COGNITIVE APPROACHES TO THE ANALYSIS OF TECHNICAL UNDERSTANDING

Nancy K. Atwood
Eva L. Baker

CSE Report No. 231
1984

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

TABLE OF CONTENTS

INTRODUCTION

One of the most central issues in learning concerns the comprehension of text. Understanding the process by which readers extract meaning from text and the factors influencing their comprehension requires inquiry on a broad range of significant areas. Included in these areas of inquiry are processes of human cognition and development such as language development, information storage, processing and retrieval, metacognitive strategies for identifying and organizing information, and individual differences in human capabilities. Beyond factors inherent to the text reader are technical issues related to text design. For example, it is necessary to understand the influence of prose varying in purpose, structure, and complexity on the nature and degree of comprehension. In addition, more metaphysical and epistemological concerns must be addressed. The simple questions of what constitutes knowing and what comprises knowledge greatly influence technical solutions and their utility. Furthermore, all of these issues are predicated on a satisfactory answer to the "how do you know?" question at the heart of any assessment or evaluation of comprehension.

To return specifically to the problem of text design, great progress has been made through the use of computer aided techniques, including word processing and text editors. Traditional approaches to the design of text still dominate the field with attention given to the idea of "readability," a catch-all concept thought to be a summary of text difficulty. Without describing the shortfalls of "readability" as a concept or a set of technical procedures, they are nonetheless based on a relatively formal set of analyses dealing with syntax, word difficulty, and density of prose. What is ultimately needed is a new view of readability, one that casts text design in relation to the function the text is to serve, the nature and complexity of the ideas and structures presented, and the explicit relationship of text features to the range of student performance, including the intellectual level of functioning the prose is to promote in the reader.

The study reported here attempts to build upon extant work in a number of specific areas related to text comprehension. It involves the use of different versions of text in order to examine the effectiveness of syntactic simplification procedures. It involves the analysis of the semantic content of the text and of student generated responses using analysis by subject matter experts as well as a more linguistically-based propositional analysis system. The study accounts for individual differences among readers based on their anticipated expertise with particular content, expectations derived from their experience with similar or identical content. It integrates the issue of text comprehension, a largely verbal phenomenon, with schematic or visual learning. In addition, it addresses a critical area of text learning, that of technical material. Thus, the end-in-view of this effort is to describe more fully the influence of textual manipulations and individual differences on comprehension of a specific set of technical text segments.

OVERVIEW

The study examined the effectiveness of text segments in conveying information about the basic operations of a technical system to military personnel. Two versions of text were contrasted with subjects of three different types of experience. After reading a short series of segments, each subject explained what he had understood from what he had read. These verbal protocols were taped, transcribed, and analyzed according to the overall ideas and specific information each contained. In addition, subjects completed schematic drawings intended to represent the functional hardware and the signal paths in the system before and after reading the text segments. While the study has only limited scope and generalizability, it provides information related to the following research questions:

1. What were the syntactic and semantic characteristics of the original and revised versions of technical text?

2. What was the differential effectiveness of the original and revised technical text on the immediate recall of subjects?

3. What was the effect of different levels and types of experience with the NATO Seasparrow System or similar systems in the immediate recall of subject?

4. What differences were there in comprehension of main ideas and specific information?

5. How did text version and experience influence comprehension of main ideas and specific information?

6. What was the relationship between spatial understanding of the system and text version, experience level, and verbal recall?

The remaining sections of this paper describe the procedures and results of the study followed by a brief discussion of the findings and directions for future research.

PROCEDURES

Subjects

Eighteen naval fire technicians served as subjects in the experiment. Six were experienced Nato Seasparrow technicians who had attended Nato Seasparrow "C" School and had an average of 24 months of on-the-job training on the Nato Seasparrow system. Six were novice Nato Seasparrow technicians who had not attended Nato Seasparrow "C" School and had an average of 15 months of on-the-job training on the Nato Seasparrow system (NSS). The remaining six were experienced technicians trained on a system other than the Nato Seasparrow system who had not attended Nato Seasparrow "C" School and had no on-the-job training on the Nato Seasparrow system.

Text Versions

Two versions of a section of technical text were used in the study. The text was drawn from the technical manual for the Nato Seasparrow system and provided a functional description of the air designate mode of the system. The text was segmented into twenty-five parts meant to reflect discrete sets of information.

The "original" version of the text was the published text taken directly from the technical manual initially prepared by the contractor. The "revised" version was subsequently prepared by the contractor in accordance with guidelines taken from the General Style and Format Requirements stated in the military specifications (MIL-M-38784A). These guidelines directed the contractor to examine the prose for vague and ambiguous terms, long or complex sentences, use of the passive voice, and the use of unnecessary technical words. More specifically, the contractor was advised that the revised version should meet the following criteria: 1) an overall grade level of 10.0 as measured by the Kincaid Readability Formula; 2) no segment with a grade level exceeding 13.0 as measured by the Kincaid formula; 3) no sentence greater than 30 words in length; 4) all sentences in the active voice; 5) no sentence that begins with a preposition; and 6) all simple sentences with no more than two phrases. While the contractor reported using the above criteria in making revisions, it was also indicated that additional revisions were made on the contractor's own initiative to simplify and clarify the text.

