Passage Recall: Schema Change and Cognitive Flexibility

Paula T. Hertel  
*Trinity University*, phertel@trinity.edu  

M. Codsen  

P. J. Johnson

Follow this and additional works at: [https://digitalcommons.trinity.edu/psych_faculty](https://digitalcommons.trinity.edu/psych_faculty)

Part of the Psychology Commons

**Publication Details**  
Journal of Educational Psychology

**Repository Citation**  

This Article is brought to you for free and open access by the Psychology Department at Digital Commons @ Trinity. It has been accepted for inclusion in Psychology Faculty Research by an authorized administrator of Digital Commons @ Trinity. For more information, please contact jcostanz@trinity.edu.
Passage Recall: Schema Change and Cognitive Flexibility

Paula T. Hertel, Merith Cosden, and Peder J. Johnson
University of New Mexico

This experiment investigated the effects of subsequent related information and individual differences in cognitive flexibility on prose recall. Subjects read a passage and then were given either consistent or contradictory incidental information. Errors in cued recall, reflecting the nature of the subsequent information, were more frequently produced after a 3-week delay than after 2 days. These results were consistent with Spiro’s findings with free recall. In addition, 3-week subjects were more confident about correct recall than errors, indicating that errors resulted, in part, from retrieval processes. The negative relationship of spontaneous flexibility and the positive relationship of adaptive flexibility to constructive error are interpreted in terms of storage and retrieval effects in memory.

The present study was concerned with the dynamic nature of memory schemata and the extent to which cognitive style may relate to recall accuracy. Evidence from several sources (Anderson, Spiro, & Anderson, 1978; Bransford & Johnson, 1972; Pichert & Anderson, 1977; Royer & Cable, 1976; Sulin & Dooling, 1974) supports the notion that information is assimilated into existing knowledge structures (schemata). From the perspective of schema theory (e.g., Piaget, 1952), a knowledge structure should provide a framework not only for the assimilation of related information consistent with its content but also for the accommodation of its content to related but disparate information.

The issue of accommodation has important implications for education, some of which have been suggested by Anderson (1977). In terms of schema theory, all knowledge acquisition is viewed as modification and differentiation of existing schemata. If the goals of education involve the change of knowledge, then we must be concerned with the modification of existing schemata when confronted with information requiring varying degrees of accommodation. Furthermore, educational theory must be concerned with individual differences that are predictive of the degree to which existing schemata assimilate or accommodate conflicting new inputs.

In an attempt to systematically investigate mechanisms of schema change, Spiro (1975, 1977) argued that varying the type of subsequent information should affect the degree of error in what is remembered about the original event. For some subjects, he presented incidental subsequent information that was either consistent with or contradictory to information in a story that subjects had previously read, predicting transformational errors in story recall for only those subjects who heard contradictory information. Half of Spiro’s subjects received memory instructions, whereas the others were deceived into believing that they would be asked to evaluate the material at a later time. Since memory instructions demand isolation of passage information, errors were predicted for the deception condition only. Retention was tested after 2 days, 3 weeks, or 6 weeks. Following story recall, all subjects were asked to rate each sentence in their recall protocol, according to their degree of confidence that it accurately represented information stated in the original story.

As expected, Spiro found that all subjects in the memory condition and those deceived subjects who heard consistent subsequent information were essentially accurate in their recall. However, after long retention intervals (3 or 6 weeks), the deceived subjects
who heard contradictory information made errors that reflected the nature of that information. When similarly treated subjects were tested after 2 days, these errors did not occur. Therefore, schemata were not immediately modified by the presentation of the subsequent information. Rather, during the passage of time, subjects may have lost the ability to distinguish information presented in the story from information inferred from the “aside” that was delivered subsequently. This loss of ability may result from confusion, or it may reflect hypothesized changes in the story schema.

