BERICHT
aus dem
PSYCHOLOGISCHEN INSTITUT
DER UNIVERSITÄT HEIDELBERG

Diskusssionspapier Nr. 32

Februar 1983

Text Processing with and without Prior Knowledge:

Knowledge- versus Heuristic-Dependent Representations

von

Franz Schmalhofer

6900 HEIDELBERG 1, HAUPTSTR. 47-51, TEL. 06221/547346

Text Processing with and without Prior Knowledge: Knowledge- versus Heuristic-Dependent Representations

Franz Schmalhofer

This paper is based on the author's dissertation submitted to the University of Colorado, in partial fulfillment of the Ph.D. requirements, and was supported by grants NIMH MH-15872 to Walter Kintsch and by USPHS program project grant HDMH11681-0181, Richard Olson, co-investigator. The author would like to thank the chairpersons Walter Kintsch and Peter Polson for their academic guidance in this research, and the committee members K.A. Ericsson, G.H. McClelland, R.K. Olson and A. Ehrenfeucht for invaluable advice. The assistance and comments of Thea Turner, Bill Walker, Raymond Guidon, and Reinhold Kliegl are also much appreciated. Special thanks are owed to Doris Glavanov for helping in all facets of this research.

Abstract

A study of expertise-differences examined whether readers encoded a text by storing its meaning units in the slots of a previously acquired hierarchical knowledge structure. Subjects with and without prior knowledge about computer programming studied a programmer's manual (LISP). For all subject groups, sentence reading times failed to show serial position effects but increased with the number of propositions in a sentence. All subjects successfully remembered the text's meaning rather than its wording. Whereas subjects without prior knowledge only remembered the text itself, subjects with prior knowledge also acquired general knowledge about LISP. It was concluded that text and knowledge representations are distinct: Whereas text memory is a by-product of general comprehension heuristics, the updating of world knowledge critically depends upon a reader's prior knowledge.

The superiority of text memory over memory for unrelated sentences has been explained in two different ways. Whereas theories of knowledge-dependent text encoding emphasize the central importance of a reader's previously acquired domain-specific knowledge structure for the encoding of the meaning of a text (Bower, Black & Turner, 1979; Schank & Abelson, 1977), theories of heuristic-dependent text representations assume that text memory is a by-product of general heuristic comprehension processes (Kintsch & van Dijk, 1978; McKoon & Ratcliff, 1980; Miller & Kintsch, 1980; van Dijk & Kintsch, 1983).

Knowledge-Dependent Text Memory

Bransford & Johnson's (1973) experiments, which found that text memory deteriorates dramatically when relevant prior knowledge is not available to the reader for the encoding of a text (e.g. washing machine paragraph), support the hypothesis that the superiority of text memory is due to a previously acquired domain-specific knowledge structure such as a schema (Bartlett, 1932), script (Schank & Abelson, 1977; Bower et al. 1979), or hierarchical goal structure (Spilich, Vesonder, Chiesi & Voss, 1979). When reading a text, the reader is assumed to select and activate such a hierarchical knowledge structure in memory. This knowledge structure will then guide the encoding of the text so that the text's meaning units are stored in appropriate slots of the reader's activated knowledge structure. The meaning of a text is thus represented by a hierarchical structure, whose superordinate entities represent the reader's previously acquired knowledge whereas the subordinate slots of this structure become filled with a text's meaning units (Schank & Abelson,1977). Scripts, schemata, or MOPs (Schank, 1980) may thus improve the encoding and retrieval of sentences in an identical way as a previously acquired encoding and retrieval structure improves memory for random digit sequences from an average of 8 to 80 remembered digits (Chase & Ericsson, 1981). Thus the superiority of text memory may indeed be due to a reader's previously acquired domain-specific knowledge structure. Consequently, when a relevant knowledge structure can not be activated in memory, such as in Bransford & Johnson's experiment, memory for text deteriorates drastically.

Heuristic-Dependent Text Memory

Kintsch & van Dijk (1978), on the other hand, have proposed that the meaning of a text is represented by a text base which is constructed from elements called propositions. A proposition is the basic meaning unit in memory and consists of a predicate with one or several arguments (Kintsch, 1974). According to the model developed by Kintsch & van Dijk, the propositions of a text are processed by general comprehension heuristics (micro- and macro-processes) which construct a representation of this text in memory. The likelihood of recalling a certain proposition is assumed to depend upon the number of processing cycles in which the heuristic comprehension processes kept this proposition in working memory. Text memory is consequently a by-product of general heuristic comprehension strategies. In support of this assumption Kintsch & van Dijk and Miller & Kintsch (1980) have found that propositions which, according to the processing model of Kintsch & van Dijk, participated in a larger number of processing cycles were indeed recalled with a higher probability than propositions with fewer processing cycles.

Contrary to knowledge-dependent encoding (Bower et al. 1979) the comprehension processes modeled by Kintsch & van Dijk are based upon general heuristic strategies. A text representation formed by these strategies may be quite different from a reader's preexisting knowledge structures. Since a text representation is assumed to be a rather veridical representation of the text itself, the superordinate entities of this structure also represent meaning units of the text rather than a reader's world knowledge.

Since the comprehension strategies of the Kintsch & van Dijk model do not necessarily depend upon a reader's previously acquired knowledge structure, it may be expected that even a reader without prior domain-specific knowledge (low knowledge or LK subject) is capable of applying these strategies to well-written texts. This expectation is supported by Kieras' (1980, 1982) experiments which found that even LK subjects may build macrostructures. Kieras showed that LK subjects can derive main ideas and main clauses surprisingly well, even when the surface structure of the text misleads the reader about its gist.

Although theories of heuristic-dependent text memory assume that the representation of the meaning of a text (text base) is constructed by general comprehension strategies (Kintsch & van Dijk, 1978), no claim is made that a reader's prior world knowledge would not affect the understanding of a text (van Dijk & Kintsch, 1983). However, contrary to theories of knowledge-driven text memory, which assert that world knowledge and meaning units are stored in one hierarchical structure in memory (Bower et al, 1979; Schank, 1980; Spilich et al., 1978), theories of heuristic-dependent text memory postulate that the meaning of a text and world knowledge which a reader had either previously acquired or derived from the current text are two related but nevertheless completely separate entities of human information processing (van Dijk & Kintsch, 1983).

Mental or Situational Models

Whereas the meaning of a text is represented by a text base, the reader's relevant world knowledge is encompassed by a mental (Johnson-Laird, in press), or situational model (van Dijk & Kintsch, 1983). A situational model is thus a model of the real or some hypothetical world (situations thereof), about which a text may present some new information. Van Dijk and Kintsch argue that the text representation and situational model must be kept separate, "because the representation of the text and the representation of the situation do not always coincide. The text representation may very well have its own, distinct existence in memory. Just as one normally remembers the situation the text refers to rather than the text itself, one can and often does remember the text per se - its organization, its macrostructure which need not share important features with the structure of the situation".

Situational models can be developed and updated not only by studying a text, but also through perception and problem solving. Through perception, a situational model can be derived by forming mental images (Kosslyn, 1975; Johnson-Laird, 1980) and in problem solving a situational model can be constructed by learning from doing (Simon, 1980). Since a situational model images the real or some hypothetical world, the structure and the elements of a situational model reflect the properties of that world rather than the properties of the text from which it was derived.

Rather than text and knowledge elements being stored in one structure theories of heuristic-dependent text encoding postulate two separate representations for the meaning of a text and for the world knowledge which a reader has either previously acquired or is deriving from the information presented in the text. The propositional text representation is assumed to be a by-product of general comprehension strategies, whereas the development or updating of a situational model which represents a person's world-knowledge is knowledge-dependent.

The present experiment attempts to distinguish between theories which postulate text memory to be knowledge-dependent and theories which assume text memory to be a by-product of general comprehension processes.

Theories of knowledge-dependent text encoding predict that LK subject's text memory would be poor, because LK subjects lack a knowledge structure for the encoding of a text's meaning units (Spilich et al.). If the text's meaning units are indeed stored in the slots of a hierarchical structure representing the reader's world knowledge, the text's meaning should be difficult to retrieve from memory, unless a subject can also retrieve the superordinate knowledge units. Memory for the meaning of a text should therefore be bound by the world knowledge which the subject has about the situations referred to by the text. In other words, a reader can only have good text memory, if he also demonstrates adequate world knowledge.

According to heuristic-dependent text encoding, every reader may form a veridical memory representation of the meaning of a text. Consequently, the text memory of HK and LK subjects may be rather similar. However, the ability to learn from a text is assumed to critically depend upon a reader's prior knowledge. In contrast to LK subjects, HK subjects may therefore acquire general domain knowledge by studying a text.

Experiment

Subjects who did not have any knowledge about computer programming (LK subjects) and subjects who had previous programming experience (high knowledge or HK subjects) read parts of a LISP programmer's manual. Two different groups of LK subjects and two different groups of HK subjects were employed. The subjects of one LK group (freshmen) were college freshmen, whereas the other LK subjects (non-progammers) were older and were matched in age and education with the HK subjects. The subjects of one HK group (LISP group) possessed rather specific prior information about the situations referred to by the text. The subjects of the other HK group (programmers) did not have any prior knowledge about LISP, but had general programming knowledge available.

Since the present experiment compares the performance of LK and HK subjects it may be classified as an expert-novice difference study on text processing. At least three methodological problems of expert-novice difference studies are known which require some special attention in the design of the experiment.

1) Because the population of HK subjects is relatively small, only a small number of HK subjects, normally ranging between one and eight, can be employed in an experiment.

2) Performance differences between LK and HK subjects may be accounted for by age, educational and general intellectual differences rather than by domain-specific knowledge differences whose effects we intend to study.

3) LK and HK subjects may apply different strategies, such as studying a text for different lengths of time or processing the test sentences for a different length of time. Because of trade-offs between reading time and text comprehension and because of speed-accuracy trade-offs, the performance measures of LK and HK subjects may consequently not be comparable.

For the present experiment three LISP subjects were available and six subjects were available for each of the other groups. Every subject was therefore required to participate in a series of experimental tasks, so that several components of text comprehension could be examined within a single reading experiment. In addition, each experimental task tested every subject with several different materials, so that the reliability of performance could be assessed relative to the joint variability between subjects and between different test materials.

In order to ensure that group differences between HK and LK subjects are not due to individual differences other than high and low domain knowledge, each subject also studied a text on Greek mythology, similar in structure and difficulty to the text from the LISP manual. No knowledge differences existed among the four subject groups with respect to this mythology text. Two control comparisons are thus available to ensure that differences in performance are indeed due to domain-specific knowledge differences: the experimental measures obtained with the LISP text may be compared across LK and HK subjects and the comprehension performance of the LISP text may be compared with the comprehension performance of the mythology text for each of the four subject groups.