Experimental Procedures

All subjects participated in the study in individual sessions. They were told that they would be reading segments of text taken from the Nato Seasparrow Technical Manual and asked to tell how they would explain the material to another technician like themselves. Subjects were asked to imagine someone they worked with who is just learning the system and actually picture themselves explaining the passages to him. Subjects were asked to identify the name of someone on their ship to whom they could imagine making these explanations.

Before beginning to read the text, subjects were provided with a sheet of paper identifying four pieces of hardware on the Nato Seasparrow system (antenna, SDC Input Channels, RF Circuit, CRT Circuit). They were asked to draw a schematic diagram indicating the major hardware components and their interconnections.

After completing the schematic drawing, subjects were provided with one of the text versions and the 7-page schematic drawing from the technical manual. They were asked to read each segment and to use the schematic as they wished. After reading and studying each segment, subjects were asked to explain aloud what they had read to the "buddy" whom they had identified earlier. The sessions were tape-recorded for subsequent analysis and the experimenter recorded the time spent reading and studying, using the schematic, and the time spent explaining. Finally, subjects were asked to make another schematic drawing of the system hardware and interconnections on a new sheet of paper (with the same identified hardware).

RESULTS

The results of the study center around analyses of three sources of data: the technical prose, transcripts of the verbal protocols of subjects audiotaped during the experiment, and the schematic drawings of subjects completed after reading the text.

Three segments were sampled for analysis of the prose and the corresponding protocols (Segments 4, 5, 7). These segments were judged by the electrical engineering experts (EEE's) as low, moderate, and more difficult in technical complexity, respectively. Furthermore, they occurred after the warm-up phase of the experiment during which subjects were getting accustomed to the procedure of speaking aloud, but before fatigue set in.

The following sections present the results emerging from the analyses of the three primary data sources. Analyses of the prose are described first, followed by a discussion of the protocol and schematic analyses.

Prose Analysis

Both the syntactic and semantic content of the sampled prose segments were analyzed. The results of the syntactic analyses are presented first, followed by the results of the semantic analyses.

Syntactic Analysis. The syntax of the sampled segments was analyzed using the Writer's Workbench program developed by Bell Laboratories. This computer program provides information on a variety of syntactic indicators including readability level, sentence length, sentence types, verb usage, and sentence beginners.

Table 1 summarizes the results for the original and revised versions of the three sampled segments. These findings indicate that in general the revised text was simplified according to the guidelines provided to the

Table 1

Syntactic Characteristics of Text

Segment

Syntactic Indicator	4		5		7	
	0*	R*	0	R	0	R
Kincaid Readability Level	13.7	10.8	14.9	10.2	16.0	12.3
Average Sentence Length	27.0	19.5	27.3	17.1	29.3	21.0
Sentence Type:						
% Simple	43	60	36	67	50	89
% Complex	29	40	18	0	38	0
% Compound-Complex	29	0	27	7	0	0
% Compound	0	0	18	27	13	11
% Passive Verbs	56	67	48	48	55	60
% Sentences Beginning with a preposition	14	0	9	0	13	0

* 0 = Original
* R = Revised

contractor. The readability levels of the original versions were reduced 3-4 grade levels in the revised versions as measured by the Kincaid readability formula. Sentence length was substantially reduced and the percentage of simple sentences was increased in the revised compared to the original versions. As directed, the contractor did modify sentences beginning with a preposition in the revised versions; however, the percentage of passive verbs was not systematically reduced.

In sum, the syntax of the revised version was simplified from the original version. However, passive verb constructions were frequent in both versions.

Semantic Analysis. The semantic content of the sampled segments of prose was analyzed using two sets of procedures. First, the electrical engineering experts identified the "gist" units or main ideas, and the information units represented by the content. In addition, the EEE's rated the information units on a number of dimensions, such as hierarchical level and type of topic. Second, a propositional analysis of the prose content was undertaken using the system developed by Kieras for use with technical text (Bovair & Kieras, 1981).

Turning first to the analyses of the subject matter experts, one of the electrical engineering experts formulated gist units to represent the macrostructure of the text and information units to represent its microstructure. The gist units were developed after Pearson and Camperell (1981). Taken together, they provide an overview of the passage and roughly approximate terms such as key episode, theme, main plot, and top-level discourse structure used by other writers.

Information units were devised using the guidelines and criteria of Van Matre (1975). Each information unit consisted of a simple sentence that expresses a single fact. In re-writing the text into information units, the electrical engineering expert (EEE) was guided by the following criteria:

1) Rewrite compound sentences as a group of simple sentences.

2) Reconstruct compound subjects, verbs, and objects as a group of simple sentences expressing a single fact per sentence.

3) Rewrite series of phrases as simple sentences that express a single fact per sentence.

4) Take into account the difference between logical "and/or" and the connective "and/or".