A stage analysis (cf. Crowder, 1976) of the locus of schema change would indicate that the effects in Spiro’s experiment occurred during information storage or at the time of retrieval. Confidence ratings were obtained to address this issue. The subjects who had heard contradictory information and were tested following 3 or 6 weeks were more confident in errors influenced by the contradictory information than they were confident in accurate recall. This finding suggested to Spiro that the same reconstructive process was used to produce accurate and inaccurate recall; he concluded that errors were not produced at retrieval, through the use of a conscious guessing strategy, but resulted from modification of the schema during storage.

The first purpose of this experiment was to determine if Spiro’s results were valid under more stringent test conditions that in turn would provide more evidence concerning the locus of schema change. The fact that Spiro did not request confidence ratings until after the story was recalled allows for the possibility that subjects who were tested after longer delays had difficulty retrieving enough information from the story to comply with task demands, accessed the subsequent information, and consciously or unconsciously used it to elaborate their recall. Regardless of the nature of the retrieval processes, the test demands would require subjects to integrate all information to write a well-connected story. Thus, it is possible that the subsequent confidence ratings reflected integration at recall, as well as the type of retrieval processing.

The following experiment departs from Spiro in two important ways. First, a cued-recall task was employed to reduce the integrative demands of the test situation; if the errors observed by Spiro reflected processes operating prior to retrieval, then the type of recall test should not affect the qualitative nature of the errors, although their magnitude may be decreased by the restrictive nature of the cues. Second, confidence ratings were requested immediately following each answer, requiring subjects to monitor their retrieval efforts.

The second purpose of this experiment was to examine individual differences in recall accuracy from the theoretical perspective of schema change. Since research concerned with knowledge structures places a large emphasis on idiosyncratic processing (Perfetti & Lesgold, 1977; Spiro, 1977), it is reasonable to suggest that identifiable cognitive styles may correspond to processes involved in schema change. Memory for events may be transformed according to the type of person who is processing the event.

Cognitive flexibility has been identified by Guilford (1967) as an important factor in discriminating among individual abilities. Defined as the ability to shift avenues of thinking in order to perceive and process information about a situation in different ways, cognitive flexibility is similar, on an intuitive level, to a description of schema change. In studies by Guilford and others (Houston & Mednick, 1963; Munsinger & Kessen, 1966; Vidler & Karan, 1975), cognitive flexibility has been associated with the ability to tolerate and structure ambiguity. Furthermore, Kaplan (1952) reported that subjects with a high tolerance for instability more readily recalled equivocal parts of stories than did other subjects. These findings suggest that at least one type of cognitive flexibility, spontaneous or adaptive, should be related to performance in this experiment.

Spontaneous flexibility is conceptualized as the ability to think in varied directions in an unstructured situation, such as the incidental learning task in this experiment. Tests for spontaneous flexibility do not require flexibility for their solution, although the magnitude of the score depends on the spontaneous use of these processes. Simi-
larly, the incidental nature of the subsequent information in this experiment does not require that it be kept separate from story information. The test for another subfactor, adaptive flexibility, does require flexible thought processes for problem solving, and therefore characterizes flexibility in situations that demand it. The test conditions of this experiment may be viewed as placing certain restrictions on subjects' recall. Logically, constructive effects that result from storage mechanisms should be related to spontaneous flexibility, whereas those occurring at testing should be related to the amount of flexibility required by the recall task, and thus to adaptive flexibility. More generally, cognitive flexibility was expected to be a reliable predictor of memory performance, in addition to the characteristics of the experimental setting.

Method

Materials

The material to be recalled was adapted from Spiro's (1975, 1977) passage about a young woman and man who met in college, developed a romantic relationship, and made marriage plans. The central theme of the story concerned the young man's desire not to have children and his hesitancy to inform the woman about this desire. He finally does tell her, and the story ends with her acceptance of his decision, and making marriage plans. The central theme of the story concerned the young man's desire not to have children and his hesitancy to inform the woman about this desire. He finally does tell her, and the story ends with her acceptance of his decision, and making marriage plans.

