A study investigated whether direct instruction in story constituents and their interrelationships could enhance children's organization in story writing. It was hypothesized that the special instruction could provide children with an internalized prototype story structure that could serve as a heuristic or planning device for organizing compositions, and that the instruction might indirectly affect other features of writing, such as creativity, cohesion/coherence, syntactic complexity, and length of protocols. Subjects, 19 fourth grade students identified (through story telling and scrambled story recall tasks) as lacking a keen sense of narrative structure, were randomly assigned to one of two treatment groups. The first group received instruction in knowledge of story structure, while the second received instruction in dictionary word study. Results showed that direct instruction in narrative structure did have a strong positive effect on organization in children's story writing. In addition, the instruction enhanced overall creativity, uniqueness, and language usage, and reduced cohesive errors. There were no meaningful effects on syntax or length of protocols.

(Author/FL)
Effects of Instruction in Narrative Structure on Children's Writing
Jill Fitzgerald, Alan B. Teasley
The University of North Carolina at Chapel Hill

NARRATIVE STRUCTURE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Jill Fitzgerald
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."
Abstract

The main goal of the present study was to investigate the possibility that direct instruction in story constituents and their interrelationships could enhance children's organization in story writing. It was hypothesized that the special instruction could provide children with an internalized prototype story structure which could serve as an heuristic or planning device for organizing compositions. The study also hypothesized that the special instruction might indirectly affect four other features of writing: creativity, cohesion/coherence, syntactic complexity, and length of protocols.

Nineteen fourth grade readers identified (through story telling and scrambled story recall tasks) as lacking a keen sense of narrative structure were randomly assigned to one of two treatments, instruction in knowledge of story structure or instruction in dictionary-word study. Instruction in narrative structure did have a strong positive effect on organization in story writing. Importantly, overall creativity, and specifically, uniqueness and language usage, were enhanced. There was only one significant effect on cohesion/coherence, i.e., cohesive errors were reduced; and there were no meaningful effects on syntax or length of protocols.
Effects of Instruction in Narrative Structure on Children's Writing

Introduction

The main purpose of the present study was to investigate the possibility that direct instruction in narrative structure would enhance children's organization in story writing. It was hypothesized that the special instruction would lead to mental internalization of a prototype story structure which could serve as an heuristic or planning device for organizing compositions. The study also explored the possibility that instruction in narrative structure might indirectly affect four other features of writing: creativity, cohesion/coherence, syntactic complexity, and length of protocols.

Narrative Structure

Knowledge of narrative structure or organization, sometimes referred to as story schema, involves identification of narrative elements and their interrelationships (cf., Mandler & Johnson, 1977; Rumelhart, 1978; Stein & Glenn, 1979; Thorndyke, 1977). The present study focussed on one way to represent narrative structure or organization, that is, by a story grammar. An example of a story grammar is Mandler and Johnson's representation (Johnson & Mandler, 1980; Mandler & Johnson, 1977) which describes six major categories of narrative information: Setting, Beginning (a precipitating event), Reaction (the-protagonist's Reaction and setting a Goal), Attempt (the effort to achieve the Goal), Outcome (the success or failure of the Attempt, and Ending (the long-range consequence of the action sequence or the added emphasis). The Beginning through the Ending make up an Episode. Rules in the grammar specify temporal relationships between categories and delineate how complex stories can
occur through options such as embedding of Episodic.

The main hypothesis of the present study was that instruction in narrative structure would provide the students with an internalized schema or prototypical narrative structure which would enable them to write better organized stories. The notion that effective writing is organized has its origins in classical rhetoric, specifically Aristotle's discussion of disposition, now more commonly called arrangement or form (cf., Corbett, 1971; Winterowd, 1975). Contemporary rhetoricians and discourse theorists acknowledge that particular modes or genres of discourse have certain prescribed patterns (Lindemann, 1982), paradigms (D'Anbgeio, 1980), or macrostructures (Kintsch & Van Dijk, 1978), which can aid a writer in the process of shaping discourse (Halliday, 1973; Lindemann, 1982; Winograd, 1979).

Several studies of text comprehension clearly document the importance of knowledge of narrative structure in understanding and remembering stories (cf., Kintsch, 1977; Mandler & Johnson, 1977; Rumelhart, 1978; Stein, 1976; Stein, 1979; Stein & Glenn, 1977). There are also a few studies involving the effects of instruction in narrative structure on reading comprehension (Dreher & Singer, 1980; Fitzgerald & Spiegel, 1983; Singer & Donlan, 1982) with conflicting results. However, studies of writers' knowledge or use of text structure or of effects of instruction in narrative structure on writing appear to be rare. The results of the few intervention studies conducted are again mixed (cf., Edmonson, 1983; Gordon & Braun, 1983).
Indirect Effects of Instruction in Narrative Structure

In addition to the main hypothesis, the present study investigated the possibility that instruction in narrative structure might have an effect on other aspects of children's writing, specifically on creativity, cohesion/coherence, syntactic complexity, and length of protocols. Consideration of potential instructional effects led to two general speculations, each of which suggested a way in which two or more characteristics of writing might be affected, and three additional specific predictions, one each for creativity, cohesion/coherence, and syntax.

The first general speculation related to creativity and length of protocols. If indeed the children in the experimental group did internalize a prototypical story structure, that structure might serve as an heuristic or aid (cf., Bereiter, 1980; Kintsch & van Dijk, 1978) to what classical rhetoricians called invention, which means discovering content for a piece of discourse (cf., Corbett, 1971). The structure could suggest "slots" to be filled as the children composed, thereby resulting in stories with richer, more elaborated content and thus more creative and longer protocols.