To ensure that LK and HK subjects studied the texts equally extensively, all subjects were required to study the texts for a predetermined time. However, they were free in how long they studied the various parts of the text. Reading time measures could thus be obtained from the first reading of each paragraph. For evaluating text memory and world knowledge a tapping speed accuracy trade-off method (Wickelgren, Corbett, & Dosher, 1980) was employed instead of simply collecting true-false responses, so that possible differences in the allocation of processing time could not affect the assessment of differences between LK and HK subjects. In the present paper the performance of LK and HK subjects will be evaluated by comparing the results at that part of the speed-accuracy trade-off curve where all subject groups had reached an asymptote in their performance. A more detailed analysis of the speed-accuracy data, including Wickelgren et al's incremental d'-analysis, has been reported in Schmalhofer (1982).

Method

<u>Subjects</u>. Twenty-one University of Colorado students, with uncorrected vision, participated in this experiment. These subjects were recruited from four different subject populations, forming four different subject groups (freshmen, non-programmers, programmers, and LISP group). The six subjects from an introductory psychology class formed the freshmen group, whereas the six subjects of the non-programmer group were obtained from an upper division psychology course. These twelve subjects of the freshmen and non-programmer group were low knowledge (LK) subjects without any training in computer programming. Subjects for the programmer group were recruited from upper division computer science courses. These subjects had been enrolled in 2 or 3 computer science courses (average 2.8) and knew between 2 and 5 different programming languages (average 3.2). Three LISP trained subjects (LISP group) were also recruited from upper division computer science courses. These subjects, in contrast to the programmer group, knew the programming language LISP before the experiment. They had written at least two programs in LISP for class assignments and knew between 2 and 5 (average 3.7) different programming languages.

The six subjects in the freshmen group participated in order to fulfill a course experiment, whereas the remaining 15 subjects were paid $16.00 for their participation in this four hour experiment.

Materials. The two expository texts used in this experiment were taken from the first pages of McCarthy, Abrahams, Edwards, Hart, and Levin's (1965) LISP 1.5 Programmer's Manual and from the first pages of Hamilton's (1940) text on Greek mythology. Some editing changes were made to shorten the texts while preserving their coherence. The two texts were divided into several pages. A page of text, which consisted of an average of 58 words, could fit on the TV screen used in this experiment. A page coincided with a paragraph of the text. For every page, a title describing the topic of the paragraph was introduced.

Each of the two texts was subdivided into two parts. Part 1 of the mythology text consisted of 4 pages with a total of 17 sentences, 133 propositions and 305 words. Part 1 of the LISP text consisted of 11 pages with a total of 37 sentences, 257 propositions and 621 words. Every formula was counted as a word.The first part of the mythology text described the relation between the gods, and the universe whereas the LISP text introduced atomic symbols and S-expressions. Sample paragraphs from the first parts of the mythology and the LISP texts are shown in Table 1.

Insert Table 1 about here

Part 2 of the mythology and LISP text consisted of 8 and 12 paragraphs with 508 and 600 words respectively. The two texts had a similar structure: The LISP text discussed the four LISP functions cons, car, cdr, and equ and the mythology text discussed the four gods Hera, Poseidon, Hades and Pallas Athena. While the LISP text specified the number and types of arguments of a given LISP function, the mythology text described favorite animals, cities or trees of a god or goddess. Two pages of text were presented for each LISP function as well as for each Greek god. However, in contrast to the mythology text, for each of the LISP functions program examples were presented on the second page. Sample paragraphs of the second part of the mythology and the LISP text are shown in Table 2.

Insert Table 2 about here

Test Stimuli for Surface Memory. Surface memory was tested by a true-false test which used 12 sentences or main clauses selected from the first part of each text. For each selected sentence, three different versions were prepared: a sentence underwent a formal change, a correctness change or remained identical to the sentence which occurred in the text. The 36 sentences were then divided into three different experimental sets with four sentences of each type. Subjects were evenly distributed across the three experimental sets.

Surface memory for LISP expressions was similarly evaluated by twelve S-expressions which were presented in the first part of the LISP text. These expressions had been constructed so that they also had a meaning in English. For example, "ORANGE" was used as an atomic symbol, and "(V.W)" as an S-expression. Half of the examples presented in the text were correct whereas the remaining six were examples of syntactically incorrect S-expressions. For every presented S-expression two distractor expressions were constructed by either changing the meaning of the expression with respect to English (formal change), or changing the correctness of the expression with respect to LISP syntax (correctness change). Sample stimuli of the verbatim memory task are shown in Table 3.

Insert Table 3 about here

Sentence Verification - Mythology Text. From the second part of the mythology text 64 sentences were extracted or constructed. Each of these sentences stated a characteristic or a family relationship of a god which was described in the text. Half of the test sentences were true and the other half were false. These false (distractor) sentences were constructed by attributing the characteristic of a given god to another god.

Test sentences could either be explicit or implicit. Explicit sentences were sentences which had the name of a God (Hera, Poseidon, Hades, or Pallas Athena) as their subject and occurred in the text, with the exception that pronouns were replaced by their proper nouns. Implicit sentences were constructed by replacing a god's name with an alternative description which occurred in the text. For example, "Hades" was substituted by "King of the Dead". These 64 test sentences were presented to every subject twice. The order of presentation was randomized each time.

Sentence Verification - LISP text. Thirty-two sentences and 32 evaluated LISP programs were extracted from the second part of the LISP text. Whenever possible these items were constructed completely analogous to the 64 test sentences of the mythology text. Thus half of the sentences and half of the programs were true and the other half false. The false (distractor) items were obtained by exchanging corresponding parts of two correct sentences or formulae.

Half of the sentences and half of the LISP programs were explicitly presented in the text whereas the other half could be inferred from the text (implicit). A false sentence or program was considered as an explicitly presented item, if a "directly" contradicting sentence occurred in the text. For example, the incorrect sentence "the function cons has one argument" is considered an explicit sentence, because "the function cons has two arguments" was presented in the text. Table 4 shows sample test stimuli for examining memory for meaning.

Insert Table 4 about here

Apparatus. Text materials, surface memory and comprehension tests were presented on a Ball Brothers Television Monitor which was under the control of a PDP 11/03 microcomputer. This microprocessor also collected various responses from each subject. A voice key and two response buttons were available to the subject for controlling the presentation of a text. By operating the voice key and the response buttons the subject could select any page for reading at any time.

During reading, an Applied Sciences Eye View Monitor Model 1996 sampled the right eye's focus point 60 times every second through a TV camera. In the present experimental set-up, the eye monitor determined the eye's fixation position with an accuracy of plus or minus two characters (Kliegl & Olson, 1981).

In the speed accuracy task (Wickelgren, Corbett & Dosher, 1980) 1000 msec long response signals were presented every 2 seconds over earphones and the subjects' button presses, their latencies and the duration of the button presses were recorded by the PDP 11/03.

Procedure. Every subject participated individually in a one hour practice and a three hour test session on two different days. On the first day the subject practiced the speed-accuracy trade-off procedure of the sentence verification task for about 40 minutes and was familiarized with the procedure used for recording eye-movements.

The three hour test session began with the presentation of the first part of the mythology text. Before every presentation of a text, the subject had to fixate on a nine point grid, which was used to convert the output of the eye-tracker to character positions on the screen. Six minutes were allocated for studying this text, which was four times the average reading time of technical texts (Just & Carpenter, 1980). The subject was not restricted to reading or rereading the text in any particular sequence, but controlled the display of the text with button presses and a voice key.

After an interfering task where the subject had to write a 60 word summary, surface memory was tested by a speed accuracy trade-off procedure. This procedure, which was modeled after the tapping speed-accuracy trade-off paradigm developed by Wickelgren, Corbett, & Dosher (1980), collected a yes-no decision together with a confidence rating every two seconds. Depending upon which of the two buttons was pressed a yes or no response was recorded and the duration of the button press was used to score the subject's confidence in his response.

The subjects started every single trial by pressing a response button. After a two second delay the first tone signal appeared. This tone signal preceeded the stimulus by one second. The subject was, nevertheless, required to respond: At this time he could only make a guess about the stimulus which would appear on the screen one second later. One second after the stimulus was presented, five more tone signals were presented in two-second intervals. Every tone signal was answered by the subject with a yes-no decision. At the end of each trial the subjects were given feedback about how well they had conformed to the experimental procedure. Subjects were not given any information about the correctness of their responses.

Upon completion of the test for verbatim memory, the subject was asked to study the second part of the mythology text for six minutes and that his or her comprehension would be tested afterwards. The subject was then presented with a sentence verification task. Two blocks, of 64 trials each, were presented in a random sequence which changed for every block. The speed accuracy methodology used in this paradigm was identical to the one used in the surface memory task.

The subject was given feedback at the end of every second trial about how well he had conformed to the experimental procedure and about the "average correctness" of his button presses. Both the correctness and the confidence of a button press entered into the calculation of this index. Subjects were told that larger numbers implied better performance and that a negative index would be obtained if most button presses with high confidence ratings were wrong.

Subjects were allowed a ten minute break before studying the LISP text. With the exception that an additional surface memory test was given for LISP expressions, the experimental procedure remained the same. After reading the first part of the LISP text for 13 minutes and after writing a summary of approximately 120 words, subjects were given two surface memory tests. Twelve sentences and twelve LISP expressions were used in this surface memory test. Next, the subject read the second part of the LISP text and performed the verification task. Half of the items presented in this task were sentences, whereas the other half of these items were LISP functions.

<u>Results</u> <u>and</u> <u>Discussion</u>

In all data analyses to be reported, a statistical significance level of .05 has been used, but probability values will also be reported for the convenience of the reader.

In order to examine whether differences in text processing among the four subject groups are due to domain-specific knowledge about computer programming or other characteristics of the four subject groups, all dependent measures obtained with the mythology text, the LISP text, and the LISP programs were entered into three separate discriminant analyses. A detailed description of these measures is presented in Schmalhofer (1982). For the mythology text, no reliable discriminant function existed for differentiating among the four subject groups, chi**2(45)=35.74, p=.84. However, exactly one significant discriminant function existed for the measures of the LISP text, chi**2(39)=75.9, p=.0004, as well as for the measures obtained with the LISP programs, chi**2(27)=61.9, p=.0002.

These results exclude the interpretation that group differences other than knowledge about computer programming are responsible for differences in the comprehension of the LISP text and LISP programs. Therefore, it is worthwhile to test more specific hypotheses by inspecting the various dependent measures separately. Text and knowledge representations, which are the cognitive products of comprehension processes, will be tested first. Afterwards on-line reading measures will be used to examine the comprehension processes themselves.

Text and Knowledge Representations. Theories of heuristic- and knowledge-dependent text encoding differ in their predictions about LK and HK subjects' memory for meaning. If memory for meaning is a by-product of general heuristic comprehension processes, HK and LK subjects may show good propositional memory. However, if detailed world knowledge is a prerequisite for the encoding of the text, LK subjects should show poor memory for meaning. If LK subjects' memory for meaning is indeed poor, one might suppose that in order to compensate for this shortcoming LK subjects would engage more actively in literal encoding and may therefore show better verbatim memory than HK subjects.

