not statistically significant. More errors tended to be made in the Questions-After than in the Interspersed Questions condition, $F(1,20) = 3.57$, $p < .075$. The interaction also approached significance $F(1,20) = 3.57$, $p < .075$. Simple effects analyses indicated more errors in the C-After than in the C-Before conditions only in the Questions-After condition ($F(1,20) = 15.23$, $p < .001$ for Questions After; $F(1,20) = 1.52$, $p > .10$ for Interspersed Questions). More errors were made with Questions-After than with Interspersed Questions only for the C-After condition ($F(1,20) = 7.13$, $p < .025$ for C-After; $F < 1$ for C-Before).

Finally, all effects were significant with the conditional probability of a B-Incorrect error given a correct B-Mention response as the dependent variable (see Table 2). An unweighted means analysis was used because the conditional probability was undefined (denominator equal to zero) for one subject in the C-After/Questions-After condition. B-Incorrect errors were more likely to be made in the C-After than in the C-Before conditions, $F(1,19) = 5.04$, $p < .04$, and in the Questions-After than in the Interspersed Questions conditions, $F(1,19) = 6.26$, $p < .025$. The interaction was again significant, $F(1,19) = 4.65$, $p < .05$. The results of the simple effects analyses paralleled those with the other two dependent variables. Errors were more likely in the C-After than in the C-Before conditions only in the
Questions-After condition \((F(1,19) = 9.24, p < .01\) for Questions After; \(F < 1\) for Interspersed Questions). Errors were more likely with Questions-After than with Interspersed Questions only for the C-After condition \((F(1,19) = 10.62, p < .005\) for C-After; \(F < 1\) for C-Before). Again, the results strongly conflict with the predictions of the immediately-corrected-representation hypothesis.

The results of Experiment II replicated those of Experiment I. Furthermore, they demonstrate that the C-After effect is not due to changing a stored representation of the B information (the explicit inference).

General Discussion

The present experiments demonstrated that inferences presented in text are superficially processed and tend to have very unstable representations in semantic memory, if they are represented at all. Various alternative hypotheses were discredited. Such a conclusion is troublesome for many current structural approaches to discourse comprehension and memory (e.g., Meyer, 1975). All semantic information presented in text, including explicit (as well as some implicit) inferences, is supposed to receive an internal representation. The present experiments create doubts about such claims. It might be argued that the inferences in the present experiment (the B information) lacked structural importance; i.e., they were low in the hierarchical text representations, thereby explaining the poor memory performance. Such an argument accounts for omissions better than it does the prevalence of remembering the opposite of what was presented. Furthermore, such an argument seems to reconceptualize structural importance. In conventional
conceptualizations, the B information in the present experiments should have the same location in a text hierarchy whether the C information comes before or after; the stories contain the same content in both conditions. Furthermore, the B information tended to be central. Rather, it might be argued that predictable information is less "important" in the sense that one need not pay as much attention to it. This sense of importance accounts for the present data but extant discourse structure theories do not provide for importance variations on such a dimension. Clearly, such theories provide no basis for predicting the differential memory for implied information found in the present study.

The present experiments further our understanding of the processes involved in accommodative reconstruction (discussed in the introduction). Spiro (1977) proposed a heuristic model, the State of Schema (SOS) model, to account for the processes involved in accommodative reconstruction. One of the questions addressed by the SOS model was how subjects could be so sure they had read what they had not read. The answer proposed was that much of what is read is predictable. When predictable information is encountered, it is taken for granted and passed over. If needed later, it could be derived from whatever other information made it predictable. Perhaps some kind of fast-fading "left-to-be-derived" marker is attached to information that would be the basis for generating the superficially processed predictable information. In any case, subjects would know (at least tacitly) that memory tends to be incomplete, with derivable information missing. When readily retrieved information at recall is insufficient to produce a coherent view of the past,
the assumption is made that the lack of coherence is attributable to the absence of superficially processed derivable information. That information would then be generated, even though it may never have been presented. The crucial assumption in this account, i.e., predictable information is superficially processed and not stably represented in memory, has been validated in the present experiments. Briefly digressing, it remains for further investigation to determine whether the sensitivity to opportunities for superficial processing, a kind of "semantic automaticity," may be a prerequisite of skilled reading absent in less able readers.

Finally, the present experiments provided further evidence for accommodative reconstruction itself. When subsequent information retroactively affected the coherence of previously comprehended information, errors in recall in the direction of enhancing coherence were evident. The accommodation was sometimes active, as in remembering (with high confidence) the opposite of what occurred (e.g., the block did not break), and was sometimes passive, as in saying nothing was mentioned (e.g., about whether the block broke).
Reference Note

References


Fredericksen, C. H. Effects of context induced processing operations on semantic information acquired from discourse. Cognitive Psychology, 1975, 7, 139-166.


Table 1
Error Frequency as a Function of Condition and Error Type in Experiment I

<table>
<thead>
<tr>
<th>Condition</th>
<th>B-Incorrect</th>
<th>B-Mention</th>
<th>Probability of B-Incorrect given correct on B-Mention</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-After</td>
<td>3.8 (1.32)</td>
<td>2.6 (1.43)</td>
<td>.21 (.15)</td>
</tr>
<tr>
<td>C-Before</td>
<td>2.2 (1.62)</td>
<td>1.9 (1.52)</td>
<td>.08 (.11)</td>
</tr>
<tr>
<td>No-C</td>
<td>3.1 (1.25)</td>
<td>2.8 (1.25)</td>
<td>.06 (.10)</td>
</tr>
</tbody>
</table>

Note: Standard deviations are given in parentheses. The maximum possible error frequency is 8.
Table 2

Error Frequency as a Function of C-Position, Question-Placement, and Error Type in Experiment II

<table>
<thead>
<tr>
<th>Condition</th>
<th>B-Incorrect</th>
<th>B-Mention</th>
<th>Probability of B-Incorrect given correct on B-Mention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B-Mention plus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question After</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-After</td>
<td>5.3 (2.58)</td>
<td>4.5 (2.26)</td>
<td>.39 (.379)</td>
</tr>
<tr>
<td>C-Before</td>
<td>1.7 (1.03)</td>
<td>1.3 (1.03)</td>
<td>.05 (.074)</td>
</tr>
<tr>
<td>Interspersed Questions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-After</td>
<td>2.5 (1.05)</td>
<td>2.3 (1.03)</td>
<td>.03 (.068)</td>
</tr>
<tr>
<td>C-Before</td>
<td>1.5 (0.55)</td>
<td>1.3 (0.82)</td>
<td>.02 (.052)</td>
</tr>
</tbody>
</table>

Note: Standard deviations are given in parentheses. The maximum possible error frequency is 8.

No. 2: Asher, S. R. Sex Differences in Reading Achievement, October 1977.

* Available only through ERIC


*No. 8: Mason, J. M. Questioning the Notion of Independent Processing Stages in Reading, February 1976. (Journal of Educational Psychology, 1977, 69, 288-297)


No. 34: Bruce, B. C. Plans and Social Actions, April 1977.


No. 54: Fleisher, L. S., & Jenkins, J. R. Effects of Contextualized and Decontextualized Practice Conditions on Word Recognition, July 1977.


No. 59: Mason, J. M. Reading Readiness: A Definition and Skills Hierarchy from Preschoolers' Developing Conceptions of Print, September 1977.


This page is intentionally blank.