The second general speculation related to creativity, cohesion/coherence, syntactic complexity, and length of protocols. If an internalized schema for narrative structure served as a planning device, then story organization could become "automatic," so the writer's attention during composing could be freed and redirected towards other aspects of text production. Current models of the composing process (Flower & Hayes, 1980, 1981; Matsuhashi, 1981; Perl, 1980; Sommers, 1979) suggest that a writer must attend to numerous facets during composing (e.g., the text written so far, the content to be expressed next, word choice, and syntactic patterns). In the words of Flower and Hayes (1980), the writer is constantly "making plans and juggling constraints," and one of the
"constraints" of story-making is narrative structure. If narrative structure became "automatic," other aspects of writing might be affected, such as creativity, cohesion and coherence, syntactic complexity, and length of protocols.

Three specific indirect effects of story structure instruction on creativity, cohesion/coherence, and syntax, respectively, were predicted. First, creativity might be either enhanced or inhibited. On the one hand, children's acquisition of a basic narrative structure could be viewed as a necessary prerequisite to creativity. It is possible that acquisition of the basics or fundamentals of a given medium facilitates or frees creativity. On the other hand, the internalized schema could have a detrimental effect, producing stories that adhere slavishly to a formula.

Second, cohesion/coherence might be improved. Knowledge of linking relationships in narrative structure, i.e., causal and temporal relationships between story parts, could result in increased use of cohesive ties in the text productions (cf., Bracewell, Hidi & Hildyard, 1981; Clancy, Jacobson, & Silva, 1976; Rentel & King, 1983). Also, some individuals define coherence as an aspect of structure (cf., Mosenthal & Tierney, 1984; van Dijk, 1980). Therefore, improving structure or organization might also improve overall coherence.

Some studies have documented developmental effects for cohesion/coherence in children's writing, revealing a movement from predominant use of unsuccessful or incorrect connections to exophoric connections (referents outside the text) to an increase in use of endophoric connections (referents inside the text) (Rentel & King, 1983). Similarly, more mature or "better" writing appears to be marked by fewer remote connections and more local, immediate ones (McCutchen & Perfetti, 1982; Witte & Faigley, 1981). Also, more mature or "better" writing
appears to be characterized by higher proportions of cohesive ties per T-unit, and higher proportions of reference, lexical and sometimes, conjunctive ties (McCutchen & Perfetti, 1982; Witte & Faigley, 1981). Based on findings from previous studies making developmental comparisons and/or investigating features of "better" or more mature writing (McCutchen & Perfetti, 1982; Rentel & King, and Witte & Faigley, 1981), it seemed reasonable to hypothesize that, in the present study, if the narrative structure instruction group had an indirect effect on cohesion/coherence, it might result in writing samples with a higher proportion of total cohesive ties per T-unit; greater proportions of reference, lexical, and perhaps conjunctive ties, and higher proportions of immediate and mediated ties and/or a reduced proportion of remote ties.

Third, protocols might be marked by increased syntactic complexity, with certain aspects of syntax being more affected than others. Since story structure knowledge includes knowledge of causal and temporal relationships between story parts, story structure instruction could result in a special increase in subordination where relationships between ideas are more clearly shown. Even more particularly, the increased subordination might be in adverb clauses, the subordinate structures most likely to express relationships of time, cause or reason, purpose or result.

The Present Study

The present study then focused on the effects of instruction in narrative structure on the organization of fourth graders' story writing. It additionally explored indirect effects on creativity, cohesion/coherence, syntactic complexity, and length of stories. Nineteen children who had little or no knowledge of story structure, and whose story writing was not well organized, were randomly assigned to one of two treatment groups. Nine children received instruction in narrative structure, while 10 others received instruction as a
control group. Children were pre-tested by writing two stories before instruction; they were tested again after a first phase of instruction and after the final phase of instruction by writing two stories at each test time. Two phases were set up to see 1) what changes, if any, could be effected with a relatively short period of instruction, as opposed to a longer one, and 2) if any initial effects could be maintained and/or enhanced through additional instruction.

**Method**

**Design**

The design of the study was a two group repeated measures design. First, 50 fourth graders in one school who returned parental permission slips were screened on knowledge of story structure. There were two screening tasks with two stories in each task. These two tasks, oral story telling and scrambled story recall, had been used in prior research on story structure knowledge (cf., Applebee, 1978; Botvin & Sutton-Smith, 1977; Buss, et al., in press; Kintsch, Mandel, & Kozminsky, 1977; Stein & Glenn, in press).

Twenty children with the lowest scores on the two tasks were selected for instruction. One student withdrew from the study after one week, and post-screening data reported here are based on 19 subjects.

Before instruction, each of the 19 children wrote two stories. These protocols were later used to analyze possible pre-treatment group differences on all writing variables in the study.

The 19 children were then randomly assigned to one of two treatment groups, special instruction in narrative structure or a control group which received instruction in dictionary usage and to a limited extent, word study. There were two phases of instruction. The first was a short-term intensive phase, which consisted of six 30- to 45-minute sessions during a two week period. The second
was long-term intermittent instruction with distributed practice consisting of ten 30- to 45-minute sessions over a five week period.

At the end of the first phase and again at the end of the second phase, the children wrote two stories.

Thus each subject had four screening scores, two pre-treatment writing samples, two interim writing samples, and two final writing samples. At each testing time, measures were counterbalanced within, and where relevant, across tasks.

Subjects

The subjects for instruction were 19 fourth grade students (7 males and 12 females) in one school who were identified through oral response screening measures as having little or no knowledge of story structure.

Task and Variable Descriptions and Scoring Procedures

Screening tasks and scoring.

To assess knowledge of story structure during screening, two tasks used in prior studies were used: a story telling task (cf., Applebee, 1978; Botvin & Sutton-Smith, 1977; Glenn & Stein, in press; Leondar, 1977; Stein & Glenn, 1977; Sutton-Smith, et al., 1976; Trabasso, et al., in press) and a scrambled story recall task (cf., Buss, et al., in press; Kintsch, Mandel, & Kozminsky, 1977; Mandler, 1978; Stein & Glenn, in press). These two tasks were used because results of previous research indicated that they were successful in revealing developmental differences; so it seemed likely they would also be sensitive to individual differences within a grade level.