Three different types of test sentences were used to examine verbatim memory for the LISP and mythology texts. Test sentences could be identical to the sentences in the text or they could undergo a formal change or a meaning change. Table 5 shows the average nine-point ratings, which were obtained from the duration of the subject's button presses, for the verbatim memory task of the mythology and the LISP texts. These nine-point ratings represent the subject's decision and his confidence in this decision: A rating of 1 means the sentence certainly occurred in the text, whereas a rating of 9 indicates that the sentence did certainly not occur in the text. Table 5 shows that both texts yielded rather similar results for LK and HK subjects. All four subject groups showed low confidence ratings for sentences whose correctness was changed and high confidence ratings for sentences which had literally occurred in the text (identical). For the LISP text HK subjects were somewhat more confident in rejecting correctness changes and more confident that identical test sentences had occurred in the text. In general, formal changes were not recognized. The average confidence rating for formal changes is therefore almost as high as the average confidence rating for identical sentences. No significant group differences existed for any type of test sentences. For the mythology text the analyses yielded for correctness changes, $F(3,77)=1.87$, $MS(e)=7.40$, $p=.14$; for formal changes, $F(3,78)=1.02$, $MS(e)=7.96$, $p=.39$; and for identical, $F(3,76)=.84$, $MS(e)=7.10$, $p=.48$. For the LISP text the analyses yielded for correctness changes, $F(3,79)=1.46$, $MS(e)=7.72$, $p=.23$; for formal changes, $F(3,80)=.29$, $MS(e)=7.40$, $p=.83$; and for identical, $F(3,79)=.91$, $MS(e)=7.24$, $p=.44$.

Insert Table 5 about here

Apparently HK as well as LK subjects based their confidence ratings on the meaning rather than the surface characteristics of the sentences. Even LK subjects relied upon a propositional rather than a verbatim encoding of the text. Memory for the surface characteristics was poor because unless a pragmatically significant surface structure motivates the reader to attend to particular wordings (Bates, Masling & Kintsch; 1978), verbatim representations exist only temporarily in memory. The present texts did not have such a pragmatically interesting surface structure.

In this verbatim memory task the subjects were asked to decide whether or not a sentence had literally occurred in the text. A more direct examination of memory for meaning is given by the verification task in which subjects had to verify the truth of sentences as opposed to judging whether sentences had literally occurred in the text. These sentences had either explicitly occurred in the text or had to be inferred from the text. Explicitly presented true sentences will be analyzed first because these sentences are a direct measure of text memory, whereas a situational model may be required for verifying implicit or false sentences.

Insert Figure 1 about here

The upper two curves of Figure 1 show the average ratings of the four subject groups for true sentences, for each text. A significant group difference was found for the LISP text, $F(3,278)=4.40$ $p=.0047$ $MS(e)=6.61$ $eta=.21$, as well as for the mythology text, $F(3,240)=4.25$ $p=.0060$ $MS(e)=4.60$ $eta=.23$.

However, Figure 1 also shows that the average ratings of the four groups are similar for the two texts. Therefore significant group differences may be caused by general differences among the four subject groups rather than by domain specific knowledge about programming or LISP. Since, for both texts, group differences account for five percent of the variance, the verification of LISP sentences does not differentiate the four groups any better than the verification of mythology sentences.

Also, for the LISP text the largest group difference occurred between the freshmen and the other three groups rather than between HK and LK subjects. When the freshmen group was excluded from the analysis, no significant difference was found among the three remaining subject groups for the LISP text, $F(2,193)=0.21$ $p=.82$ $MS(e)=6.11$ $eta=.046$. Therefore it may be suspected that for true and explicitly presented sentences text memory is good even when a reader has only low domain knowledge. Analyses of the subjects' text summaries also supported this hypothesis.

Because a reader cannot infer all possible inferable correct and incorrect test sentences during the encoding of the text, a situational model may be required for successfully verifying sentences which were not explicitly presented in the text. The lower two curves of Figure 1 show the average ratings of explicit false sentences for the four subject groups and the two texts. Whereas no significant group differences were found with the explicit false mythology test sentences, $F(3,272)=1.10$ $p=.3478$ $MS(e)=5.35$ $eta=.1097$, reliable differences between HK and LK subjects existed for the LISP test sentences, $F(3,370)=13.21$ $p<.00001$ $MS(e)=8.33$ $eta=.31$. For the LISP sentences ten percent of the variance was accounted for by these group differences whereas for the mythology sentences only one percent of the variance was accounted for by group differences.

Similarly, a situational model may also be required for verifying implicit sentences. Figure 2 shows the average ratings for true and false implicit mythology sentences. As Figure 2 suggests, significant group differences did not exist for true mythology sentences, $F(3,270)=1.03$ $p=.38$ $MS(e)=7.47$ $eta=.1063$, or false mythology sentences, $F(3,283)=1.74$ $p=.158$ $MS(e)=8.53$ $eta=.1347$.

Insert Figure 2 about here

However, for the ratings of the LISP sentences, which are shown in Figure 3, significant group differences existed for true as well as for false sentences, $F(3,287)=8.50$ $p<.00001$ $MS(e)=8.74$ $eta=.2857$ and $F(3,221)=8.10$ $p<.00001$ $MS(e)=8.81$ $eta=.3147$, respectively. Again, each of these effects accounted for approximately 10 percent of the variance.

Insert Figure 3 about here

It has thus been demonstrated that domain-specific knowledge about computer programming has an important influence upon the verification of false and implicit sentences. Apparently, LK subjects can construct a representation of the meaning of a text whereas HK subjects, in addition, derive a general model about the programming language LISP. This hypothesis may be further supported by demonstrating that HK subjects are better in verifying small LISP programs. In order to determine whether group differences exist in the verification of the four different types of LISP programs, a separate ANOVA was performed upon the subjects' nine-point ratings. Significant group differences existed for all four program types: explicit true, $F(3,283)=10.18$ $MS(e)=4.1$ $p<.0001$; explicit false, $F(3,287)=6.14$ $MS(e)=7.4$ $p=.0005$; implicit true, $F(3,281)=5.27$ $MS(e)=4.6$ $p=.0015$; implicit false, $F(3,277)=6.42$ $MS(e)=4.1$ $p=.0003$.

These performance differences between HK and LK subjects demonstrate that while the textbase depends upon general comprehension processes, the development of a situational model relies upon a reader's domain-specific knowledge. Whereas HK subjects developed a situational model which allowed them to verify any statement about the programming language LISP, LK subjects had only constructed a textbase. The information stored in a textbase is specific to the presented text whereas a situational model about LISP represents a general understanding of the programming language LISP itself. The structure of the mental model reflects the properties of the programming language LISP, whereas the text base reflects the characteristics of a particular text.

The mental model of the programming language LISP allowed the HK subjects to become successful programmers in LISP, whereas LK subjects did not acquire any programming skills from studying the LISP manual. High and low domain knowledge may thus determine whether a student can develop a situational model and learn a new skill whereas the influence of domain knowledge upon text encoding and text memory may only be minor. Although the analyses of the verification tasks indicate that LK subjects constructed a text base and HK subjects had in addition developed a situational model about the programming language LISP, the processes involved in forming a text base and constructing a situational model have not as yet been examined.

<u>Cognitive</u> <u>Processes</u> <u>of</u> <u>Text</u> <u>Encoding</u>. Theories of knowledge- and heuristic-dependent text encoding also differ in their predictions about sentence reading times, reflecting different hypotheses about the cognitive processes of text encoding. If text processing is knowledge-driven, early sentences of a text will require more processing than late sentences. Because the selection and activation of a knowledge structure, which occurs early in a text, requires additional processing time and because once a knowledge structure has been selected, it will then faciliate the encoding of the following sentences (Bower et al., 1979), in knowledge-driven encoding reading times decrease with the serial position of a sentence (Cirilo & Foss, 1980).

Theories of heuristic-dependent text memory, on the other hand, predict that every proposition requires some processing time. In heuristic-driven text encoding, sentence reading times therefore increase with the number of propositions of a sentence rather than with its serial position in a text. If text encoding critically depends upon a reader's previously acquired domain-specific knowledge, reading times of HK subjects may also differ from the LK subjects' reading times. However, if LK and HK subjects apply similar heuristic comprehension processes no group-differences may be expected between the reading times of LK and HK subjects.

The average reading time per character was chosen as a measure of comprehension time in all reading time analyses, in order to eliminate the effects of perceptual encoding. Since the number of characters in a sentence was highly correlated with the number of words ($r=.97$ for each text), this measure was also systematically related to the average reading time of a word. In order to evaluate whether differences in cognitive processing exist among the four subject groups and between the two texts, a text by subject group analysis of variance was performed. For every subject the average reading time per character of the sentences from the first part of each text was entered as a separate observation into the analysis.

Although both main effects were significant, $F(3,319)=5.45$ $p=.0011$ $MS(e)=.00147$, $F(1,319)=9.90$ $p=.0018$ $MS(e)=.00113$, while their interaction wasn't, $F(3,319)=2.15$ $p=.0944$ $MS(e)=.00113$, Tukey's method of multiple comparisons showed that only the LISP group read the LISP text reliably faster than the mythology text, ($q=4.73$). No reliable differences between the two texts were found for any other subject group, $q<1.06$, and there were also no reliable differences in the reading times of the mythology text among the four groups, $q<2.87$. Similarly for the freshmen, non-programmer and programmer groups, no reliable differences were found in the reading times of the LISP text, $q<2.34$. However, the LISP group, although not significantly different from the programmer group, $q=3.42$, differed from the freshmen, $q=4.22$, and the non-programmer group, $q=5.64$. Thus with the exception of the LISP subjects who read the LISP text faster than the mythology text, differences exist neither among the subject groups nor between texts. The average reading times from the second part of the mythology and the LISP texts are shown in Figure 4. Except for the programmers, who read the second part of the LISP text faster than the second part of the mythology text, the results obtained from the second parts of each text replicated the described results.

Insert Figure 4 about here

Since the mythology text was selected to be comprehensible for all subjects of the experiment, we may assume that these subjects processed the mythology text as any other expository text for which they are intended readers. The comparison of the reading measures for the

LISP and the mythology text showed that the reading times are identical for the two texts when the subjects have not previously acquired any programming or LISP knowledge, whereas the reading times of HK subjects are faster for the LISP text than for the mythology text. HK subjects might have read the LISP text faster because they knew already much of the information presented in this text.

The analysis of average reading times thus seems to indicate that even without domain-specific knowledge a technical text can be processed like a normal text. This hypothesis was further evaluated by examining the relations between sentence reading times, its serial position and the number of propositions in a sentence.

In order to determine whether reading times increase or decrease with the serial position of a sentence in a text, regression estimates of reading times on the serial position were calculated. Spearman's correlation coefficients indicated that no systematic trend existed between the serial position of a sentence and its reading time for any of the subject groups or texts ($p>0.05$).

Regression estimates were also calculated under the hypotheses that reading times increase or decrease only monotonically with the serial position of a sentence (Barlow, Bartholomew, Bremner & Brunk, 1972). In order to determine which hypothesis is more adequate these estimates were then compared to the observed reading times by calculating chi-square deviation values. Although the distribution of this statistic is unknown, it may help to distinguish the two conflicting hypotheses. For the hypothesis of increasing reading times, chi-square values of 45.3 and 40.2 (summed over all subject groups) were obtained for the mythology and the LISP text, respectively, whereas for the hypothesis of decreasing reading times values of 46.2 and 67.0 were calculated. Since the calculated values do not differ much for the two alternative hypotheses, and since there was no significant correlation, it is argued that the average reading times do not show any monotonic serial position effect. Nethertheless, if a serial position effect is assumed a priori, the obtained result slightly favors the hypothesis that reading times increase with serial position. This result was also replicated by analyzing serial position effects within paragraphs.