Test booklets consisted of 10 questions, prompts, or statements with a blank to be filled in; each of the 10 items was presented on a separate page. The first 5 filler items concerned details from the first part of the story, and answers were not viewed as potential sources of constructive errors. The remaining 5 items addressed the nature of the relationship, and were constructed to allow for intrusions emanating from the subsequent bias. All items are presented in the following display:

Cued Recall Test
1. When Bob and Margie met they were both ____ years old.
2. Bob was majoring in ____.
3. Margie was majoring in ____.
4. They didn't know each other until ____.
5. Bob began to think he would like to marry Margie, after he had known her for ____.
6. Margie's feelings about Bob could be characterized as: ____.
7. When Bob asked Margie to marry him, she: ____.
8. Bob and Margie's feelings about having children were: ____.
9. What were Margie's goals for the future?
10. How did the story you read end?

For Items 6–10, a priori scoring continua were developed to reflect the direction of possible errors. Zero points were established for essentially accurate recall, the negative range was reserved for errors leading to a distortion of the conflict in the relationship, and the positive range was relegated to errors that would increase the harmony. For Item 7 the continuum ranged from hesitancy to say yes (−5 to −1), to acceptance (0), to enthusiasm (+1 to +5). Since magnitude was not considered as important an aspect of the replication as direction of error, the numerical value of the error score was not predetermined apart from establishing a positive correspondence between the degree of error and the magnitude of the score.

Two tests of spontaneous flexibility were employed (Object Naming and Utility). For the Object Naming test, subjects were required to name instances of each of two successively presented categories (liquids and plants), and scores were determined by the number of subcategorical shifts (from fruit juices to medications, e.g.). The Utility test was similar; subjects must name uses for each of two objects (brick and pencil), and the number of functional shifts was scored. For the Match test of adaptive flexibility, subjects were instructed to vary their solutions to the problem of removing a specified number of matches and leaving only matches that contributed to squares; scores were tabulated by counting the number of different solutions (removing corner matches vs. center matches, e.g.).

Subjects

Seventy students volunteered for two sessions, participating in groups of 6–10 during the acquisition phase and 2–5 during the test phase. They received credit toward their introductory class grades. Groups were assigned to acquisition conditions on the basis of maintaining an equal male–female ratio across conditions.

Procedure

All subjects were told that data would be collected for several short, unrelated experiments during the two sessions, and that all experiments would be explained at the end of the second session. Spiro's (1975) incidental instructions for processing the story were then delivered:

This is an experiment concerned with changes in the way people react to stories involving interpersonal relations when there is a delay prior to giving the reactions. You will read a story about two people. The story is true in all respects. I knew both of the people and can vouch for the accuracy of the story. What I would like you to do is think about and react to the story. At the second session I will ask you various kinds of questions concerning your reactions to the story. Are there any questions?
Subjects were told to use 3 minutes to read the story. After the allotted time the stories were collected, and approximately 8 minutes were employed by assigning subjects to second-session dates, collecting phone numbers for reminder calls, and instructing subjects not to discuss the experiment between sessions. As the experimenter reviewed the story requirements for the second session, she very casually delivered one of the two types of subsequent bias. (a) Bob and Margie did get married and are very happy together to this day (harmony) or (b) they never did get married; the engagement was broken, and they never saw each other again (conflict). Therefore, consistent and contradictory subsequent information was presented to subjects reading each story end. For example, the harmonious bias was consistent with the harmonious story end and contradictory to the story end that reported conflict.

The Object Naming test, the Utility test, and the Match test were then administered in that order; procedures included reading the printed instructions aloud and observing the time limitations for each section. (For additional details, see French, Ekstrom, & Price, 1963.)

Half of the subjects in each acquisition group returned for testing 2 days later; the remaining subjects, after 3 weeks. The cover page of the test booklet contained the following instructions, which were read aloud by the experimenter:

I'm sorry, but we deceived you. This is not a study of how people react to situations involving interpersonal relations. It is a study of memory. As you will recall, at the last session you read a story. What we would like you to do is to try to recall the story as best you can. The following pages contain questions concerning the story you read. Base your answer to each question on your memory for the story and not on your personal reactions. You must answer every question, and you must answer them in the order in which they are presented. Do not look ahead or behind the question you are working on. After you have answered a question, please rate your confidence in the answer by placing a number on the line at the bottom of the page. Use the scale below. For example, if you write “1” you will be indicating that you are very uncertain that the meaning of the sentence you wrote was explicitly expressed in the story, “5” will indicate moderate certainty, and “9” shows absolute certainty. Are there any questions?