For each task, two stories were developed. Students were tested individually, and oral responses were tape recorded and later transcribed.

For the story-telling task, subjects were instructed to read a story setting (such as "Once there was a candy store owner named Charles") and finish
the story, making it into the best story they could. They were also told to ask the examiner for help on any words they didn't know when reading. A standard set of prompt statements, such as "I know it's hard, but I'd like you to try your best" was used. Approximately 15 to 20 minutes were allotted for each story. All children finished within the time limit.

Individuals' knowledge of story structure was revealed by the complexity of their story telling. Responses were scored using a modification of a system developed by Stein and Glenn (1977). The following five degrees of complexity were scored 1 to 5 points, respectively: 1) descriptive sequence (no clear Goal; causal connections between statements rare); 2) reactive sequence (no clear Goal; causal connections between several statements); 3) abbreviated Episode (Goal stated explicitly or easily inferred; certain categories, but not all categories, stated (e.g., a Beginning, implied Goal, Outcome sequence would be an abbreviated Episode); 4) complete Episode (Goal stated explicitly or easily inferred; all categories other than Goal and Ending explicitly stated); and 5) complex story (Goal stated explicitly or easily inferred; multiple Episodes present). Overall interrater reliability for scoring the two stories was .95.

For the second knowledge of story structure screening task, subjects were instructed to read and orally recall scrambled stories written at the second grade level according to the Fry (1977) readability formula. Instructions stressed that the stories the children read were all mixed up and that their task was to try to remember everything they could and put the pieces back in order. The children were also told to ask the examiner for help on any unknown words while reading and to read the story straight through once. Again, a standard set of prompt statements was used. After retelling the story 24 hours later, the children were asked if they had talked to anyone about the story.
None indicated that they had done so. Overall interrater reliability for parsing the two stories was .97.

One of the scrambled stories was a "then-connected" two-Episode story which was interleaved this way: Beginning from Episode one, then Beginning from Episode two, respectively; Attempt from Episode one, then two, and so on (cf., Mandler, 1978). The second scrambled story at each testing time was a two-Episode story in which there was an embedded Outcome Episode. The embedded Outcome story was scrambled using a random order. The two types of scrambling, one interleaved and one random, were used in order to represent a wider array of stimuli types.

Individuals' knowledge of story structure was revealed by the degree to which subjects reordered the stories. Kendall's Tau rank order correlations were computed for each subject (cf., Buss et al., in press; Stein & Glenn, in press; Stein & Nezorski, 1978). The Kendall's Tau scores indicated the degree of concordance between the order of categories recalled and the original order of the categories in the canonical version of the story. Overall interrater reliability between the two investigators for scoring the two story retellings was .85.

Writing task.

Two writing samples were acquired before treatment, at the end of the first phase of instruction, and again after the final phase, for a total of six writing samples for each child. For the writing task the children (in groups of nine or ten) were given a written story setting such as: "There was a boy named Sam. Sam was in the fourth grade and lived in a small brick house." The examiner read the setting out loud and instructed the children to: "...write a story for me. Make it a story you could tell someone else if you wanted to. Make it the best story you can." At each testing time one story setting
involved people and one involved animals. A standard set of prompt statements was used (e.g., "I know it's hard, but I'd like you to try your best"). Care was taken to avoid using words like "start," "finish," and "ending." The children were told not to be concerned about spelling. If they asked for help in spelling, standard prompt statements such as "I'll help you with that later" were used. The children were not allowed to talk to each other during the writing. Twenty minutes were allowed for each story. All children finished within that time period.

Writing: Story organization.

The modification of Stein and Glenn's system for revealing structural complexity was used. The system is described above under "Screening tasks and scoring." Each child received one score for each of the six writing samples. Overall interrater reliability was .79.

A supplementary category adjacency analysis was also carried out to investigate story organization. Individuals' knowledge of temporal relationships between two specific categories was revealed by a conditional analysis in which each child's protocols were examined to see if the second category predicted by the grammar was immediately produced given that the first member of the pair was present (e.g., given that an Attempt was present, did an Outcome follow?). Two judges separately determined the presence/absence of category adjacencies for each protocol. Agreement on this proved difficult, and disagreements were resolved through discussion (cf., Mandler & Johnson, 1977, for similar procedures).

Writing: Creativity.

Moslemi's (1975) four part scale for creativity was used. Five creativity scores were obtained: uniqueness, idea production, language usage, originality, and the average across the four subscales. Uniqueness was defined as: "a
reflection of the student writer's unique individuality, ... particular preferences, tastes or beliefs, ... particular use of humor or wit. A unique blend of emotions, moods and personal philosophy." Idea production was "quality, quantity, fluency or diversity of ideas or precise detailed elaboration or description of one person, experience object or idea." Language usage was: "the use of imagery, lively description and figures of speech (metaphor, simile, personification, etc.) and the coining of new words. Fresh or colorful word combinations. The use of vivid terminology in appealing to the senses." Originality was: "the use of new, imaginative or unusual ideas or a common idea used in a new and imaginative fashion. A breaking away from the original stimulus in the production of an uncommon response." For each subscale, holistic scores could range from one (low) to five (high). Each child then received five creativity scores for each of the six writing samples. Following Moslemi's (1975) procedures for securing reliability estimates, using four judges, Cronbach's coefficient alpha for reliability for the average of judges' ratings for uniqueness was .96. For idea production it was .98, and for language-usage, originality, and the average across the four subscales it was .94, .96, and .98, respectively.

Writing: Cohesion/coherence.