5) Treat compound nouns representing a single term as a single word rather than as a noun preceded by a group of modifiers.

6) Use the active voice where possible.

Part A of Table 2 summarizes the results of the initial analyses of the EEE. Each version of the sampled segments contained the same number of gist units. Further, the gist units were identical for each pair of segments suggesting that the revision did not alter the overall nature or amount of information presented.

The original and revised versions did differ in the number of information units that they contained. While both versions of segment 4 were equivalent on this dimension, the revised version of segment 5 contained more information units than the original version. This pattern was reversed for segment 7. However, the revised versions did contain more sentences than the original versions. Thus, when a measure of idea density was constructed (a ratio of the number of information units to the number

Table 2

Semantic Analysis of Prose Using

Information and Gist Units

Syntactic Indicator	4		5		7	
	0	R	0	R	0	R
A. Summary of Initial Analysis						
No. of Gist Units	3	3	5	5	2	2
No. of Information Units	24	24	26	34	27	20
No. of Sentences	6	10	10	15	8	9
Density	4.0	2.4	2.6	2.3	3.4	2.2
B. Nature of Information Units*						
Hierarchical Level*	2.11	1.83	1.96	1.70	2.30	1.75
Topic-% System	15.8	16.7	23.1	39.4	29.6	25.0
% Signal	21.1	29.2	11.5	18.2	18.5	5.0
% Hardware	21.1	0	38.5	12.1	11.1	10.0
% Signal + Hardware	42.0	54.1	19.2	30.3	22.2	40.0
% Signal + System	0	0	3.85	0	14.8	10.0
% System + Hardware	0	0	3.85	0	3.8	10.0
Salience*	1.79	1.33	1.42	1.52	1.93	1.30

* Rated on a 3-point scale where 1=high and 3=low.

of sentences), the density of the revised versions was lower than that of their respective original versions.

Subsequently, the other EEE examined the nature of the information units by rating each one on hierarchical level and salience as well as identifying the type of topic presented. Part B of Table 2 summarizes the results of this analysis.

Generally, the information units in the revised versions tended to present ideas at a higher hierarchical level (i.e., main ideas rather than supporting details) and with greater salience to the overall passage. Further, information units in the revised versions tended to address the overall system or the interrelationship between the signal and the hardware to a greater extent than those in the original version.

Table 3 summarizes the results of the propositional analyses of the sampled segments. The propositional analyses constitute a more fine-grained examination of the text microstructure. As can be seen in Table 3, the revised versions contained fewer propositions per sentence than the original versions. Furthermore, there was less variability on this dimension in the revised versions. When the type of propositions per sentence was examined, the revised versions tended to show a reduction in verb-based propositions (and to a lesser extent, modifiers) compared to the original versions. Further, when embedded propositions (propositions referring to other propositions within the same sentence) were examined, their incidence was also reduced in the revised version. Finally, the number of propositions with missing subjects, i.e., passive constructions, was slightly reduced in the revised versions in two of the sampled segments.

Table 3

Propositional Analysis of Prose
(after Kieras)

Segment

Measure	4				5				7			
	O		R		O		R		R		O	
	$\overline{X}$	S	$\overline{X}$	S	$\overline{X}$	S	$\overline{X}$	S	$\overline{X}$	S	$\overline{X}$	S
1. Propositions per sentence	1.17	1.60	4.70	1.77	8.80	4.59	5.87	2.10	8.63	4.31	5.00	1.58
2. By type per sentence:												
verb-based	2.83	1.47	1.60	1.07	3.70	1.77	2.47	1.41	2.88	1.55	1.67	0.71
modifiers	1.83	1.83	1.60	1.58	2.60	2.88	1.93	1.44	3.38	2.07	1.44	1.13
prepositions	2.67	0.82	2.00	1.15	2.30	1.34	2.27	3.61	2.75	1.67	1.78	1.09
3. Embedded propositions per sentence	2.50	1.05	2.00	1.25	2.80	3.33	1.40	1.59	3.38	2.83	2.00	0.87
4. Propositions with missing subject per sentence	1.50	1.05	1.10	1.10	1.60	2.72	0.80	1.15	0.63	0.74	0.67	0.71

Thus, the results of both sets of analyses suggest that the original and revised versions of the sampled segments were equivalent in the overall amount and nature of the information provided. However, the sentences in the revised version were simpler in their semantic content. They were lower in "idea density" and higher in hierarchical level and salience. Further, they tended to concern the overall system or the relationship between signal and hardware. The revised versions also showed fewer propositions per sentence and less embedding of propositions within sentences than the original versions.

Protocol Analysis

The amount and nature of comprehension was examined through an analysis of subjects' verbal protocols. Scoring procedures and reliability estimates are described first. Then analyses of information unit comprehension, gist unit comprehension, and their relationship to individual difefereneces among subjects are presented.