Ten minutes were allowed for recall.

Results and Discussion

Cued Recall

Scoring. Two scorers, blind to the type of subsequent information and retention interval, independently determined error scores for Items 6–10. Agreement about direction of error was 100%; agreement about error magnitude was 89%. Where differences existed, their absolute value was 1, and all differences were resolved by the scorers so that complete agreement was finally reached.

The constructive error score (CES) for each subject was computed by subtracting from the individual item score with the highest absolute value the absolute value of the highest item score with the opposite sign. For example, if a set of item scores were 0, −3, 1, −1, 0, the CES would be −2. (Item scores of opposite directions rarely occurred within a subject's recall booklet.) This method of error scoring was chosen by Spiro (1975, 1977) to reflect that evidence for reconstructive memory does not depend on the number of errors. Rather, it is determined by the magnitude of qualitative change, which may be as likely to occur in one sentence alone as in several sentences.

Constructive errors. Table 1 presents the mean CES for each condition of this experiment, along with the means for the corresponding free recall condition in Spiro's experiment. The direction of errors (positive or negative) for cued recall are identical to those for free recall, with the exception of cells representing the harmonious story end, consistent subsequent information, 2-day delay. However, neither of these means is apparently different from zero.

For the cued-recall experiment, the major finding was that the type of subsequent information interacting with the length of the retention interval reliably influenced memory for the story. This interaction was tested by allowing for all other effects in the linear model, due to the nonorthogonality of the design (Appelbaum & Cramer, 1974); $F(1, 62) = 11.237, MS_e = .0084, p < .005$. Of the remaining possible main effects and interactions, only the main effect of the subsequent information was reliable beyond the .10 level (allowing for the other effects), $F(1, 62) = 39.574, MS_e = .0084, p < .001$. The interpretation of these effects is clear. When subsequent information indicated eventual disharmony (breaking up), subjects tended to impose or emphasize conflict in the story they read. When subsequent information indicated eventual harmony (marriage and happiness), conflict in the
Table 1
Mean Constructive Error Scores for Cued Recall and Free Recall (Spiro, 1977) Under Incidental Memory Instructions

<table>
<thead>
<tr>
<th>Story end</th>
<th>Bias</th>
<th>2 days</th>
<th>3 weeks</th>
<th>2 days</th>
<th>3 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harmony</td>
<td>Consistent</td>
<td>.5</td>
<td>1.6</td>
<td>-.2</td>
<td>.9</td>
</tr>
<tr>
<td></td>
<td>Contradictory</td>
<td>-.1</td>
<td>-2.1</td>
<td>-1.4</td>
<td>-3.1</td>
</tr>
<tr>
<td>Conflict</td>
<td>Consistent</td>
<td>-.5</td>
<td>-1.2</td>
<td>-.4</td>
<td>-.3</td>
</tr>
<tr>
<td></td>
<td>Contradictory</td>
<td>1.0</td>
<td>2.3</td>
<td>1.0</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Differences between the two experiments in magnitude of CES are apparent in Table 1. For the 3-week contradictory condition, the magnitude of the cued-recall CES was lower than free-recall CES, indicating that the scoring procedure was perhaps more conservative, or that the cuing procedure in some way limited error magnitude. In addition, cued recall CES was greater for the 3-week consistent conditions than was free recall CES. This difference may be accounted for by the change in scoring procedures or by the possibility that in this experiment the effect of consistent subsequent information was to exaggerate the harmony or conflict in the passage.