Halliday and Hasan's (1976) system for scoring cohesion/coherence was the main resource for scoring cohesion/coherence. Briefly, they define cohesion as the linguistic features which make a sequence of sentences a text, i.e., give it "texture." Halliday and Hasan look at only those cohesive ties which operate between or across sentences. They classify these ties in five main categories: reference, substitution, ellipsis, conjunction, and lexical ties. To qualify as a cohesive tie, a word phrase must pre-suppose the existence of another element elsewhere (in another sentence) in the text. Halliday and Hasan also provide a
way for quantifying the distance between a cohesive tie and its presupposed element. The distance may be immediate (that is, the tie and its referent are in contiguous sentences), mediated (the tie and its referent are more than one sentence apart, but the intervening sentences also contain a tie that refers to the element), or remote (the tie and referent are more than one sentence apart, but the intervening sentence contains no mediating ties.)

Thirteen cohesion variables were scored: reference, substitution, ellipsis, conjunctive, and lexical ties; total number of ties; cohesive errors (cases of overt ambiguity, e.g., references to "this" or "that knife" when a knife had not been previously mentioned, or use of a pronoun with an unclear referent); proportion of all ties that were immediate, mediated, and remote; and average mediated distance, average remote distance, and average overall distance. To control for length of protocols, for the first seven variables the number of that type of tie was divided by the number words in the protocol and multiplied times 100 to obtain the number of the type of tie per 100 words (Cooper, 1983). Substitution and ellipsis ties were dropped from the study because they were so rarely given in the protocols. Each child then received 11 cohesion scores for each of the six writing samples. Overall interrater reliability for types of ties, for agreement for scores within plus or minus two was .92. Overall interrater reliability for cohesive errors for agreement for scores within plus or minus two was .90. Overall interrater reliability for types of distance for agreement for scores within plus or minus two was .59. Results for distances, with reliability .59, should be interpreted cautiously.

Writing: Syntax.

Thirteen syntax variables were scored: noun clauses, adjective clauses, adverb clauses, total clauses, participle phrases, gerund phrases, infinitive phrases, total verbal phrases, prepositional phrases, two measures of syntactic
density, proportion of T-units with subordination, and mean T-unit length. To control for length of protocols, all variables were created by dividing the number of occurrences of the syntax type by the number of T-units (cf., Witte & Faigley, 1981). Adjective clauses, participle phrases, and gerund protocols were dropped from the study because they were rarely given in the protocols. Each type of clause and phrase was defined according to Warriner (1983). One syntactic density measure was the number of total clauses plus the number of verbal phrases, all divided by the number of T-units. The second syntactic density measure was the same as the first except that the number of prepositional phrases was also added in the numerator. Proportion of T-units with subordination was defined as number of T-units containing either a subordinate clause or a verbal phrase divided by total T-units in the story. Mean T-unit length was defined according to Hunt (1965), i.e., it was the number of words divided by the number of T-units. Each child then received 10 syntax scores for each of the six writing samples. Overall interrater reliability for syntax was .77.

Writing: Length.

Length was measured in two ways, number of words and T-units. Number of words was obtained using a simple count. Each child received one score for each of the six writing samples. Overall interrater reliability for agreement for scores within plus or minus one was .93. T-units were defined according to Hunt (1965), i.e., a T-unit was a main clause plus all the subordinate clauses attached to it. Each child received one score for each of the six writing samples. Overall interrater reliability was .80.

Instruction

One of the investigators, a trained colleague, and a trained graduate student taught both treatment groups. Each group met with each of the three
instructors an approximately equal number of times.

So that potentially intervening variables could be controlled, the following factors were equated for the two groups: (1) instructional time, (2) amount of individual versus group work, (3) degree to which the lessons were structured, (4) broad content area of instruction (language arts), and (5) amount of practice reading and writing stories. In addition, the same stories were presented to both groups.

The instruction for the story structure group centered around two purposes. The primary purpose was to help the children to form a sort of "cognitive blueprint" or structure for stories and to elaborate the "blueprint" (cf., Rubin, 1980; Stein & Trabasso, in press). The secondary purpose was to increase the children's awareness of the ways in which knowing the structure of stories could help them understand and write stories (cf., Brown, Campione, & Day, 1980; Collins & Smith, in press).

As noted, there were two phases of instruction. During phase one of instruction, each lesson focused on one story constituent and its temporal relationship to other story parts. In a typical phase one lesson, the instructor first gave a review of the story parts learned in previous lessons and an overview of the new lesson. Next, the instructor told about the story element (e.g., Attempt) by describing it, pointing out the element in a story on a wall chart, and giving two or three other examples of the element that would be appropriate for the story on the chart. Then, the instructor elicited two or three oral examples of the story element from the children. Next, the instructor gave non-examples and asked why these were not good examples of the element being studied. The non-examples might be different story parts (e.g., an Outcome for an Attempt), or they might be the right story part but misplaced within the story. Last, the students participated in one or more group or
individual activities designed to reinforce understanding of the element being taught that day. Examples of reinforcement activities are prediction and macro-cloze tasks. For prediction tasks, the students read part of a story (e.g., Setting and Beginning), told what part should come next, and provided an appropriate example of that part. In the macro-cloze activities, the children filled in missing chunks of stories after reading or listening to an entire story. For example, a story might contain a Setting, Beginning, Reaction, Outcome, and Ending. After reading the story, the student would identify what part was missing and supply an appropriate Attempt.

Phase two instruction consisted of individual and group activities designed to provide continued reinforcement of knowledge of story constituents and to make the children aware of the relationship between knowledge of specific story parts and their temporal relationships and story production. Examples of additional activities used to reinforce knowledge of story constituents during phase two are scrambled story tasks, sorting tasks, and retelling activities. The scrambled story tasks required students to reorder stories in which the story had been jumbled. For the sorting tasks, the students sorted sentences and phrases from the story into piles to show which pieces "went together." They then ordered the piles to make a well-sequenced story. In the retelling activities, the children recalled stories they had just read or heard and discussed ways in which the stories deviated from well-formed stories, such as through omission of necessary elements or addition of extra elements. In each of the reinforcement activities, there were at least two or three variations for individual and small group work. The instructor stressed that knowing about the story parts could lead to better understanding of and memory for the story and to better story writing.