Scoring Rules and Reliability. Transcripts of the verbal protocols of subjects explaining aloud the material were used to examine the degree of comprehension of both gist and information units. Since the gist and information units contain multiple propositions and partial comprehension was possible, rating scales were developed to assess the degree of comprehension (rather than a binary yes/no judgment). Somewhat different procedures were used for scoring the gist and information units. A 5-point scale was used to score the gist units and scoring was conducted liberally based on the literal statements in the transcripts as well as inferences about their comprehension that followed from what they had said. During training, the two scorers were able to reach a high level of agreement (in fact 100%) using these procedures (see Table 4).

Table 4

Reliability Estimates for Protocol Scoring

Type of Unit	N	% Agreement	Mean Discrepancy Over all Items	Mean Discrepancy Over Items with Disagreements
I. Actual Scoring				
Information Units (3-point scale)	155	90.87	0.12	1.05
Gist Units (5-point scale)	20	67.23	0.51	1.33
Other (Frequency Counts)	38	88.73	0.11	1.00
II. Final Training Segment				
Information Units (3-point scale)	28	85.71	0.81	1.25
Gist Units (5-point scale)	3	100.00	0.0	0.0
Other (Frequency Counts)	6	100.00	0.0	0.0

For the information units, a more stringent policy was adopted that required the same wording or synonyms in the transcripts. The two scorers were able to score the information units reliably using a 3-point scale where 1 = little or no comprehension, 2 = partial comprehension, and 3 = full or almost full comprehension. Their level of agreement was about 86% on the final training segment (see Table 4).

In addition, the scorers identified the frequency of connecting statements (with previous text segments and between the text and the schematic), elaborations beyond the text (system-specific or general electronics), and meta-statements about the audience or the comprehension task. During training, a 100% level of agreement was attained by the two scorers.

Table 4 also summarizes the reliability of scoring for the actual scoring. As shown in the table, high reliability was maintained on the information units and the other frequency counts. Interrater agreement fell on the gist units to 67%; however, disagreements were generally only one point off on a 5-point scale.

Comprehension of Information Units. Subjects' comprehension of the information units contained in the text were examined using two strategies. First, the relationship of experience and text version (the two primary independent variables in the study) on the amount of comprehension of information units was examined using two-way analyses of variance. Second, the influence of these variables on the type of information units comprehended was examined.

The results of the first set of analyses are summarized in Table 5. Separate analyses were conducted for each text segment with amount of comprehension measured by subjects' average comprehension ratings over all

Table 5

Influence of Experience and Text on Information Unit Comprehension (3-point Scale)

Segment 4

a. ANOVA

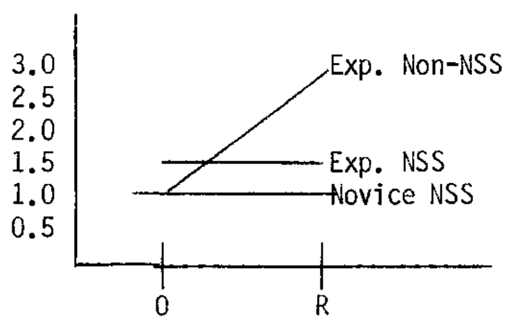
Source	df	SS	F
E	2	2.7089	17.23***
T	1	2.0842	26.53***
EXT	2	2.1082	13.41***
Error	12	0.9433	

***p<.001

b. Cell Means

	O	R	
Exp. Non-NSS	1.22	2.86	2.04
Exp. NSS	1.25	1.57	1.41
Novice NSS	1.07	1.15	1.11
	1.18	1.86	1.52

c. Graphic Display



Segment 5

a. ANOVA

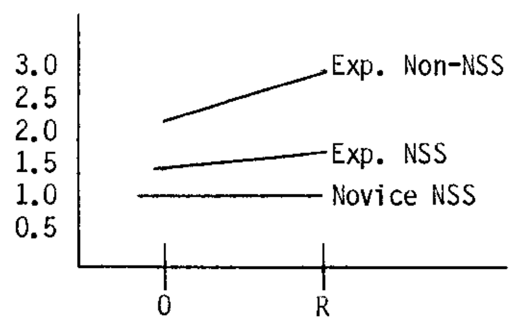
Source	df	SS	F
E	2	2.6099	22.93***
T	1	1.5744	27.66***
EXT	1	1.4349	12.61***
Error	12	0.6830	

***p<.001

b. Cell Means

	O	R	
Exp. Non-NSS	1.53	2.87	2.20
Exp. NSS	1.35	1.78	1.57
Novice NSS	1.29	1.28	1.29
	1.39	1.98	1.69

c. Graphic Display



Segment 7

a. ANOVA

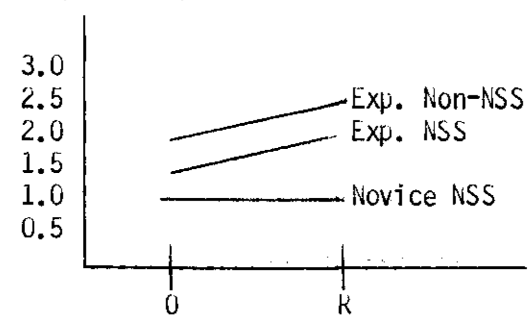
Source	df	SS	F
E	2	1.4726	13.35***
T	1	2.3835	43.22***
EXP	2	0.6627	6.01**
Error	12	0.6618	