Confidence ratings. Spiro (1975) examined confidence ratings for sentences pertaining only to the issue of having children; these sentences were presumed to be best examples for reflecting the subsequent bias. Only ratings for absolute error values of four or five were contrasted to ratings for error values of one or less. Spiro reasoned that if errors were a result of conscious fabrication, those that were greater in magnitude should be more easily detected. In the cued recall experiment, no responses to Item 8 (concerning the issue of children) were given error scores large enough to adopt Spiro's criterion. Indeed, it is questionable if responses to this item are comparable to the types of freely recalled sentences that were judged to concern the issue of having children; the cuing procedure is quite likely more restrictive.

Due to these differences in scoring and procedure, the locus of schema change was tested by comparing ratings for correct versus incorrect responses. Such a comparison should be a very conservative test of Spiro's locus hypothesis, in that higher ratings for nonextreme errors should reduce the difference between correct and incorrect response ratings. All five responses were employed, since all five items were designed to reflect the incorporation of the subsequent information, and since they formed the basis of all prior analyses.

As can be seen in Table 2, there was a slight tendency in all conditions for more subjects to rate correct responses higher than incorrect responses. (For Spiro's, 1975, contradictory 3-week condition, the number of subjects showing greater confidence in incorrect responses was five times greater than those showing less or equal confidence.)

Table 2
Number of Subjects Showing Two Patterns of Confidence Ratings of Correct and Incorrect Responses

<table>
<thead>
<tr>
<th>Direction of difference in M ratings</th>
<th>2 days</th>
<th>3 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Consistent</td>
<td>Contradictory</td>
</tr>
<tr>
<td>Incorrect ≥ correct</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Incorrect &lt; correct</td>
<td>11</td>
<td>7</td>
</tr>
</tbody>
</table>
Furthermore, for all 3-week subjects, confidence in correct sentences was reliably greater than confidence in incorrect sentences; paired-comparison $t(25) = 3.230, SD = .409, p < .01$. Since Spiro emphasized the rating comparison for subjects hearing contradictory information, a separate test was performed for these subjects; once again, the results indicated that confidence for correct was greater than for incorrect recall, $t(13) = 2.923, SD = .491, p < .02$. Therefore, subjects tested after 3 weeks could reliably distinguish between their errors and correct recall, when they were asked to do so immediately after retrieving specific information. However, these results do not demonstrate that all reconstructive effects occurred during retrieval; the mean difference in ratings between correct and incorrect was slightly greater than one. Nor do the results implicate a conscious guessing strategy; lower ratings could logically reflect the difficulty of retrieving and integrating poorly formulated aspects of story memory. Nevertheless, the existence of a rating difference under the conditions of this experiment makes it difficult to argue that all constructive effects occurred during the retention interval, and not during the conscious processing of story-related information. The most reasonable conclusion concerning these results, as well as the results of other experiments demonstrating memory changes, is that changes occur due to processes operating at all stages of information processing, and are guided by the task demands.

Individual Differences

The second purpose of this experiment was to examine the relationship between constructive memory performance and scores on tests of cognitive flexibility. A multiple regression analysis was performed, using group membership variables (type of subsequent information, time of testing, and their interaction) and scores on the flexibility tests as predictors, and the absolute value of the largest error (AES) as the outcome measure. AES is the appropriate measure for the individual difference analyses because direction of error is no longer relevant.

The most important finding was that the addition of the individual difference variables reliably increased the predictive ability of the regression equation, $F(3, 63) = 3.195, SE = .012, p < .03$; knowledge about differences in cognitive flexibility provided additional information about constructive error performance apart from knowing the parameters of the experimental situation.

The predictors with weights reliably different from zero were type of subsequent information, $F(1, 63) = 5.305, SE = .162, p < .025$; the Object Naming test, $F(1, 63) = 4.438, SE = .048, p < .05$; and the Match test, $F(1, 63) = 5.176, SE = .039, p < .05$. These three variables provided independent and reliable sources of information about the variance of AES performance. Since the magnitude of the standardized beta weights was approximately equal for the reliable predictors, interpreting the regression equation is straightforward: Occurrence of contradictory subsequent information and high scores on the adaptive flexibility test led to a greater degree of constructive error, and spontaneous flexibility was positively related to recall accuracy.