Instruction for the control group centered on three main dictionary skills:
using phonemic respellings to determine the pronunciation of a word, using a
dictionary to identify the correct spelling of an inflected word, and choosing a
meaning for a word from the choice listed in the dictionary. A typical lesson
plan included the following elements: (1) an overview and review were provided;
(2) the instructor introduced the new skill for the day; (3) the instructor led
the children through the use of the skill in a step by step manner; and (4) the
children participated in follow-up practice exercises involving worksheets,
games, and/or team competition. The stories used for instruction with the story
structure group were read by the children in the control group in many of the
follow-up exercises. Similarly, practice in writing stories was included in
many of the follow-up exercises through activities such as writing stories
incorporating the words introduced earlier in the lesson.

Analyses

The major analyses were multivariate and univariate repeated measures
analyses of variance. Procedures outlined in Bock (1975) and Finn and Mattsson
(1978) were followed. For all repeated measures analyses, preliminary tests of
compound symmetry were done to determine whether the repeated measures analyses
were performed as mixed models or not. Depending on the number of dependent
variables, a univariate or multivariate mixed model repeated measures analysis
was used when the condition of compound symmetry was met. A multivariate
repeated measures model was used when the condition of compound symmetry was not
met.

Pre-treatment group differences on all writing variables were assessed
prior to testing interim and post-treatment effects. Where there were
pre-treatment differences, the two pre-treatment scores were used as covariates
in the pertinent repeated measures analysis.

When covariates were included in the models, the appropriate tests of
parallelism of regression planes were performed. There were only three instances of significant parallelism tests. These are discussed in the "Results" section on length and on syntax.

Twenty-nine analyses were performed for the 29 writing variables in the study.

Results

Story Organization

Story organization scores.

To assess the instructional effects on organization of written narratives, a mixed model repeated measures analysis was done in which the between-factor was treatment (experimental and dictionary-word study groups). The within factors were time (interim and final) and story within time (two stories at each time).

There was an overall significant difference between treatment groups (univariate $F(1,17) = 11.25, p < .01$) favoring the story structure group. Means for each interim and final measure are shown in Table 1. The scores indicate that, after instruction, on the average, the story structure group tended to write stories that were organized most like abbreviated episodes, while the control group stories still tended to be reactive sequences.

There were no other significant effects.

Insert Table 1 about here

Adjacency analysis for production task.

To determine whether there were treatment group differences in awareness of temporal relationships between any two adjacent story categories, adjacency figures were calculated for the proportion of occasions that the second category
predicted by the grammar was immediately produced, given that the first member of the pair was provided. The story structure group demonstrated greater awareness of temporal relationships between all category pairs. The proportions for the adjacency figures for the interim and final stories taken together for the story structure group were Setting-Beginning, .65; Beginning-explicitly stated Goal, .46; explicitly stated Goal-Attempt, .87; Attempt-Outcome, .00; and Outcome-Ending, .86. The corresponding proportions for the dictionary-word study group were .25, .22, .50, .78, and .78, respectively. There were no apparent differences in patterns over time.

Creativity

There were important pre-treatment differences only for the idea production subscale. These scores needed to be used as covariates, and it was not sensible to adjust all of the subscales for pre-treatment differences on one subscale. Consequently, for interim and post-treatment analyses, rather than treating the four subscales as a multivariate set in one run, five separate univariate runs were done for the four subscales and for the mean creativity score. The models for uniqueness and language usage were mixed model repeated measures analyses in which the between factor was treatment (experimental and dictionary-word study group), and the within factors were time (interim and final) and story within time (two stories at each time). The models for mean creativity and idea production were mixed model repeated measures analyses of covariance in which the between factor was treatment (experimental and dictionary-word study group), the within factors were time (interim and final) and story within time (two stories at each time), and the covariates were the two pre-treatment scores on the respective variables. The model for originality was a multivariate repeated measures analysis in which the between factor was treatment (experimental and dictionary-word study), and the within factors were time (interim and final) and story within time (two at each time).
There were significant effects for mean creativity, uniqueness, and language usage. First, for mean creativity, there were three significant effects. One was for treatment (univariate F [1, 15] = 5.70, p < .04), with the story structure group being more creative (overall mean, 2.90) than the other group (overall mean, 2.29). (See Table 2 for creativity means and standard deviations.) The second significant effect was for time (univariate F [1, 15] = 7.25, p < .02). The time effect can only be interpreted in light of the third significant effect which was for a disordinal interaction for treatment by time (univariate F [1, 15] = 5.21, p < .04). The story structure group's mean creativity scores remained about the same over time. (The average across the two final stories was 2.23 versus 2.35 for the two interim stories.) There were no other significant effects.

Next, for uniqueness, there was a significant effect for treatment (univariate F [1, 17] = 7.01, p < .02), with the story structure group outscoring the other group. (The overall average for the story structure group was 2.68. For the other group it was 2.03). There were no other significant effects.

Finally, for language usage, there was a significant effect for treatment (univariate F [1, 17] = 6.49, p < .03), with the story structure group producing more creative language usage. (The overall average for the story structure group was 2.15. For the dictionary-word study group, it was 1.67). There were no other significant effects.

**Cohesion/coherence**

To assess the instructional effects on cohesion/coherence, 11 separate repeated measures models were done, one for each of the 11 cohesion dependent
variables. For reference ties, conjunctive ties, total number of ties, and proportion of immediate ties, the models were mixed model repeated measures analyses in which the between factor was treatment (experimental and dictionary-word study group), and the within factors were time (interim and final) and story within time (two stories at each testing time). For proportion of mediated ties, proportion of remote ties, mean mediated distance, and mean remote distance, the models were mixed model repeated measures analyses of covariance in which the between factor was treatment (experimental and dictionary-word study group), the within factors were time (interim and final) and story within time (two stories at each testing time), and the covariates were scores for the two pre-treatment writing samples. For lexical ties and cohesive errors, the models were multivariate repeated measures in which the between factor was treatment (experimental and dictionary-word study group), and the within factors were time (interim and final) and story within time (two stories at each testing time). Finally, for mean overall distance, the model was a multivariate repeated measures analysis of covariance in which the between factor was treatment (experimental and dictionary-word study group), the within factors were time (interim and final) and story within time (two stories at each testing time), and the covariates were scores for the two pre-treatment writing samples.