**p<.01
***p<.001

b. Cell Means

	O	R	
Exp. Non-NSS	1.32	2.55	1.94
Exp. NSS	1.26	1.92	1.59
Novice NSS	1.09	1.38	1.23
	1.22	1.95	1.59

c. Graphic Display



information units in the segment. The main effects for experience (E) and text (T) as well as their interaction (EXT) were statistically significant for all three sampled segments. When the cell means were graphed and examined using Scheffe's test for post hoc comparisons (Kirk, 1982), a similar pattern emerged in all three cases. The experienced non-NSS technicians showed significantly greater information unit comprehension with the revised compared to the original text. Further, with the revised text the experienced non-NSS technicians performed significantly above the experienced and novice NSS technicians. The three experience groups did not differ significantly with the original version nor did the experienced and novice technicians differ with the revised version.

The second set of analyses examined the type of information units comprehended based on ratings and categorizations made by one of the EEE's. These included hierarchical level of the information unit, salience of the unit to the overall gist of the segment, and type of topic. The first two types were measured on a 3-point rating sale where 1 = high and 3 = low, while the last was based on a category system of six possible topics or combinations of topics.

Separate analyses were conducted for each text segment. Each analysis included two between-subjects variables (experience and text) and one within-subjects variable (type of information unit). The dependent measures for each subject were average comprehension ratings for information units at each level or category of the scale being examined (e.g., hierarchical levels 1, 2, 3).

Table 6 summarizes the results for hierarchical level and Table 7 for topic. (No significant differences were obtained for the salience ratings, so the results are not presented here.)

As shown in Table 6, the between-subjects portion of the analyses replicated the effects of experience level, text version, and their interaction described earlier. The within-subjects portion revealed a significant main effect for hierarchical level for segments 5 and 7. Post hoc comparisons among the marginal means using Scheffe's test indicated that subjects tended to show more recall of the units rated high in hierarchical level (i.e., main ideas) compared to those rated lower in hierarchical level (i.e., supporting details).

As shown in Table 7, the effects of experience level, text version, and their interaction were again replicated as expected. In addition, statistically significant interactions between topic and text were obtained for segments 5 and 7. (A main effect for topic with segment 7 was also obtained.) Post-hoc comparisons of the means for the topic by text interaction were conducted using Tukey's HSD test (Kirk, 1982).

The pattern of results was somewhat different for segments 5 and 7. With segment 7, system-related topics tended to be comprehended more than more specific topics, particularly with the revised version. With segment 5, there was greater comprehension of hardware with the revised version compared to the original text.

Comprehension of Gist Units. The influence of experience and text version on subjects' overall comprehension of gist units was also examined using two-way analyses of variance. Again, analyses were conducted separately for each segment where gist unit comprehension was measured by average gist unit comprehension ratings for each segment. A somewhat complicated pattern of results emerged, as shown in Table 8. For segment 5, a similar pattern of results occurred as described above for amount of

Table 6

Influence of Hierarchical Level on Information Unit Comprehension

Segment 4

a. ANOVA

Source	df	SS	F
E	2	6.2881	5.63**
T	1	6.8152	12.20**
ExT	2	5.2440	4.70**
Error	12	6.7014	
HL	2	1.2825	2.95
HLxE	4	1.8232	2.10
HLxT	2	0.3600	0.83
HLxExT	4	5.2196	
Error	24	5.2196	

** p < .01

b. Marginal Means for HL

HL1	1.81
HL2	1.44
HL3	1.66

Segment 5

a. ANOVA

Source	df	SS	F
E	2	8.7549	16.15**
T	1	4.0057	14.78***
ExT	2	5.2518	9.69**
Error	12	3.2528	
HL	2	0.7748	4.25*
HLxE	4	0.0633	0.17
HLxT	2	0.6057	3.32
HLxExT	4	0.2721	0.75
Error	24	2.1868	

* p < .05
** p < .01
*** p < .001

b. Marginal Means for HL

HL1	1.83
HL2	1.67
HL3	1.53

Segment 7

a. ANOVA

Source	df	SS	F
E	2	5.9826	19.29***
T	1	6.8603	44.23***
ExT	2	2.8030	0.94**
Error	12	1.8613	
HL	2	2.4432	9.46***
HLxE	4	0.9288	1.80
HLxT	3	0.4295	1.66
HLxExT	4	1.2110	2.35
Error	24	3.0979	

** p < .01
*** p < .001

b. Marginal Means for HL

HL1	1.95
HL2	1.48
HL3	1.53

Table 7

Influence of Topic on Information Unit Comprehension

Segment 4

a. ANOVA

Source	df	SS	F
E	2	5.8682	7.29**
T	1	5.8246	14.46**
EXT	2	5.4860	6.81**
Error	12	4.8328	
Top	2	1.1753	2.22
TopxE	4	1.1179	1.05
TopxT	2	0.3901	0.74
TopxExT	4	0.4449	0.42
Error	24	6.3601	