These results can be explained by assuming that all subjects constructed propositional representations for the mythology text as well as for the LISP text. Because integrating new propositions into a partially constructed text base becomes more difficult and more reinstatement searches are required when this text base grows, reading times may increase with the serial position of a sentence rather than decrease. Although theories of heuristic-dependent text encoding may predict this relationship, a more critical hypothesis of these theories is that reading times should depend upon the number of propositions in a sentence (Kintsch, 1974).

The average reading time per character as a function of the number of propositions in the sentence was analyzed only for the LISP text, because in the mythology text the number of characters was highly correlated with the number of propositions, $r=.93$. For the LISP text this correlation was only moderate, $r=.53$, and the correlation between serial position and number of propositions of a sentence was .04. When only the sentences of the LISP text, which did not contain any LISP expressions, were included in the analysis, the Spearman correlation coefficients indicated a positive correlation between the number of propositions and the reading time of a sentence: freshmen $r(85)=.30$ $p=.0057$; non-programmers $r(102)=.19$ $p=.0512$; programmers $r(85)=.25$ $p=.0208$; LISP group $r(51)=.14$ $p=.3130$. For the freshmen group and the programmer group these correlation coefficients were found to be significant. When all sentences of the first part of the LISP text (now including LISP expressions) were entered into the analysis a reliable positive correlation was found for all four subject groups: freshmen $r(185)=.43$ $p=.0001$; non-programmers $r(222)=.37$ $p=.0001$; programmers $r(185)=.44$ $p=.0001$; LISP group $r(111)=.49$ $p=.0001$.

Regression estimates were also calculated under the assumption that reading times increase only monotonically with the number of propositions in a sentence. The results of this calculation are shown in Figure 5. This Figure shows that the average reading time per character increases in a similar way, though at different rates, for all four subject groups as a function of the number of propositions in a sentence.

Insert Figure 5 about here

The analysis of reading times thus indicated that all four subject groups processed the LISP text by its propositional representation. Even the LK subjects apparently executed heuristic comprehension processes which yielded as a by-product propositional text memory (text base), rather than encoding the LISP text by its surface characteristics.

The processes involved in constructing a text base were further examined by regressive eye-movements. Regressive eye-movements may occur when a reader cannot attach some new proposition to the text base which is being constructed. When a reader fails to attach a new proposition to the propositions which are held in working memory, he must search long term memory and/or external memory, i. e. the written text, to find a proposition which is related to the new incoming proposition. These search processes may be reflected or assisted by regressive eye-movements, which search the externally presented text for an appropriate proposition. Since LISP subjects can rely upon their previously acquired LISP knowledge, the LISP subjects should show fewer regressive eye-movements for the LISP text than for the mythology text. However, for LK subjects, the number of regressive eye-movements should be approximately equal for both texts, because similar comprehension processes are assumed for both texts. The results of this analysis are shown in Figure 6.

Insert Figure 6 about here

For the freshmen, non-programmers, and programmers, the average number of regressive eye-movements in a sentence was about equal for both texts, whereas the LISP group showed significantly fewer regressive eye-movements for the LISP text than for the mythology text, $t(89)=2.67$ $p=.009$. As expected, complete and detailed knowledge about LISP reduced the number of regressive eye-movements, which may be an indicator of reinstatement searches occurring during the construction of a text base. However, HK subjects who had only some superordinate knowledge about the subject area of the text (programmers) showed the same pattern of regressive eye-movements as LK subjects. It thus appears that moderate domain-knowledge does not facilitate the construction of a text base, whereas readers who possess the knowledge expressed by the text beforehand (LISP group) may have applied quite different reading strategies.

The more interesting finding, however, is that the LK subjects processed the LISP text by its meaning, similar to the programmers (HK subjects). It thus appears that with respect to the encoding of the LISP text, LK subjects execute similar micro- and macroprocesses as HK subjects. For all four subject groups reading times increased with the number of propositions of a sentence, indicating the construction of a propositional text base. Only those HK subjects who had previously acquired detailed knowledge about LISP, showed fewer regressive eye movements. Because text memory is a by-product of general comprehension processes which may be executed by a reader with and without domain-specific knowledge, HK as well as LK subjects may build a veridical representation of the meaning of a text in memory. Contrary to HK subjects, however, LK subjects were unable to acquire a general understanding of the programming language LISP through studying the text.

These results demonstrate that text and knowledge representations are two separate structures which may exist in memory independent from each other. If a text's meaning units were stored in selected slots of a previously acquired knowledge structure such as a script or schema, LK subjects would have poor memory for the meaning of a text. Therefore, it is concluded that a text representation is created in memory as a by-product of general comprehension processes, which may even be executed in the absence of domain-specific knowledge. The construction of a situational model, on the other hand, critically depends upon a reader's world knowledge.

The present experiment also examined the cognitive processes which are executed during the construction of a situational model. Since previous research has emphasized text memory rather than learning from texts the predictions about the cognitive processes of developing a situational model are necessarily less specific.

Cognitive Processes of Acquiring Knowledge from Texts. The study times of example programs of HK and LK subjects may help to assess the importance of learning from examples for the development of a situational model. Whereas LK subjects might attempt to encode example LISP programs similarly to words or sentences in English, HK subjects might know about the importance of learning from examples (Simon, 1980) and therefore process them differently.

The average study time per character of example LISP programs (including explanation of the example) in relation to the LISP text and the mythology text were shown in Figure 4. All four subject groups processed the LISP programs significantly slower than the LISP text: freshmen $t(30)=7.01$ $p<.00001$; non-programmers $t(46)=4.13$ $p=.0002$; programmers $t(46)=2.24$ $p=.0302$; and LISP group $t(22)=3.61$ $p=.0015$. In addition, reliable differences in the study times of LISP programs were observed among the four subject groups, $F(3,72)=7.91$ $p=.0001$ $MS(e)=.00183$. Figure 4 shows that the freshmen and non-programmers processed the LISP programs slower than the programmers and LISP group.

Since LK subjects took more time for processing example LISP programs than HK subjects, it may be conjectured that either they are aware of the importance of examples or they processed the example programs in a very inefficient way, requiring long study times. Since our previous analyses have already determined that the HK subjects developed a better understanding of the programming language LISP, the inefficient processing explanation appears to be more appropriate. Furthermore, if it is assumed that cognitive processes generated a related memory trace as a by-product of this processing, it might be possible to determine some characteristics of these information processes by examining their memory trace.

An evaluation of the memory for meaning has already been reported. Verbatim memory of LISP expressions is another memory trace which may yield important information about the underlying cognitive processes. Every LISP expression presented in the text built a meaningful word or expression in English and in addition was either correct or incorrect with respect to LISP syntax. By either changing the meaning with respect to English (formal change) or the correctness with respect to LISP syntax (correctness change), two different distractor "LISP expressions" were constructed.

A one-way ANOVA was performed upon the nine-point verbatim memory ratings for each of the three different types. For the LISP expressions which were identical to the LISP expressions presented in the text, no significant group differences were found, $F(3,79)=1.14$ $MS(e)=7.9$ $p=.3393$. Also no group differences were found for LISP expressions whose correctness with respect to LISP syntax was changed, $F(3,79)=.33$ $MS(e)=7.6$ $p=.8026$.

Only LISP expressions whose meaning (in English) was changed yielded significant group differences, $F(3,78)=5.65$ $MS(e)=6.6$ $p=.0015$. Thus at least some subjects encoded the English meaning of the LISP expressions in memory. Table 6 shows the average verbatim memory ratings for the four subject groups and the three different types of test items of verbatim memory. Whereas the freshmen and the non-programmers correctly rejected changes of the meaning of a LISP expression, the LISP group did not identify these distractors as accurately and the programmers did not perform better than chance. It thus appears that LK subjects may process LISP expressions similarly to words in English, whereas HK subjects may process the LISP expressions in a more computer-oriented way.

Insert Table 6 about here

Although the analyses of study time and verbatim memory of LISP expressions do not provide enough evidence in order to decide about the significance of learning from examples upon the construction of a situational model, important processing differences were found between HK and LK subjects. Whereas LK subjects showed memory traces of encoding processes for English words, the HK subjects' verbatim memory for LISP expressions did not show any processing traces, neither with respect to English words not with respect to LISP syntax. LK subjects also studied LISP programs significantly longer than HK subjects. The previously developed hypothesis that LK subjects process a technical text like normal English prose, whereas HK subjects attempt to develop a situational model about the technical subject domain, is consistent with these results.

LK subjects processed the LISP text by general comprehension processes, which develop a text base. Since LK subjects relied upon general comprehension processes and did not derive a situational model of the programming language LISP, they even processed LISP expressions and LISP programs like English words. Since LISP programs and LISP expressions do not adhere to the rules of English grammar, the processing of LISP programs required more time than the processing of sentences of the LISP text. As a memory trace of this processing LK subjects recognized the "English meaning" of LISP expressions. HK subjects did not show such memory traces because they processed the program examples differently. Rather than processing LISP expressions by their meaning in English, HK subjects developed a general model about the programming language LISP. Since HK subjects spent relatively little time studying program examples and since no memory traces were found for the specific example programs presented in the text, we may assume that they derived a situational model more from the general explanations of the text than by learning from the program examples presented in the text. Based upon some informal reports of HK subjects, we may speculate that HK subjects used the program examples to test their understanding of the general explanation of the LISP text.

Conclusions

Theories of knowledge-driven text encoding (script, schema theories) postulate that a text's meaning units are stored in the slots of a reader's hierarchically structured domain knowledge (e.g. about a restaurant). Because for storing the text's meaning units no slots are available to readers without domain-specific knowledge, these LK subjects should consequently show poor memory for meaning. However, the present study shows that LK like HK subjects have poor verbatim memory but good memory for the meaning of a text. This result contradicts theories of knowledge-driven text encoding but supports theories of heuristic-dependent text encoding. These theories assume that similar to episodic and semantic information (Tulving, 1972), text base and knowledge representations (situational model) are separate entities in memory. A representation of the meaning of a text is constructed in memory as a by-product of general comprehension heuristics such as micro- and macroprocesses, whereas the development of a situational model depends upon a reader's knowledge. Since general comprehension heuristics are domain-independent, they can even be executed by LK subjects. In support of this hypothesis, the present study found that similar to HK subjects even LK subjects formed a propositional text base. HK as well as LK subjects' sentence reading times increased with the number of propositions in a sentence further validating the execution of general comprehension processes.

Previous experiments which demonstrated a severe deterioration of memory when a reader could not activate relevant domain knowledge in memory (Bransford & Johnson, 1973) employed texts which were developed for the particular purposes of these experiments. However, these texts may have violated some fundamental characteristics of natural texts such as text coherence. Contrary to the results obtained with pseudo-texts in which the writer did not seriously attempt to communicate information to the reader (Bransford & Johnson, 1973; Collins, Brown & Larkin, 1977), the present study showed that for real texts general comprehension processes can even be executed in the absence of domain-specific knowledge and that text memory is a by-product of these processes.