Only two models had significant effects. For the model for cohesive errors, there were significant effects for treatment (univariate $F [1.17] = 6.51, p < .03$), with the special instruction in narrative structure resulting in, on the average, fewer cases of cohesive errors (group mean of 1.45 per 100 words per story for interim and final stories) than were produced in the other group (group mean of 2.66 per 100 words per story for interim and final stories). There was also a significant time effect (univariate $F [1.17] = 6.31, p < .03$, with both groups decreasing the number of cohesive errors over
time. As can be gleaned from Table 3, the narrative structure instruction group decreased cohesive errors from an average of 1.72 across the two interim stories to an average of 1.19 across the two final stories. The comparable figures for the dictionary-word study group were 3.54 and 1.79, respectively. There were no other significant effects in the cohesive error model.

For the model for proportion of ties that were remote, after controlling for pre-treatment differences, there was a significant effect for time (univariate $F [1,15] = 7.21, p < .02$), with both groups decreasing the proportion of ties that were remote. As can be gleaned from Table 3, the narrative structure instruction group decreased the proportion of remote ties from an average of .37 across the two interim stories to an average of .35 across the two final stories. The comparable figures for the dictionary-word study group were .39 and .30, respectively. There were no other significant effects in the model.

Syntax

To assess the instructional effects on syntax, 10 separate repeated measures models were done, one for each of the 10 syntax dependent variables. For noun clauses, infinitive phrases, total verb phrases, prepositional phrases, the second syntactic density measure, and proportion of T-units with subordination, the models were mixed model repeated measures analyses in which the between factor was treatment (experimental and dictionary-word study group), and the within factors were time (interim and final) and story within time (two stories at each testing time). For adverb clauses, total clauses, and the first syntactic density measure, the models were mixed model repeated measures analyses of covariance in which the between factor was treatment (experimental
and dictionary-word study group), the within factors were time (interim and final) and story within time (two stories at each testing time), and the covariates were scores for the two pre-treatment writing samples. Finally, for mean T-unit length, the model was a multivariate repeated measures analysis of variance in which the between factor was treatment (experimental and dictionary-word study group), and the within factors were time (interim and final) and story within time (two stories at each testing time).

Only two models had significant results, but none of the results seemed meaningful. First, there was a significant effect for time in the adverb clauses model (univariate F [1,15] = 4.85, p < .05). After controlling for pre-treatment scores on adverb clauses, both groups gave significantly more adverb clauses in the final stories than they did in the interim stories. (As can be gleaned from Table 2, the final total across both stories for the experimental group was .25 versus .11 for interim; for the dictionary-word study group, the respective figures were .21 versus .14). There were no other significant effects in the adverb-clauses model.

Second, there was a significant parallelism test on treatment group for the model for total clauses (multivariate F [2,13] = 4.63, p < .05). There was also a significant effect for treatment by story within time (multivariate F [2,14] = 5.57, p < .02); for treatment by story in interim testing (univariate F [1,15] = 4.99, p < .05); and for treatment by story in final testing (univariate F [1,15] = 7.4014, p < .02). Again, the significant parallelism test made interpretation of results difficult (M. Applebaum, personal communication, December, 1983). Inspection of the group means for total clauses in Table 4 suggests one pattern: the story structure group appeared to include
more clauses than the dictionary-word study group in story two (a story about people) than on story one (a story about animals) at interim and final testing, and the dictionary-word study group included more clauses on story one than on story two at interim and final testing. The pattern was reversed prior to treatment.

**Length**

To assess the instructional effects on the number of words produced in the writing samples, two mixed model repeated measures analyses of covariance were done, one for number of words, and one for number of T-units. In each model, the between factor was treatment (experimental and dictionary-word study group), the within factors were time (interim and final) and story within time (two at each testing time). The covariates were scores for number of words on the two pre-treatment writing samples and scores for number of T-units on the two pre-treatment writing samples, respectively.

In both models there were significant tests of parallelism of regression planes on treatment groups, indicating that single covariate adjustments could not be made for both groups (for the test of parallelism on treatment for number of words, multivariate $F[2,13] = 7.65, p < .01$; for the test of parallelism on treatment for T-units, multivariate $F[2,13] = 4.67, p < .03$). No other significant effects emerged. The significant parallelism tests made results difficult to interpret. No clear statistical remedy is evident for such a situation (M. Applebaum, personal communication, December, 1983). As Table 1 shows, the group means for number of words and for number of T-units at the final testing time appear to favor the experimental group, but the experimental group had such a large advantage prior to treatment that it is likely that no meaningful statement regarding effects of instruction on length can be made.

**Discussion**

In conclusion, the most important finding of the study was that instruction
in narrative structure did have a strong positive effect on organization in children's writing. With regard to indirect results of such instruction on writing, importantly, overall creativity, and more specifically, certain aspects of creativity, i.e., uniqueness and language usage, were enhanced as a result of the special instruction. There was one significant effect on cohesion/coherence, i.e., on the average, the story structure group produced fewer cohesive errors than did the dictionary-word study group. Finally, there were no clear or meaningful effects on syntax or on length of protocols.

Concerning the effect of instruction on story organization in writing, it appears that the special instruction did result in an internalized representation of a prototypical story structure. The findings support the conjecture that such an internalized representation may serve as a valuable heuristic or planning device for composing stories. At least one other study of the effects of story structure instruction on children's writing showed similar results in that, after instruction, fifth graders tended to include more text structure categories in written stories (Gordon & Braun, 1983).