** $p < .01$

Segment 5

a. ANOVA

Source	df	SS	F
E	2	9.1671	16.62***
T	1	6.2143	22.53***
ExT	2	5.0573	9.17**
Error	12	3.3096	
Top	3	0.4039	1.43
TopxE	6	0.8806	1.55
TopxT	3	1.3642	4.82**
TopxExT	6	1.3800	2.44
Error	36	3.3993	

* $p < 0.5$
** $p < .01$
*** $p < .001$

b. Means for TopxT Interaction

	O	R
System	1.56	1.98
Signal	1.59	1.85
Hardware	1.32	2.31
Signal & Hardware	1.27	1.96

Segment 7

a. ANOVA

Source	df	SS	F
E	2	7.9206	21 867***
T	1	19.3661	107.06***
ExT	2	4.1333	4.42**
Error	12	2.1708	
Top	5	5.7686	5.90***
TopxE	10	2.8106	1.44
TopxT	5	4.6316	4.74***
TopxExT	10	1.1552	0.59
Error	60	11 7261	

* $p < .05$
** $p < .01$
*** $p < .001$

b. Means for TopxT Interaction

	O	R
System	1.46	2.07
Signal	1.16	1.56
Hardware	1.07	1.72
Signal & Hardware	1.00	1.68
System & Signal	1.33	2.39
System & Hardware	1.00	2.67

Table 8

Influence of Experience and Text on Gist Unit Comprehension (5-point Scale)

Segment 4

a. ANOVA

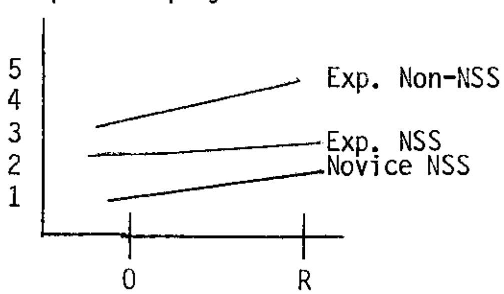
Source	df	SS	F
E	2	10.7160	3.91*
T	1	2.0000	1.46
EXT	2	2.3704	0.86
Error	12	16.4444	

*p<.05

b. Cell Means

	O	R	
Exp. Non-NSS	3.22	4.77	4.00
Exp. NSS	3.11	2.89	3.00
Novice NSS	1.78	2.44	2.11
	2.70	3.37	3.04

c. Graphic Display



Segment 5

a. ANOVA

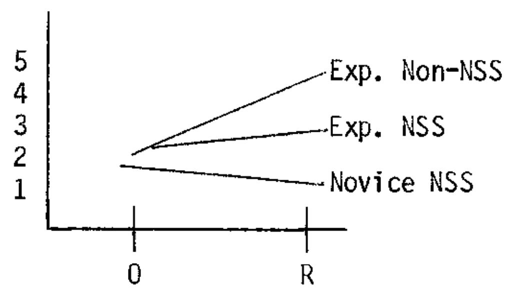
Source	df	SS	F
E	2	10.0844	9.22**
T	1	7.2200	13.21**
EXT	2	8.0133	7.33**
Error	12	6.5600	

**p<.01

b. Cell Means

	O	R	
Exp. Non-NSS	1.87	4.80	3.33
Exp. NSS	1.80	3.00	2.40
Novice NSS	1.67	1.33	1.50
	1.78	3.04	2.41

c. Graphic Display



Segment 7

a. ANOVA

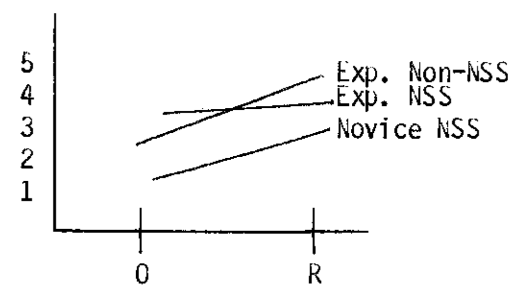
Source	df	SS	F
E	2	7.1111	4.13*
T	1	7.3472	8.53**
EXP	2	1.4444	0.84
Error	12	10.3333	

*p<.05
**p<.01

b. Cell Means

	O	R	
Exp. Non-NSS	3.00	4.83	3.92
Exp. NSS	3.67	4.17	3.92
Novice NSS	1.83	3.33	2.58
	2.83	4.11	3.47

c. Graphic Display



information unit comprehension. Both experienced non-NSS and NSS technicians showed significantly greater gist unit comprehension with the revised compared to the original text version. The novice NSS technicians did not show significant differences in performance with the original and revised versions and they performed significantly below the two groups of experienced technicians on the revised version. For segment 7, all three experience groups showed greater gist comprehension with the revised compared to the original version. In addition, the two experienced groups (non-NSS and NSS) showed significantly greater gist comprehension than the novice NSS technicians. For segment 4, only a main effect for experience was obtained suggesting that the groups were ordered on degree of gist comprehension as follows: experienced non-NSS technicians, experienced NSS technicians, and novice NSS technicians.