When readers cannot construct a text base for pseudo-texts they may visualize some real world situation in order to make any sense out of the presented pseudo-text. Readers may therefore rely upon some situational model, which may or may not be adequate, for interpreting the presented pseudo-text. Verbal protocols collected by Collins, Brown & Larkin (1977) showed how readers selected a wrong situational model for interpreting the following pseudo-text.

> "He plunked down $ 5 at the window. She tried to give him $ 2.50, but he refused to take it. So when they got inside, she bought him a large bag of popcorn."

These sentences with unresolved pronomial references supposedly describe how a guy treats his date to a movie ticket and she buys him popcorn in return. One of Collin et al.'s subjects used the situational model of a racetrack for interpreting these sentences: at a racetrack a guy was giving $ 5 to a lady behind a betting window, who returned $ 2.50 as the change for the betting ticket. Only when reading the last two sentences of this pseudo-text did the subject realize that his situational model was inconsistent with the presented sentences.

Obviously, the application of the racetrack model was contingent upon the reader's domain knowledge. However, the interaction of a reader's domain knowledge and his text encoding is not caused by text and knowledge elements being stored in one hierarchical knowledge structure but by the reader's processing strategy. A reader may utilize a text base and/or a situational model when he is asked to encode or retrieve text information. A reader may decide that he already knows the information presented in the text. This reader will consequently rely upon his previously acquired knowledge (situational model). He will only substantiate the information of his situational model and will therefore read the text faster than a reader who step by step constructs a text base by micro- and macroprocesses. The LISP subjects of the present experiment seemed to use this strategy. Consequently, the LISP subjects read the LISP text faster and made fewer regressive eye-movements. Regressive eye-movements indicate processing difficulties during the construction of a text base. These HK subjects, who had previously learned LISP, relied upon their situational model and did therefore not engage in constructing a text base to the same degree as LK subjects. The dependency of reading times upon the number of propositions in a sentence was therefore weakest for LISP subjects and strongest for LK subjects. Because LISP subjects knew the topic of the text beforehand, they did not require any extra processing time for instantiating their LISP knowledge when reading the first sentences of the text, so that no serial position effect was found in their sentence reading times.

When a reader instantiates relevant knowledge during the reading of the first sentences of a text, serial position effects are observed in the reading of the text, because instantiated knowledge primes the processing of later text units (Bower et al.; Cirilo & Foss). The LK subjects of the present experiment could not activate relevant knowledge during the reading of the first sentences, so that the reading times of sentences did not depend upon its serial position but upon the number of propositions in a sentence.

When a reader is tested, he may again apply different strategies: he may either utilize the knowledge of his situational model or the information stored in his text base. An experiment by Bransford, Barclay & Franks (1972) showed that for familiar situations, subjects may rely upon a situational model rather than upon propositional text memory. Bransford et al. presented subjects with sentences like a) and b), which differred with respect to their propositional representations, but were identical on account of the underlying situational model.

a) Three turtles rested on a floating log and a fish swam beneath it.

b) Three turtles rested on a floating log and a fish swam beneath them.

Since subjects neither rejected sentence a) nor sentence b) as a new sentence, independent of whether they were presented with sentence a) or b) at study time, they must have based their decisions upon a situational model rather than a propositional representation of the encoded sentence. The experiment by Bransford et al. demonstrated that a confusion between world and text information occurs when a reader relies upon a situational model rather than the text base itself. For example, Bower et al.'s subjects confused world knowledge about a doctor's office with text information because they relied upon the situational model of a doctor's office in the same way as Bransford et al.'s subjects relied upon a situational model of a fish swimming below a log with turtles.

Previous studies of differences in domain-knowledge (Chiesi, Spilich, & Voss, 1979; Spilich et al., 1979) did not clearly separate between the reader's information derived from the presented text and a reader's previously acquired knowledge. In these studies, HK subjects' better "text memory" may thus be due to their previously acquired knowledge per se rather than to superior text processing.

In the present experiment the domain knowledge stored in a situational model and the text's meaning units were examined independent from each other. The experimental results contradicted the assumption that a text's meaning units are stored in the slots of a previously acquired knowledge structure but supported a model which assumed that text and knowledge representations are separate (Johnson-Laird, 1980; van Dijk & Kintsch, 1983). It was concluded that memory for the meaning of a text is a by-product of general comprehension processes whereas the development and updating of a situational model depends upon a reader's knowledge.

Reference Notes

Collins, A., Brown, J. S.,and Larkin, K.M., Inference in text understanding. University of Illinois at Urbana-Champaign, 1977, Report No. 3684.

Kieras, D.E., Abstracting main ideas from technical prose: a preliminary study of six passages. Technical Report No.5, University of Arizona, 1980.

Kieras, D.E., Thematic processes in the comprehension of technical prose: final report. Technical Report No.UARZ/DP/TR-82 ONR-10. University of Arizona, 1982.

Schmalhofer, F. Comprehension of a Technical Text as a Function of expertise. Dissertation. University of Colorado, 1982.

References

Anderson, J.R. Language, memory and thought. Hillsdale, New Jersey: Erlbaum, 1977.

Barlow, R.E., Bartholomew, D.J., Bremner, J.M., and Brunk, H.D., Statistical inference under order restrictions: the theory and application of isotonic regression. New York: Wiley, 1972.

Bartlett, F.C., Remembering: A study in experimental and social psychology. Cambridge: University Press, 1932.

Bates, E., Masling, M., and Kintsch, W. Recognition memory for aspects of dialogue. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 187-197.

Bobrow, D.G. and Collins, A. (Eds.) Representation and understanding. New York: Academic Press, 1975.

Bower, G.H., Black, J.B., and Turner, T.J. Scripts in memory for text. Cognitive Psychology, 1979, 11, 177-220.

Bransford J. D., Barclay, J. R., and Franks, J.J., Semantic memory: A constructive versus interpretive approach. Cognitive Psychology, 1972, 3, 193-209.

Bransford J. D., and Johnson, M. K., Considerations of some problems of comprehension. In W.G. Chase (Ed.) Visual information processing. New York: Academic Press, 1973.

Chase, W. G. and Ericcson, K. A. Skilled memory In J.R. Anderson (Ed.) Cognitive skills and their acquisition. New Jersey: Erlbaum, 1981.

Chiesi, H.L., Spilich, G.J., and Voss, J.F. Acquisition of domain-related information in relation to high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 1979, 18 , 257-273.

Cirilo, R.K., and Foss, D.J. Text structure and reading time for sentences. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 96-109.

De Jong, Prediction and substantiation: A new approach to natural language processing. Cognitive Science, 1979, 3, 251-273.

Hamilton, E., Mythology, Timeless Tales of Gods and Heroes New York: The New American Library, Inc., 1940.

Jonson-Laird, P.N. Mental Models in Cognitive Science. Cognitive Science, 1980, 4, 71-111.

Johnson-Laird, P.N. Mental Models. Cambridge: Harvard University Press, in press.

Just, M.J., and Carpenter, P.A. A theory of reading: from eye fixations to comprehension. Psychological Review, 1980, 87, 329-354.

Kieras, D.E., Initial mention as a signal to thematic content in technical passages. Memory and Cognition, 1980, 8, 345-353.

Kintsch, W. On modeling comprehension. Educational Psychologist, 1979, 14, 3-14.

Kintsch, W. Comprehension and memory of text. In W.K. Estes (Ed.), Handbook of learning and cognitive processes. New York: Erlbaum, 1978, vol 6.

Kintsch, W. The representation of meaning in memory. New York: Erlbaum, 1974.

Kintsch, W., and van Dijk, T.A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363-394.

Kliegl, R. Automated and interactive analysis of eye-fixation data in reading. Behavior Research Methods and Instrumentation, 1981, 13, 115-120.

Kliegl, R., and Olson, R. K., Reduction and calibration of eye monitor data. Behavior Research Methods and Instrumentation, 1981, 13, 107-111.

Kosslyn, S.M. Information representation in visual images. Cognitive Psychology, 1975, 7, 341-370.

McCarthy, J., Abrahams, P. W., Edwards, D.J., Hart, T. P., and Levin, M. I., LISP 1.5 Programmer's Manual. Cambridge: The M.I.T. Press, 1965.

McKoon, G., and Ratcliff, R. Priming in item recognition: The organization of propositions in memory for text. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 369-386.

Miller, J.R., and Kintsch, W. Readability and recall of short prose passages: A theoretical analysis. Journal of Experimental Psychology: Human Learning and Memory, 1980, 6, 335-354.

Rumelhart, D.E. Notes on schema for stories. In D.G. Bobrow and A.M. Collins (Eds.), Representation and understanding, New York: Academic Press, 1975.

Schank, R.C. Language and Memory. Cognitive Science, 1980, 4, 243-284.

Schank, R.C. and Abelson, P.R. Scripts, plans, goals, and understanding. Hillsdale, New Jersey: Erlbaum, 1977.

Simon, H.A. Problem solving and education. In D.T. Tuma and F. Reif (Eds) Problem solving and education, issues in teaching and research. New York: Erlbaum, 1980.

Spilich, G.J., Vesonder, G. T., Chiesi, H. L., and Voss, J.F. Text processing of domain-related information for individuals with high and low domain knowledge.

Journal of Verbal Learning and Verbal Behavior, 1979, 18, 275-290.

Tulving E. Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds), Organization and Memory. New York: Academic Press, 1972, 382-404.

van Dijk T.A. and Kintsch, W. Strategies of Discourse Comprehension. New York: Academic Press, 1983.

Wickelgren, W.A., Corbett, A.T. and Dosher, B.A. Priming and retrieval from short-term memory: A speed accuracy trade-off analysis. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 387-404.

Table 1

Sample Paragraphs of the First Part of the Mythology

and LISP Texts

==

Olympus - Paragraph 3

Within Olympus were the gods' dwellings, where they lived and slept and feasted on ambrosia and nectar and listened to Apollo's lyre. It was an abode of perfect blessedness. Homer says no wind ever shakes the untroubled peace of Olympus; no rain ever falls there or snow; but the cloudless firmament stretches around it on all sides and the white glory of sunshine is diffused on its walls.

==

LISP data

In LISP all data are in the form of symbolic expressions usually referred to as S-expressions. S-expressions are either atomic symbols or composite S-expressions. Atomic symbols are the most elementary type of S-expressions. Composite S-expressions are built out of atomic symbols. The rules for building these atomic symbols will be described, next.

==

Table 2

Sample Paragraphs of the Second Part of the Mythology

and the LISP Texts

===

Hades - Paragraph 1

Hades was the third brother among the Olympians, who drew for his share the underworld and the rule over the dead. He was also the God of Wealth, the God of the precious metals hidden in the earth. He had a far-famed cap or helmet which made whoever wore it invisible. It was rare that he left his dark realm to visit Olympus or the earth, nor was he urged to do so. He was not a welcome visitor.