The results for creativity are especially encouraging. Concerns about formulaic application of the story structure appear unwarranted. Acquisition of a narrative structure enhanced certain aspects of creativity. Possibly, students used the structure as a framework, a rudimentary aspect of composition, which, once acquired, did in fact release attention and processing time for more creative aspects. Interestingly, the narrative structure instruction resulted in a significant treatment effect for language usage, even though the control group received dictionary-word study training.

For cohesion/coherence, the finding of only one treatment effect, i.e., for cohesive errors, is interesting. One explanation of the failure to affect patterns in specific cohesive ties is that the children in the present study had originally demonstrated low level organizational skills in writing and might
have been among the poorest writers in their grade. An earlier developmental study of cohesion supports such a contention. When compared to third and fifth graders, the children in the present study had remarkably little skill in producing cohesive/coherent text. The proportions of lexical, reference, and conjunctive ties per composition for the children in the present study (prior to instruction) were about half the size of the proportions reported in the earlier study (Bracewell, et al., 1981). For children with very low levels of ability, it would seem particularly difficult to indirectly affect specific writing skills, such as cohesion; such children might require direct instruction in cohesion.

A second intriguing aspect of the results for cohesion and coherence is that although changes in several specific cohesive tie patterns did not emerge, there were other indications of parallels to developmental trends in previous studies. As a result of instruction, the narrative structure group produced fewer cohesive errors than the other group and both groups decreased the proportion of cohesive errors and of remote ties over time. Such results do suggest a move from exophoric to endophoric reference, paralleling developmental findings from other work (cf., McCutchen & Perfetti, 1982; Rentel & King, 1983).

A final speculation with regard to the cohesion/coherence results is that the special narrative instruction may have affected overall coherence rather than specific cohesive ties. If a text is coherent, it should be unambiguous at all points (Halliday & Hasan, 1976). Cohesive ties indicate lack of ambiguity, whereas cohesive errors indicate ambiguity. Perhaps as indicators of ambiguity, cohesive errors are more sensitive reflectors of overall coherence (or the lack of coherence) than are cohesive ties. There is some evidence to support such conjecture. Tierney and Mosenthal (1983) found no significant relationship between ratings of coherence and measures of cohesion following Halliday and
Hasan's procedures. Similarly, Bertrand (1983) found a low correlation (.12) between holistic ratings of what was called cohesion and scores secured following Halliday and Hasan; but there was a high correlation (.49) between the holistic ratings and cohesive errors per 100 words.

For syntax, the only clearly significant effect was an overall increase over time, for both groups, in use of adverb clauses. It was anticipated that the instruction in temporal relationships between story parts could be manifested in the children's use of subordination when writing, which would be largely reflected by use of adverb clauses. However, such an outcome was not anticipated for the dictionary-word study group, nor is it readily explained.

An explanation for the lack of effects on length is not evident. It appears that internalization of a story structure enhanced the quality of invention or elaboration (at least as measured by the creativity subscales), but not the extent of invention or elaboration.

A further important aspect of the instructional effects on organization of compositions and on creativity and cohesive errors was that the effects were realized quickly, i.e., at the end of the first phase of teaching. The effects for organization and cohesive errors were maintained over time, but the magnitude of difference between the two groups did not increase; for overall creativity, the magnitude of difference did increase over time, but this was because the dictionary-word study group's scores decreased. From a practical viewpoint, it is encouraging to find that such effects can be secured in a relatively short period of time.

One remarkable result of the study is the rarity of differences in findings between stories. There were no important differences due to presentation of different story stems. In story structure studies involving comprehension and/or recall, story stem or story content effects are not uncommon (cf., Fitzgerald & Spiegel, 1983; Whaley, 1981). The reliability of results across
stories in the present study attests to the strong influence of the internalized
narrative structure.

It is important to note that dictionary-word study also decreased the
group's proportion of cohesive errors and of remote ties over time and increased
use of adverb clauses over time. It is difficult to tell whether the effects
were due to practice, to dictionary-word study, or to both.

As might be expected, on the whole, the special instruction did appear to
have a greater effect on macro-level characteristics of text, i.e.,
organization, overall coherence (at least as indexed by a decrease in cases of
cohesive errors), and overall creativity and uniqueness. (Originality, which
might be considered a macro-level characteristic was not affected.) Micro-level
characteristics, i.e., length, syntax, cohesiveness, and creative idea
production were not affected at all. (Creative language usage was the only
micro-level characteristic that was significantly affected.)

Finally, although the effects of the special story structure instruction on
story organization in writing and on some other variables were striking, it is
not clear from the present study whether any of the treatment effects could be
attributed more to certain aspects of the instruction than to others. For
example, it is possible that either the teacher-pupil direct instruction
component or abundant use of special activities was primarily responsible for
the positive findings. At present, demonstration of effects of controlled, yet
practical and ecologically valid, instruction is a worthwhile endeavor.
Follow-up studies might now attempt to pinpoint the critical facets of such
instruction.

In summary, the findings of the present study suggest that direct
instruction in narrative structure appears to provide children with a powerful
heuristic or planning device for organizing compositions. Such instruction also
appears to particularly benefit creativity in story writing.
References


Collins, A., & Smith, E. E. (in press). Teaching the process of reading
comprehension. Intelligence.


Glenn, C. G. & Stein, N. L. (in press). Syntactic structures and real world themes in stories generated by children (Tech. Rep.). Champaign, IL:
Narrative Structure

University of Illinois, Center for the Study of Reading.