A similar analysis was conducted for the frequency of connectives, elaborations, and meta-statements. No statistically significant effects were isolated for experience or text. In fact, the frequency of such statements was generally quite low. The most frequent type of statement was system-specific elaborations which averaged 1.39 per segment (s = 2.89).

Relation to other Individual Difference Measures. The relationship of information and gist unit comprehension to time spent reading and studying, use of the provided schematic diagrams, and individual differences among subjects as measured by their Armed Forces Qualification Test scores (AFQT) and their Armed Services Vocational Aptitude Battery (ASVAB) scores was examined. No significant relationships were identified.

In sum, the revised text version contributed to greater comprehension of information units for the experienced non-NSS and NSS technicians but did not appear to facilitate the information unit comprehension of the novice NSS technicians. There was also some evidence that subjects tended to comprehend more global ideas rather than supporting details, particularly with the revised text.

Schematic Analysis

The schematic drawings prepared by subjects before and after reading the text were scored using a system developed by one of the EEE's. The scoring system produced two sub-scores and a total score. The first sub-score was concerned with the functional topology or nodes of the drawing, e.g., the functional boxes such as the Computer Complex. The nodal hierarchies were additive and weighted in a geometric fashion by level of specificity to emphasize the level of knowledge. Thus, a person could earn a score of 1, 2, 4, 8, or 16 for the computer complex, for example, depending on the level of detail. The second sub-score concerned the signal flow or connections between nodes. One point was assigned for a connection between boxes, two points for the direction of the path, and four points if the connection was named. The scoring system also contained a penalty for incorrect nodes and signals such that points were subtracted for errors drawn. The total score was simply the sum of the function and signal sub-scores.

Because the sub-score distributions were highly skewed, the first step in the analysis was to truncate the distributions and to transform the original scores to a five-point scale. This transformation was accomplished by examining the frequency distributions of the sub-scores and dividing them at naturally occurring breaks between score values.

Analysis of the pre-test drawings indicated a significant main effect for experience on both the function sub-score and the signal sub-score (see Table 9). Post-hoc comparisons between means using Scheffe's method revealed that, on the function sub-score, the experienced and novice NSS technicians performed significantly above the experienced non-NSS technicians. On the signal sub-score, the experienced NSS technicians showed significantly higher performance than either of the other two groups.

Table 10 presents an analysis of the posttest schematic drawings by experience and text version. There was a significant main effect for experience on the function sub-scores. Post-hoc comparisons among means using Scheffe's method showed that, as in the analyses of the pretest drawings, both the experienced and novice NSS technicians performed higher on this dimension than the non-NSS technicians.

The main effect for experience on the signal sub-scores approached statistical significance ($p<.07$); and the pattern of means was somewhat different than at the pretest. In this case, the experienced NSS technicians performed significantly higher than the novice NSS technicians, as at pretest. However, the performance of the experienced non-NSS technicians was not significantly different from the experienced NSS technicians.

Finally, pretest and posttest scores were compared for each of the three experience groups using t-tests for dependent groups. While no statistically significant differences were observed, examination of the means suggests that the novice NSS technicians tended to show more improvement on the function sub-score from pretest to posttest, while the experienced

Table 9

Relationship of Experience and Text
To Pretest Schematic Drawings

A. Function Sub-Score

1. ANOVA

Source	df	SS	F
E	2	12.6627	8.19**
T	1	3.1405	4.06
EXT	2	1.2261	0.79
Error	11	8.5000	

** $p < .01$

2. Experience Group Means

Exp. Non-NSS	2.17
Exp. NSS	4.17
Novice NSS	3.60

B. Signal Sub-Score

1. ANOVA

Source	df	SS	F
E	2	9.3000	5.48**
T	1	0.9444	1.11
EXT	2	4.4333	2.61
Error	11	9.3333	

** $p < .01$

2. Experience Group Means

Exp. Non-NSS	1.50
Exp. NSS	3.00
Novice NSS	1.40

Table 10

Relationship of Experience and Text
to Post-test Schematic Drawings

A. <u>Function Sub-Score</u>

1. ANOVA

Source	df	SS	F
E	2	9.9353	4.15*
T	1	0.1381	0.12
EXT	2	0.9953	0.42
Error	11	13.1667	

*$p<.05$

2.

Experience Group	Means
Exp. Non-NSS	2.50
Exp. NSS	4.00
Novice NSS	4.20

B. <u>Signal Sub-Score</u>

1. ANOVA

Source	df	SS	F
E	2	10.2510	3.28*
T	1	2.5621	1.64
EXT	2	4.1379	1.33
Error	11	17.1667	

*$p<.07$

2.

Experience Group	Means
Exp. Non-NSS	2.33
Exp. NSS	3.33
Novice NSS	1.40

non-NSS technicians tended to show improvement on the signal sub-score.