Hades - Paragraph 2

Hades was unpitying, inexorable, but just; a terrible, not an evil god. His wife was Persephone whom he carried away from the earth and made Queen of the Lower World. He was king of the Dead - not Death himself, whom the Greeks called Thanatos.

===

The function car

The function car has one argument and is a function that is used to extract the first S-expression from a larger S-expression. The value of the function car is the first S-expression of its composite argument.The argument of the function car must be a composite S-expression. car of an atomic symbol is undefined.

Examples of the function car

```
Examples of the function car:  car<(A.B)>     =  A
                               car<((A.B).C)> =  (A.B)
                               car<(A.(B.C))> =  A
 The following examples are incorrect because in
    car<A>    =  A          the argument is an atomic symbol;
car<(A.(B.C))> = (A.B)      A is the first S-expression;
car<(A.B);C>  =  A          car has too many arguments.
```

===

Table 3

Sample Stimuli From the Verbatim Memory Task

correctness changes:		
	Mythology:	Homer makes Poseidon say that he rules the dead.
	LISP:	Atomic symbols are a type of composite S-expressions.
	Programs:	((C.B).S)
formal changes:		
	Mythology:	Homer describes Poseidon saying that he he rules the sea.
	LISP:	The most elementary type of S-expressions are atomic symbols.
	Programs:	((C.B).C))
identical:		
	Mythology:	Homer makes Poseidon say that he rules the sea.
	LISP:	Atomic symbols are the most elementary type of S-expressions.
	Programs:	((C.B).S))

Table 4

Sample Stimuli For Testing Memory For Meaning

explicit true:

Mythology: Hades was the god of precious metals.

LISP: The fuction car has one argument.

Programs: car<(A.B)> = A

explicit false:

Mythology: Hades was called "Earth-shaker".

LISP: The function car has two arguments.

Programs: car<(X.Y)> = Y

implicit true:

Mythology: The King of the Dead was a terrible God.

LISP: The value of car may be an atomic symbol.

Programs: car<((A.B).(C.D))> = (A.B)

implicit false:

Mythology: The King of the Dead had some connections with bulls.

LISP: The value of car is always an atomic symbol.

Programs: car<((A.B).(C.D))> = ((A.B).C)

Table 5
Average Nine-Point Ratings of Verbatim Memory for the Three Sentence Types of the LISP Text and the Four Subject Groups

	FRESHMEN	NON-PROGRAMMERS	PROGRAMMERS	LISP-GROUP
Correctness Changes	4.3(3.5)	4.1(4.4)	3.5(3.2)	2.4(2.2)
Formal Changes	5.6(5.5)	5.8(4.3)	6.0(5.3)	6.4(5.8)
Identical	5.8(5.3)	5.9(6.6)	6.9(6.2)	6.7(6.3)

In parentheses the ratings obtained with the mythology text are shown.

Table 6

Average Nine-Point Ratings of LK and HK subjects'

Verbatim Memory for the Three Types of LISP Expressions

	FRESHMEN	NON-PROGRAMMERS	PROGRAMMERS	LISP-GROUP
Change of LISP syntax	5.5	5.1	5.5	4.6
Change of English Meaning	3.8	2.2	5.3	4.1
Identical	5.6	6.4	4.9	5.7

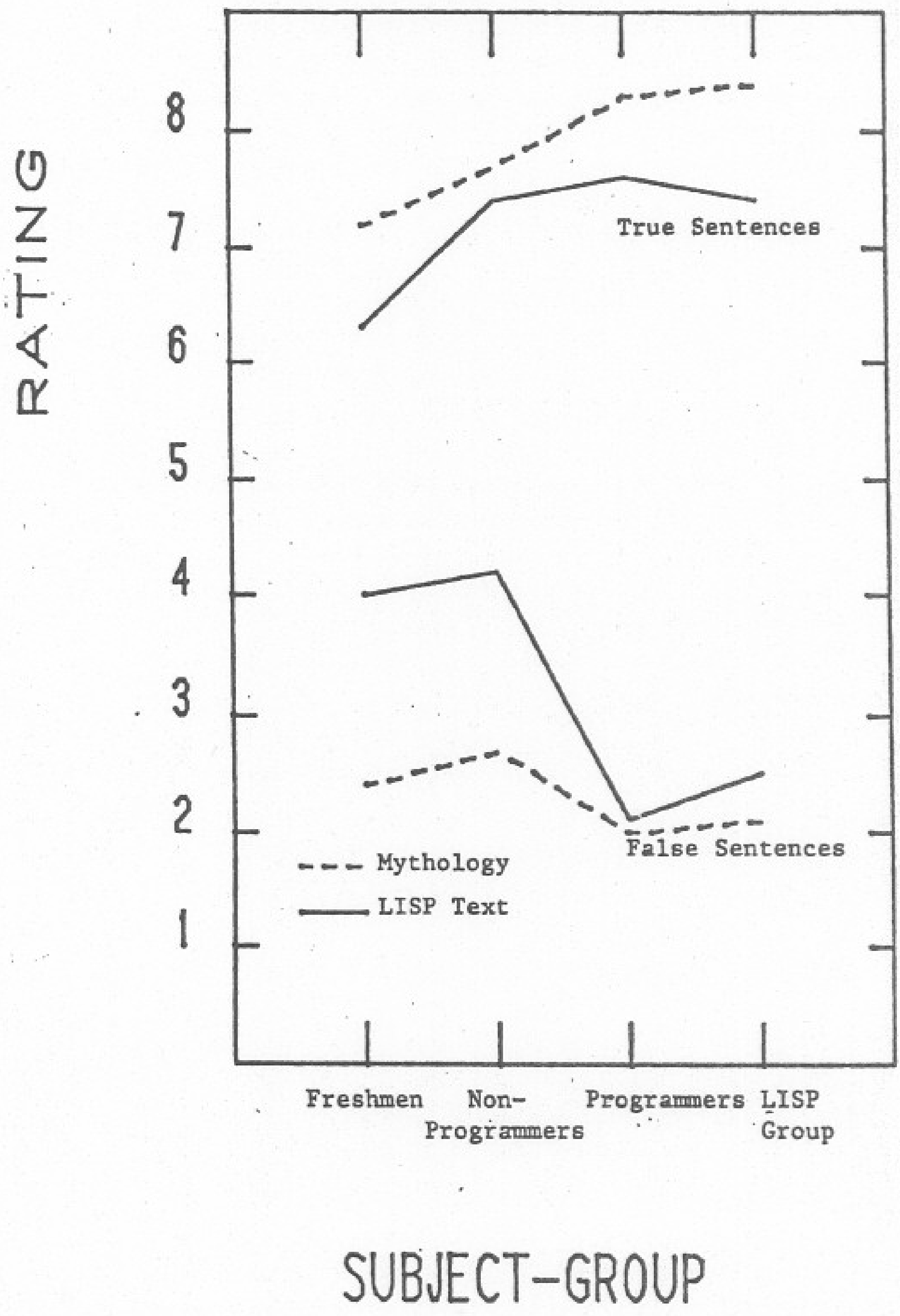

Figure 1. Average nine-point ratings on the explicit sentences for each subject group on both texts.

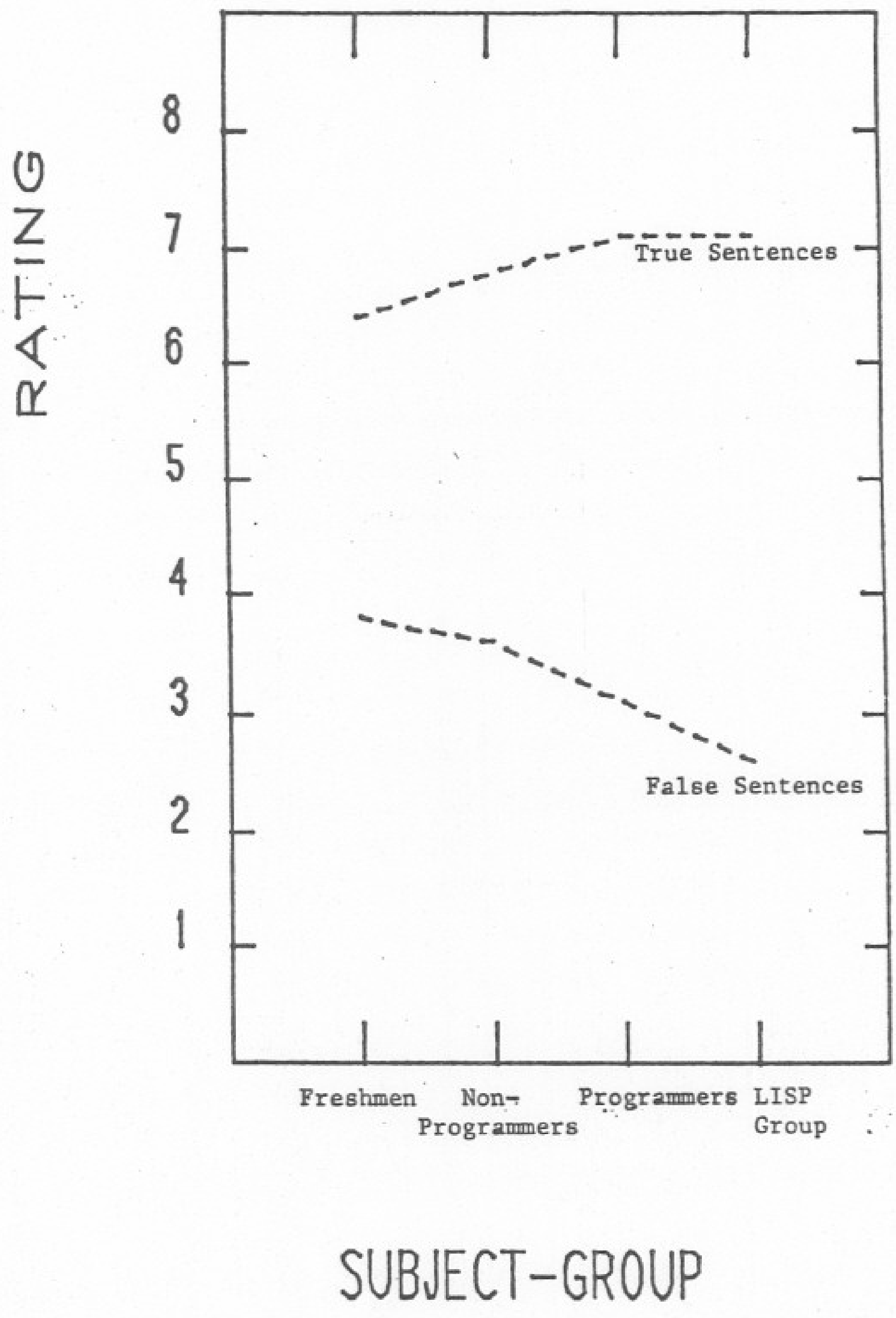

Figure 2. Average nine-point ratings on the implicit sentences for each subject group on the mythology text.