The relationships of the flexibility tests to constructive memory performance can be loosely interpreted as follows: Spontaneous flexibility appears to indicate individual abilities in maintaining separate memory stores for the experimental materials; after the influence of context was determined, subjects with higher scores more accurately recalled the story. Thus, spontaneously flexible individuals may tend to engage in less automatic restructuring of their experience. In addition, adaptive flexibility may be indicative of the tendency to combine separately stored information at retrieval, as a function of the amount of integration required by the task. Presumably, task demands are a matter of individual perception; for those who perceived the cues as functioning for the related information, adaptive flexibility predicted their tendency to incorporate it. More generally, support has been provided for the assumption that the two types of flexibility tests tap different characteristics of cognitive behavior.

Another possible view of the relationship of memory performance to flexibility scores
is to attribute this relationship to a third variable, general intelligence. There are three objections to this approach. First, several sources have identified measures of flexibility as tapping individual characteristics that are different from those characteristics measured by intelligence tests (Anastasi & Schaefer, 1971; Lindeman & Fullagar, 1975; Yamamoto, 1965). The reported correlations are generally quite low (approximately .20), especially for individuals with above average scores on intelligence tests. Second, if the relationship reported in this study could be partially accounted for by a general intelligence factor, even less information about individual differences in recall accuracy would be obtained. This argument is based on the issue of intelligence testing and what it tells us about specific abilities. Guilford's factor-analytic approach to intelligence, as well as the individual-difference approach of cognitive psychology, emphasizes a detailed analysis of individual characteristics and a convergence of these characteristics with behaviors in certain experimental settings. Finally, Meehl (1970) has called attention to the fallacy of using a third factor as a covariate for investigating relationships between a naturally occurring characteristic and some outcome variable. In the present study, removing the variance accounted for by a general intelligence factor would automatically allow some unknown fourth factor to systematically affect the relationship between flexibility and recall. In summary, the relationship between cognitive flexibility and recall merits attention in its own right, apart from investigating the more general relationship between intelligence and memorial performance.

Conclusions

The cued-recall results of this experiment are consistent with Spiro's (1975, 1977) findings concerning the influence of subsequent information on memory for an event. As the testing interval increased, the magnitude of constructive errors increased, and their direction reflected the nature of the subsequent information. However, after 3 weeks, subjects indicated more confidence in correctly recalled sentences than in sentences containing errors. This difference in recall confidence probably indicates that either some constructive effects occurred during retrieval or subjects engaged a conscious guessing strategy. At the very least, both types of recall were not produced by directly accessing a stable, well-integrated schema for all information related to the story.

Second, the relationships between cognitive flexibility and constructive errors may illuminate the issue concerning the locus of constructive effects. Both the terms used to distinguish the two types of flexibility tests and the nature of the tasks employed suggest that spontaneous flexibility may indicate a reduced tendency to restructure experience, whereas adaptive flexibility may correspond to characteristics of integrative recall. Thus, the tests of cognitive flexibility may be relatively independent, and furthermore, they may reflect relatively independent processing with regard to storage and retrieval of real world information.

Finally, this research has implications for educational theory and practice. Spiro (1977) has suggested that an emphasis on knowledge updating (incorporating new information into previous structures) rather than isolation for the purpose of test taking (as in typical memory experiments) could provide for richer, more particularized interpretations of information. In a general sense, such an emphasis may encourage inaccurate memory of specific information but increase the tendency to approach learning and testing situations with a more flexible cognitive set. The results from this experiment further suggest that knowledge structures are modified both during storage and when information from the structures is retrieved. In addition, individuals probably differ according to the conditions under which they modify their schemata for events, and one way of specifying this difference is in terms of cognitive flexibility.

References

Anderson, R. C. The notion of schemata and the edu.


Received October 23, 1978
Revision received March 1, 1979