In sum, the NSS technicians exhibited greater understanding of the functional hardware in their schematic drawings than the non-NSS technicians, probably a function of their direct experience with the NSS system. Further, experienced NSS technicians showed a greater understanding of the signal paths at pretest than the other two groups. However, the experienced non-NSS technicians improved at posttest such that their signal performance was not significantly different from the experienced NSS technicians.

DISCUSSION

This study examined the effects of textual manipulations and previous experience on comprehension of technical text. Measures of comprehension included verbal indicators of the amount and type of information based on the audiotaped protocols as well as spatial indicators of function and signal knowledge based on schematic drawings.

Results supported the initial distinctions drawn between text versions and experience groups. The revised version was indeed syntactically simpler on a number of dimensions such as readability level, sentence length, and sentence structure. Furthermore, while the overall semantic context or gist of corresponding segments remained comparable, the revised version was lower in density of ideas and ideas presented were higher in hierarchical level and salience than the original version.

Further, analyses for the pretest schematic drawings supported that there were pre-existing differences among the three groups of technicians. NSS technicians showed greater functional knowledge in their drawings than non-NSS technicians, while experienced NSS technicians showed greater signal knowledge than either of the other two groups.

Analyses of the verbal comprehension protocols suggest that the revised version had the greatest facilitating effects for experienced non-NSS technicans. Technicians in this group showed performance superior to both experienced and novice NSS technicians in analyses of both the amount and types of information comprehended.

One might speculate that the simpler revised version of the text allowed the experienced non-NSS technicians to apply their general knowledge about electronic systems and to extract the information necessary to understand the specific operations of the Seasparrow System. Such an explanation aligns with work by Rumelhart and Ortong (1977) which suggests that efficient learners call up schema corresponding to the particular type of discourse and instantiate the scheme based on what is read. The experienced non-NSS technicians were familiar with technical manuals generally and the simplifications in the revised version may have facilitated their ability to instantiate their pre-existing schema with specific facts related to the Seasparrow System.

One might further speculate that the hands-on experience of the experienced NSS technicians actually interfered with their instantiation of facts from the text. In several cases, members of this group noted that the text was incomplete or inaccurate (e.g., "it doesn't really work that way"). Meyer (1984) reports a similar phenomenon in which readers disagreed with an author's message, rejected his schema, and substituted their own schema with instantiations based on information outside the text.

The analyses of the verbal comprehension protocols also suggested that main ideas rather than supporting details tended to be better comprehended. This finding replicates the work of Meyer (1975) and others.

Analyses of the posttest schematic drawings revealed the same effects for experience as the pretest drawings. Changes in performance from pretest to posttest suggest that technicians with different types and amounts of experience may be extracting different types of information from the text. Novice NSS technicians tended to show improved functional understanding of the system while technicians experienced in systems other than the Seasparrow System showed improvements in signal knowledge.

The study reported here was exploratory and the results should be cautiously interpreted. However, there are some indications that simplifying the syntax and semantic content of technical text does contribute to increased comprehension, particularly with experts who have prior background knowledge and experience with other electronic systems. But the findings also suggest that such manipulations are not sufficient to facilitate the comprehension of novices with minimal background and experience.

Clearly, additional research is needed before specific recommendations for improving the comprehensibility of technical text can be made. For example, the processes by which experts and novices extract meaning from technical prose and complex schematic drawings needs further attention and their implications for technical writing need to be assessed. Further, the range of purposes and tasks for which such manuals are intended should be clarified and the effects of textual manipulations on the performance of these tasks should be examined. For example, Kieras, Tibbits, and Bovair (1984) examined the effects of different presentation formats on the operation of various electronic equipment by individuals varying in prior experience. Their work suggests that such presentations of technical information should match the informational requirements of the task and expertise of the users. Clearly, study is required to identify such requirements as a basis for making recommendations for the writing and revision of technical manuals.

REFERENCES

Bovair, S., & Kieras, D.E. A guide to propositional analysis for research on technical prose. Technical Report #8. San Diego, CA: NPRDC, 1981.

Kieras, D.E., Tibbits, M., & Bovair, S. How experts and nonexperts operate electronic equipment from instructions. Technical Report #14. Arlington, VA: Office of Naval Resarch, 1984.

Kirk, R.E. Experimental design, 2nd edition. Belmont, CA: Brooks/Cole Publishing Co., 1982.

Meyer, B.J.F. The organization of prose and its effects on memory. Amsterdam: North Holland, 1975.

Meyer, B.J.F. Text dimensions and cognitive processing. In H. Mandl, N. Stein, & T. Trabosso (Eds.), Learning and understanding texts.

Pearson, P.O., & Camperell, K. Comprehension of text structures. In J.T. Guthrie (Ed.), Comprehension and teaching: Research reviews. Newark, Delaware: International Reading Association, 1981.

Rumelhart, D. & Ortony, A. The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), Schooling and the aquisition of knowledge. Hillsdale, NJ: Erlbaum, 1977.

Van Matre, N.H., et al. Learning from lecture: Investigation of study strategies involving note-taking. San Diego, CA: NPRDC, 1975.