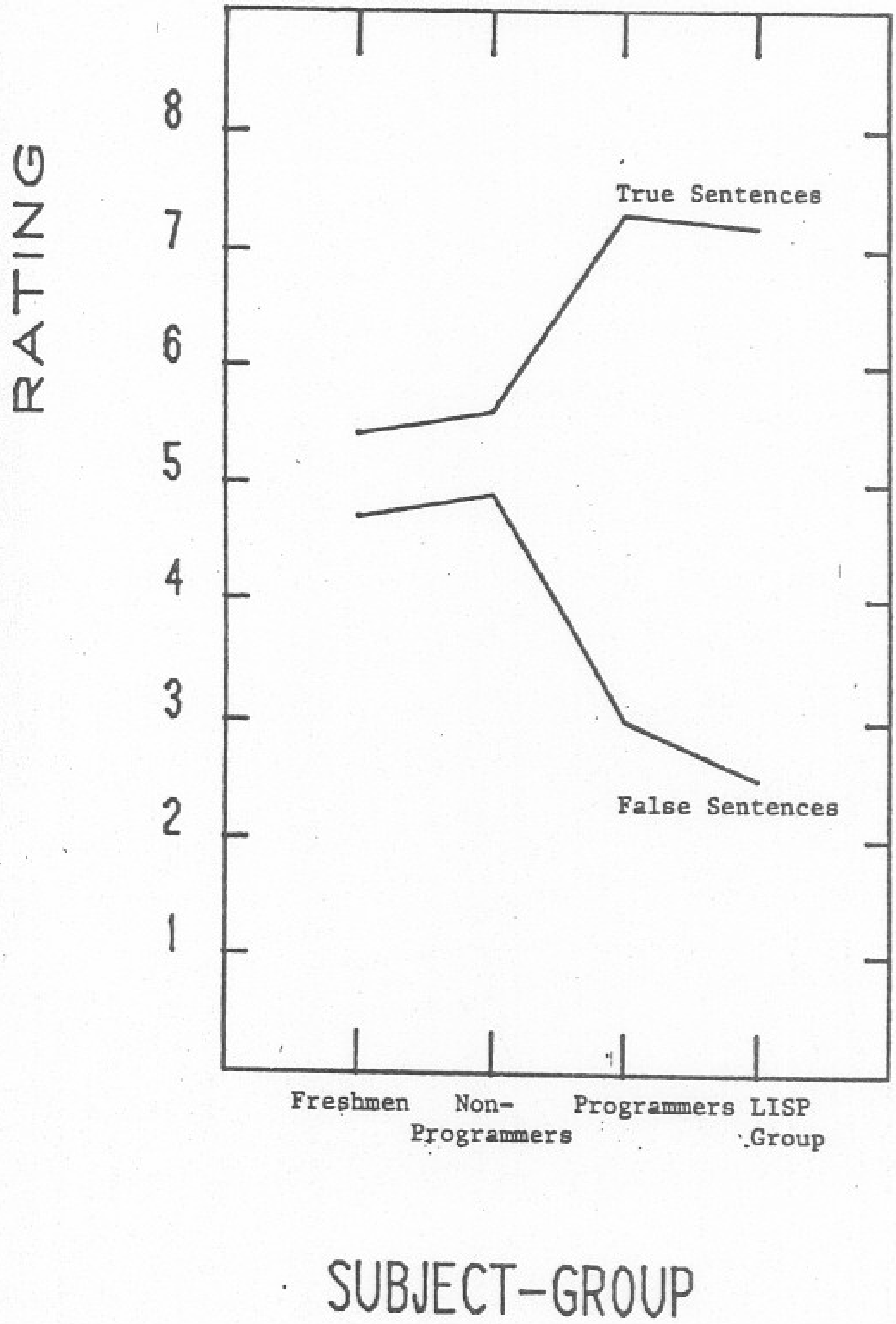

Figure 3. Average nine-point ratings on the implicit sentences for each subject group on the LISP text.

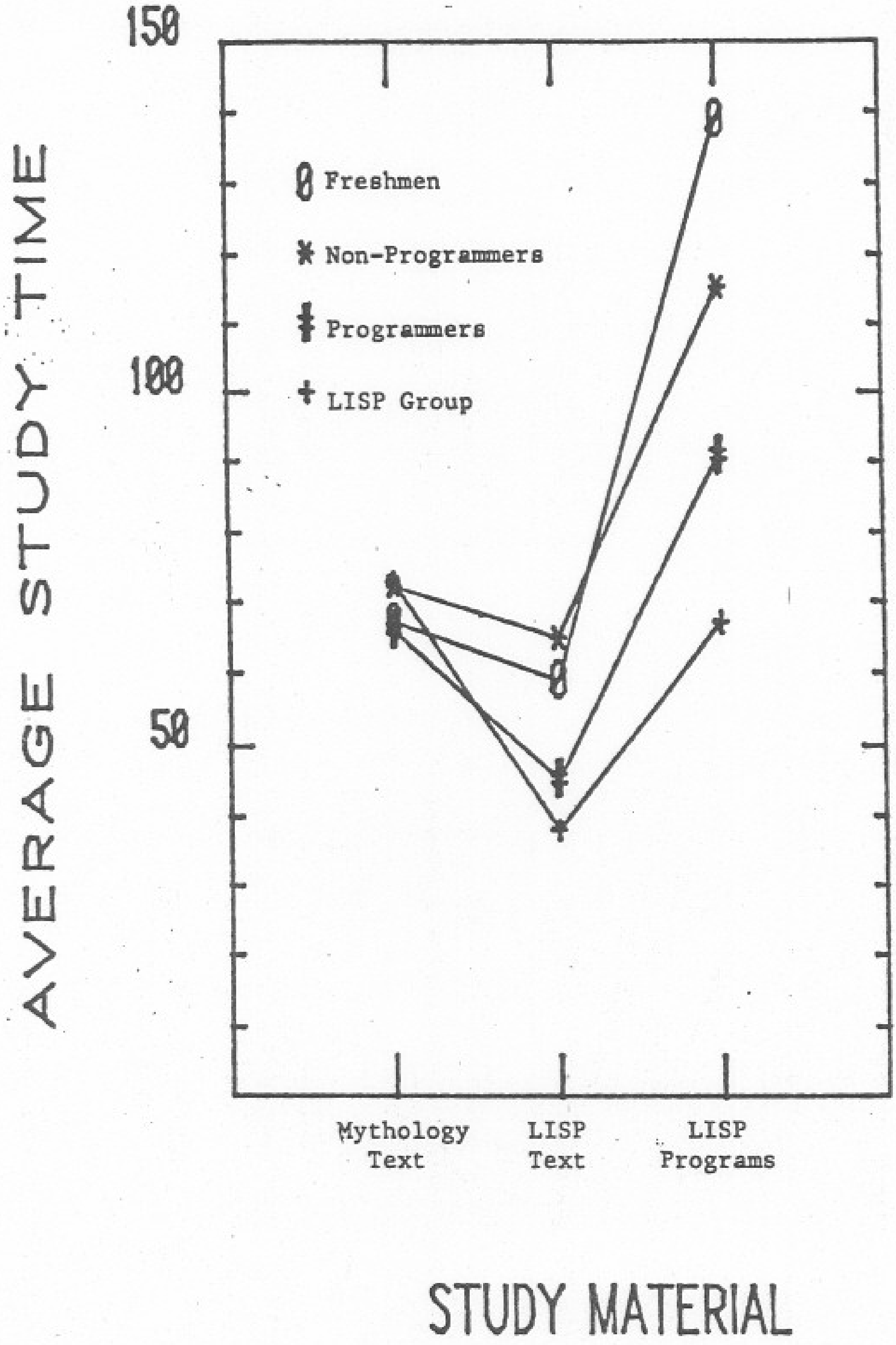

Figure 4. Average study time per character (in msec) for the mythology and LISP texts and LISP programs.

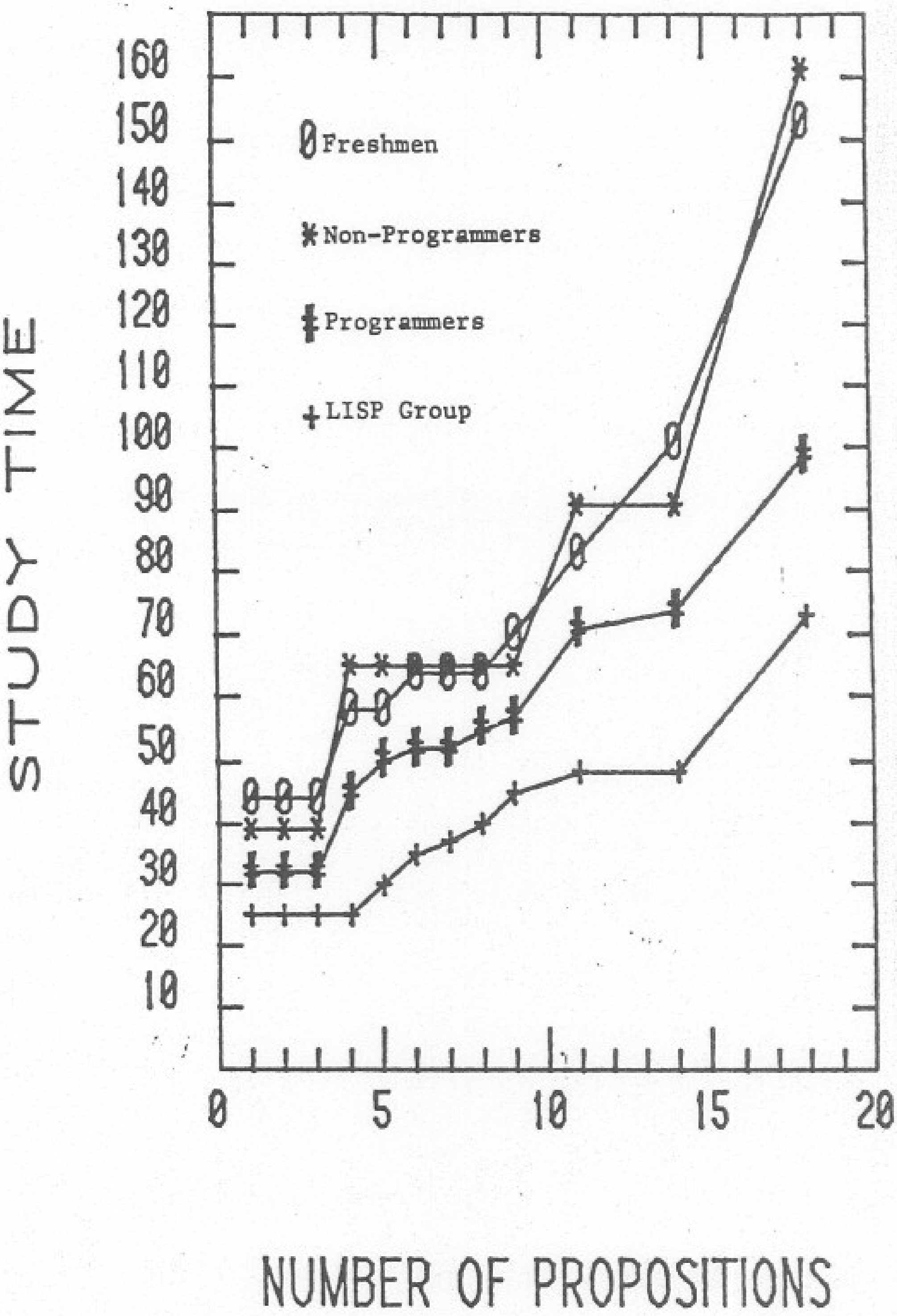

Figure 5. Estimated average study time per character (in msec) as a function of the number of propositions in a sentence in the LISP text.