Table 1
Means (and Standard Deviations) for Organization and Length

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment Group</th>
<th>Pretreatment</th>
<th>Interim</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Organization</td>
<td></td>
<td>(1.20)</td>
<td>(1.80)</td>
<td>(3.60)</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>(1.23)</td>
<td>(1.23)</td>
<td>(1.32)</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>2.33</td>
<td>1.56</td>
<td>2.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.23)</td>
<td>(1.13)</td>
<td>(1.45)</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>124.30</td>
<td>103.80</td>
<td>74.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(47.00)</td>
<td>(34.62)</td>
<td>(21.37)</td>
</tr>
<tr>
<td>Number of Words</td>
<td></td>
<td>97.22</td>
<td>49.11</td>
<td>70.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(43.33)</td>
<td>(24.77)</td>
<td>(29.28)</td>
</tr>
<tr>
<td></td>
<td>NS</td>
<td>16.60</td>
<td>15.20</td>
<td>10.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6.55)</td>
<td>(4.39)</td>
<td>(3.50)</td>
</tr>
<tr>
<td>Number of T-units</td>
<td></td>
<td>13.33</td>
<td>6.00</td>
<td>8.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6.78)</td>
<td>(3.54)</td>
<td>(3.50)</td>
</tr>
</tbody>
</table>

1 NS=Narrative Structure; DW= Dictionary-Word Study
Table 2
Means (and Standard Deviations) for Creativity

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment Group</th>
<th>Pretreatment</th>
<th>Interim</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Creativity: Uniqueness</td>
<td>NS</td>
<td>2.70</td>
<td>2.50</td>
<td>2.70</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>2.44</td>
<td>1.56</td>
<td>2.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.95)</td>
<td>(.85)</td>
<td>(.68)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.01)</td>
<td>(.88)</td>
<td>(.44)</td>
</tr>
<tr>
<td>Creativity: Idea Production</td>
<td>NS</td>
<td>3.49</td>
<td>4.10</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>3.33</td>
<td>2.22</td>
<td>3.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.32)</td>
<td>(.97)</td>
<td>(1.41)</td>
</tr>
<tr>
<td>Creativity: Language Usage</td>
<td>NS</td>
<td>1.60</td>
<td>2.10</td>
<td>2.60</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>1.56</td>
<td>1.56</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.73)</td>
<td>(1.01)</td>
<td>(.53)</td>
</tr>
<tr>
<td>Creativity: Originality</td>
<td>NS</td>
<td>2.60</td>
<td>2.40</td>
<td>3.30</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>2.44</td>
<td>2.00</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.33)</td>
<td>(1.12)</td>
<td>(.87)</td>
</tr>
<tr>
<td>Mean Score on Creativity</td>
<td>NS</td>
<td>2.63</td>
<td>2.78</td>
<td>3.05</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>2.44</td>
<td>1.83</td>
<td>2.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.91)</td>
<td>(.81)</td>
<td>(.61)</td>
</tr>
</tbody>
</table>

1NS = Narrative Structure; DW = Dictionary-word study
<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment Group</th>
<th>Pretreatment</th>
<th>Interim</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Total of Number of Ties per 100 Words</td>
<td>NS</td>
<td>37.98</td>
<td>40.20</td>
<td>35.18</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>35.64</td>
<td>30.73</td>
<td>33.61</td>
</tr>
<tr>
<td>Number of Reference Ties per 100 Words</td>
<td>NS</td>
<td>10.76</td>
<td>15.73</td>
<td>13.64</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>12.95</td>
<td>12.93</td>
<td>13.51</td>
</tr>
<tr>
<td>Number of Conjunctive Ties per 100 Words</td>
<td>NS</td>
<td>6.02</td>
<td>6.19</td>
<td>8.06</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>6.89</td>
<td>3.82</td>
<td>5.39</td>
</tr>
<tr>
<td>Number of Lexical Ties per 100 Words</td>
<td>NS</td>
<td>20.68</td>
<td>17.67</td>
<td>12.87</td>
</tr>
<tr>
<td>Number of Cohesive errors per 100 Words</td>
<td>NS</td>
<td>1.56</td>
<td>2.01</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>1.66</td>
<td>2.19</td>
<td>3.33</td>
</tr>
<tr>
<td>Proportion of Ties that were Immediate</td>
<td>NS</td>
<td>.54</td>
<td>.49</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>.51</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>Proportion of Ties that we Mediated</td>
<td>NS</td>
<td>.23</td>
<td>.31</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>.30</td>
<td>.19</td>
<td>.31</td>
</tr>
<tr>
<td>Proportion of Ties that were Remote</td>
<td>NS</td>
<td>.40</td>
<td>.31</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>.39</td>
<td>.24</td>
<td>.28</td>
</tr>
</tbody>
</table>

^NS=Narrative Structure; DW=Dictionary-Words Study
Table 4
Selected Means (and Standard Deviations) for Syntax

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pretreatment</th>
<th>Interim</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment Group</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Adverb Clauses per T-unit</td>
<td>NS</td>
<td>.17 (.15)</td>
<td>.04 (.05)</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>.07 (.11)</td>
<td>.21 (.19)</td>
</tr>
<tr>
<td>Total Clauses per T-unit</td>
<td>NS</td>
<td>.33 (.19)</td>
<td>.14 (.17)</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>.17 (.19)</td>
<td>.25 (.19)</td>
</tr>
<tr>
<td>Syntactic Density&lt;sup&gt;1&lt;/sup&gt;</td>
<td>NS</td>
<td>.57 (.35)</td>
<td>.31 (.24)</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>.30 (.20)</td>
<td>.37 (.26)</td>
</tr>
<tr>
<td>Proportion of T-units</td>
<td>NS</td>
<td>.63 (.19)</td>
<td>.57 (.20)</td>
</tr>
<tr>
<td>with Subordination</td>
<td>DW</td>
<td>.59 (.14)</td>
<td>.48 (.24)</td>
</tr>
<tr>
<td>Mean T-unit Length</td>
<td>NS</td>
<td>7.92 (2.19)</td>
<td>7.08 (1.80)</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>7.59 (1.69)</td>
<td>8.76 (2.35)</td>
</tr>
</tbody>
</table>

<sup>1</sup>NS=Narrative Structure; DW=Dictionary-word Study