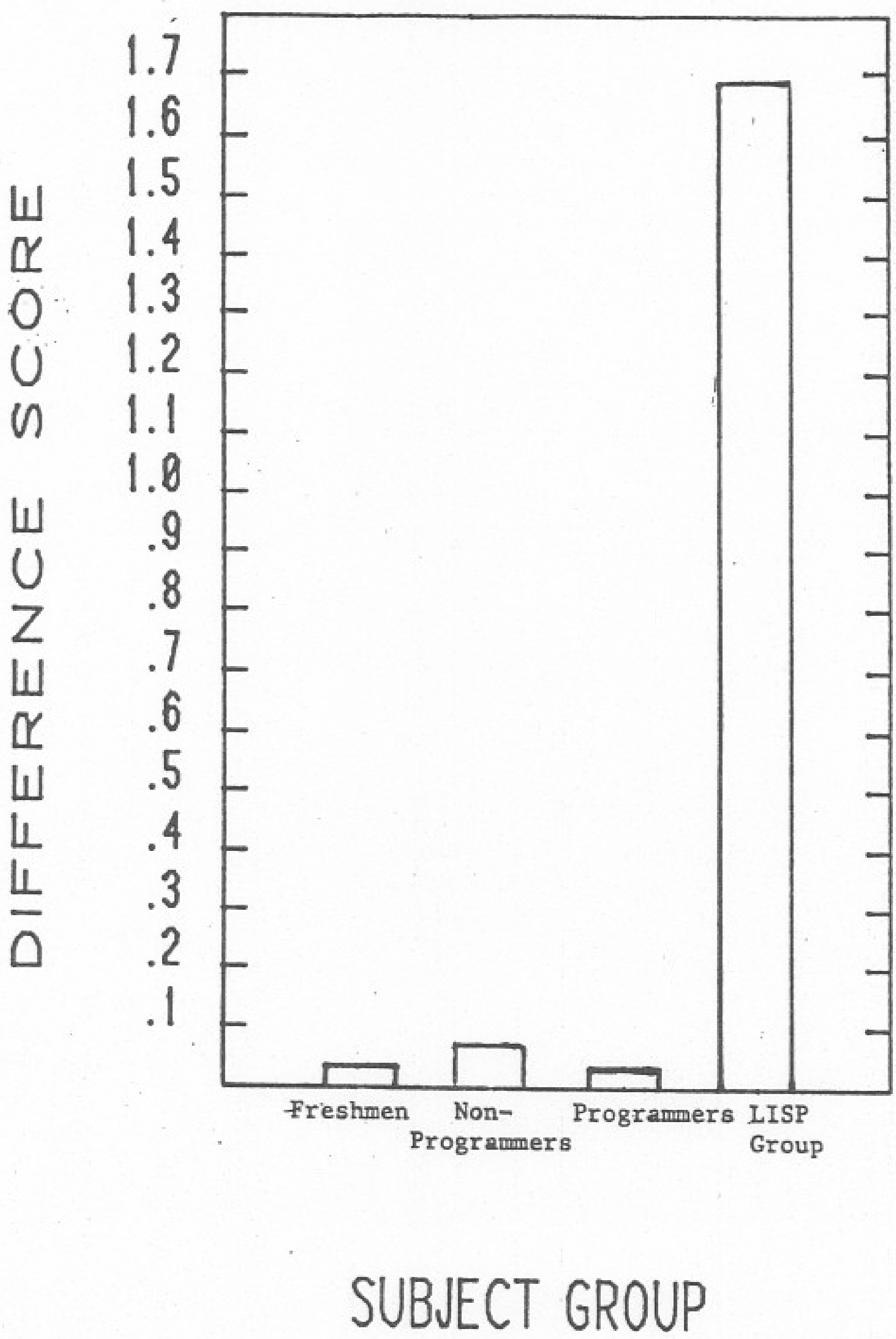

Figure 6. Mean difference between the number of regressive eye-movements per sentence in the LISP and mythology texts for each subject group.

Bisher erschienene Berichte aus dem Psychologischen Institut der Universität Heidelberg

Diskussionspapier Nr. 1: GROEBEN, N.: Vom behavioralen zum epistemologischen Subjektmodell: Paradigmawechsel in der Psychologie? September 1975

Diskussionspapier Nr. 2: MÖBUS, C. & SIMONS, H.: Zur Fairness psychologischer Intelligenztests gegenüber ethnischen und sozialen Gruppen: Kritik klassischer Konzepte. Oktober 1975

Diskussionspapier Nr. 3: WOTTAWA, H.: Skalenprobleme bei probabilistischen Meßmodellen. März 1976

Diskussionspapier Nr. 4: TREIBER, B. & PETERMANN, F.: Zur Interaktion von Lernermerkmalen und Lehrmethoden: Rekonstruktion und Normierung des ATI-Forschungsprogramms. April 1976

Diskussionspapier Nr. 5: MÖBUS, C. & WALLASCH, R.: Zur Erfassung von Hirnschädigungen bei Kindern: Nichtlineare Entscheidungsregeln auf der Basis von Veränderungsmessungen. August 1976

Diskussionspapier Nr. 6: SCHEELE, B. & GROEBEN, N.: Voraussetzungs- und zielspezifische Anwendung von Konditionierungs- vs. kognitiven Lerntheorien in der klinischen Praxis. Dezember 1976

Diskussionspapier Nr. 7: MÖBUS, C.: Zur Analyse nichtsymmetrischer Ähnlichkeitsurteile: Ein dimensionales Driftmodell, eine Vergleichshypothese, TVERSKY's Kontrastmodell und seine Fokushypothese. Juni 1977

Diskussionspapier Nr. 8: SIMONS, H. & MÖBUS, C.: Veränderung von Berufschancen durch Intelligenztraining. Juli 1977

Diskussionspapier Nr. 9: BRAUNMÜHL, C. v. & GRIMM, H.: Zur Kommunikationspsychologie: Über Versuche der methodischen Konstitution eines genuin humanwissenschaftlichen Forschungsansatzes zur Entwicklung der Verständigungsfähigkeit. November 1977

Diskussionspapier Nr. 10: HOFER, M.: Entwurf einer Heuristik für eine theoretisch geleitete Lehrer- und Erzieherbildung. November 1977

Diskussionspapier Nr. 11: SCHEIBLER, D. & SCHNEIDER, W.: Probleme und Ergebnisse bei der Evaluation von Clusteranalyse-Verfahren. Juni 1978

Diskussionspapier Nr. 12: SCHEELE, B.: Kognitions- und sprachpsychologische Aspekte der Arzt-Patient-Kommunikation. September 1978

Diskussionspapier Nr. 13: TREIBER, B. & SCHNEIDER, W.: Mehrebenenanalyse sozialstruktureller Bedingungen schulischen Lernens. Oktober 1978

Diskussionspapier Nr. 14: AHRENS, H.-J. & KORDY, H.: Möglichkeiten und Grenzen der theoretischen Aussagekraft von multidimensionalen Skalierungen bei der Untersuchung menschlicher Informationsverarbeitung. Teil I: Formale und wissenschaftstheoretische Grundlagen. März 1979

Diskussionspapier Nr. 15: GROEBEN, N.: Entwurf eines Utopieprinzips zur Generierung Psychologischer Konstrukte. Juni 1979

Diskussionspapier Nr. 16: WEINERT, F.E. & TREIBER, B.: School Socialization and cognitive development. Juni 1979

Diskussionspapier Nr. 17: GUNDLACH H.: Inventarium der älteren Experimentalapparate im Psychologischen Institut Heidelberg sowie einige historische Bemerkungen. 1978

Diskussionspapier Nr. 18: SCHEELE, B. & GROEBEN, N.: Zur Rekonstruktion von subjektiven Theorien mittlerer Reichweite. Eine Methodik-Kombination von halbstandardisiertem Interview (einschließlich Konfrontationstechnik) und Dialog-Konsens über die Theorie-Rekonstruktion mittels der Struktur-Lege-Technik (SLT). Dezember 1979

Diskussionspapier Nr. 19: GLOGER-TIPPELT, G.: Subjektive Theorie von Frauen über ihre erste Schwangerschaft: Theoretische Konzepte und methodische Möglichkeiten. Januar 1980

Diskussionspapier Nr. 20: KÄMMERER, A.: Das Konzept 'psychotherapeutische Strategie' am Beispiel des Problemlösens. Juli 1980

Diskussionspapier Nr. 21: SCHEELE, B.: (unter Mitarbeit von B.Tuschen und C.Maier): Subjektive Theorien über Ironie - als Heuristik für einen wissenschaftlichen Hypothesenkorpus. August 1980

Diskussionspapier Nr. 22: TREIBER, B.: Erklärung von Förderungseffekten in Schulklassen durch Merkmale subjektiver Unterrichtstheorien ihrer Lehrer. Oktober 1980

Diskussionspapier Nr. 23: RÖHRLE, B. & KOMMER, D.: Handlungstheoretische Betrachtungen zur primären Prävention psychischer Störungen. Februar 1981

Diskussionspapier Nr. 24: VOIGT, F.: Die Entwicklung des Zahlbegriffs. Teil I: Entwicklungslinien des Zahlbegriffs im Vorschulalter: Übersicht über theoretische Probleme und empirische Untersuchungen, mit einer Bibliographie zur Zahlbegriffsentwicklung. Teil II: Entwicklungslinien des Zahlbegriffs im Vorschulalter: Deskriptive Untersuchung des kindlichen Zahlverständnisses und verwandter Konzepte. April 1981

Diskussionspapier Nr. 25: SCHNEIDER, G. & WEIMER, E.: Aspekte der Kategorisierung städtischer Umwelt - Eine empirische Untersuchung. Juni 1981

Diskussionspapier Nr. 26: SCHNEIDER, W. & SCHEIBLER, D.: Zur Evaluation numerischer Klassifikation: Probleme beim Vergleich von Clusteranalysen. August 1981

Diskussionspapier Nr. 27: DRINKMANN, A. & GROEBEN, N.: Techniken der Textorganisation zur Verbesserung des Lernens aus Texten: Ein metaanalytischer Überblick. November 1981

Diskussionspapier Nr. 28: GRAUMANN, C.F.: Theorie und Geschichte. November 1982, Historische Reihe Nr. 1

Diskussionspapier Nr. 29: WOODWARD, W.R.: From the Science of Language to *Völkerpsychologie:* Lotze, Steinthal, Lazarus, and Wundt. November 1982, Historische Reihe Nr. 2

Diskussionspapier Nr. 30: SOMMER, J.: Dialogische Forschungsmethoden. Dezember 1982

Diskussionspapier Nr. 31: WINTERMANTEL, M. & CHRISTMANN, U.: Textverarbeitung: Empirische Untersuchung zum Verstehen einer Personbeschreibung. Januar 1983