INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND

7913176

ZURAWSKI, VALERIE B. R.
THE SOLUTION OF TWO-TERM SERIAL PROBLEMS: EFFECTS OF PROBLEM STRUCTURE, QUESTION LOCATION, AND DISPLAY DELAY.

CITY UNIVERSITY OF NEW YORK, P.H.D., 1979

University Microfilms International 300 N. ZEEB ROAD, ANN ARBOR, MI 48106

THE SOLUTION OF TWO-TERM SERIAL PROBLEMS: EFFECTS OF PROBLEM STRUCTURE, QUESTION LOCATION, AND DISPLAY DELAY

by

VALERIE B. R. ZURAWSKI

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1979

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Jan 27, 1979
date

Louis J Gerstman
Chairman of Examining Committee

January 29, 1979
date

Florence L. Denmark
Executive Officer

Prof. Ethel Weiss

Prof. Alden Wessman

Supervisory Committee

The City University of New York

Abstract

THE SOLUTION OF TWO-TERM SERIAL PROBLEMS:
EFFECTS OF PROBLEM STRUCTURE, QUESTION LOCATION,
AND DISPLAY DELAY

by

Valerie. B. R. Zurawski

Adviser: Professor Louis J. Gerstman

There has been an extended controversy between linguistic and placement theorists concerning the basic mental operations and principles used in the solution of serial problems. These investigations have concentrated on the three-term problem and have produced findings which are often inconclusive and contradictory. The present study was based on the assumption that two-term serial problems would prove a better way to assess the cognitive demands upon, and strategies employed by, the problem solver, especially if, in addition, the location of the question with respect to the premise and delays between question and premise were also varied.

Positive and negative (POSNEG) two-term problems were constructed to represent a question before premise or QPRE condition (e.g., "Who is best? Sam is better than Joe.") and a question after premise or QPOST condition (e.g., "Joe is not as bad as Sam. Who is best?"). Drawing on evidence presented by linguists such as Clark, these problems were classified by the principles of congruence (COMPAT) and

linguistic adjectival marking (ADJECT) and investigated under four display delays (0, .5, 1, and 2 seconds--TIMES) in both the QPRE and QPOST conditions.

Sixteen female and 15 male psychology students at CCNY were tested on 128 problems representative of the above effects as well as on 16 problems which defined a 4 second QPOST delay condition. Problems were presented visually on a Tektronix 4010-1 Telex screen. A PDP-8E computer, programmed to emit and control the timing of the stimuli, also recorded solution times.

Repeated measure analyses of variance on the correct and adjusted solution time data confirmed the following predictions: 1) The POSNEG, COMPAT, and ADJECT main effects and the POSNEG x COMPAT (PO x CO) interaction were highly significant at $p < .001$ in both the QPRE and QPOST conditions. That is, positive, congruent, and "better/not as good as" problems were solved faster than negative, incongruent, and "worse/not as bad as" problems, regardless of question placement. 2) The main effect TIMES was also highly significant in the QPRE and QPOST conditions. This inhered in an initial priming effect (0 to .5 sec.), but no further benefit when longer exposures of the question were allowed (1, 2 sec.); and a significant decrease in solution time as premise preexposure increased (0 to 2 sec.).

Results from separate analyses of the QPOST condition also showed the magnitude of effects to vary with display delay. For example, POSNEG and COMPAT reached their peak

effect in the .5 second delay; the ADJECT effect only became significant at the longer durations (1 and 2 seconds); and all main effects except for POSNEG (and POSNEG x ADJECT) lost significance by the added 4 second delay.

These findings may indicate that subjects rely on structural factors (POSNEG, COMPAT) for solution when there is limited processing time, or that processing of the adjective dimension requires more processing time than both the logical "not" and match for congruence. These findings may speak for a structure to meaning switch within the encoding process, and indicate that by the 4 second QPOST delay, subjects are prepared to answer either question.

ACKNOWLEGEMENTS

I wish to express my appreciation to each member of my committee: Professors Louis J. Gerstman, Chairman; Ethel Weiss; and Alden Wessman.

I especially wish to thank Dr. Gerstman for the many hours he devoted to overseeing the completion of my project. I also wish to thank Dr. Weiss for her help and guidance during my initial work with three-term serial problems which was presented as a paper at the 83rd meeting of the American Psychological Association.

Special gratitude and acknowledgement is also due to Dr. Ira Kaplan who continued to encourage and advise me on the design of the present study even after his relocation to Kansas City.

As a special note, I wish to dedicate this paper to my mother and to express appreciation to her, John, and Cal for their love and support.

Table of Contents

Table of Contents

List of Tables

I

The Problem

Asked to solve a problem such as "Who is best? A is better than B," the reader would be apt to produce a fairly quick response. However, when asked to solve "A is not as bad as B. Who is best?", a reader might pause, restate, and even reformulate the problem before achieving a solution. Problems of this nature are defined as the two-term serial problem. They consist of one premise, which compares two-terms such as "A" and "B" on some adjectival dimension, and a question which asks the reader which term demonstrates most of that quality. Although the above examples compare terms on a "good-bad" dimension, other adjective pairs could be used. Structurally, however, the basic forms of the premise are:

I. A is better than B

II. A is worse than B

I'. A is not as bad as B

II'. A is not as good as B.

Names such as Sam and Joe are usually substituted for the "A" and "B" terms and the question asked conventionally is the superlative form "Who is best?" or "Who is worst?"

Higher order problems which can vary in structure can also be constructed from this problem type. Three-term problems (also known as linear syllogisms or transitive inferences) consist of two premises and three terms (A, B,

and C). They take the eight forms (I to IV and I' to IV') presented in Table 1. Similarly, four-term problems which contain three premises and five-term problems which contain four premises could also be built from the simplest two-term problems.

Although the number of premises which appear in a problem can vary, researchers have tended to favor the three-term problem (with question after premises). Recent studies of this problem, however, have led to a heated controversy. The heavily debated issue is whether the encoding process is linguistic or visual.

Borrowing linguistic principles from the study of transformational grammar, theorists such as Clark (1969a; 1969b) have developed a linguistic model to explain the encoding and solution of serial problems. Other theorists such as Huttenlocher (1968) and De Soto, London, and Handel (1965) have developed a placement, spatial, or imagery model which proposes that subjects encode and order serial problems on a mental array.

A further area of dispute between the two models has been the nature of the adjective. The linguists propose that the "lexical markings" in deep structure of the adjectives contained in a problem relate to the ease of retrieving a solution. Conversely, the spatial theorists contend that subjects' tendencies to assign particular adjectives to one end of a physical or imagined array is important for problem success.

Table 1

Three-Term Serial Problems

Positive Comparatives		Negative Equatives	
I	A > B, B > C	I'	A is not as < as B, B is not as < as C
II	C < B, B < A	II'	C is not as > as B, B is not as > as A
III	A > B, C < B	III'	A is not as < as B, C is not as > as B
IV	B < A, B > C	IV'	B is not as > as A, B is not as < as C

Note. In the positive comparatives, " >" can stand for "better," "taller," or "lighter (haired)," while "<" stands for "worse," "shorter," or "darker (haired)."

In the negative equatives, " >" can stand for "good," "tall," or "light (haired)," while "<" stands for "bad," "short," or "dark (haired)."

Despite their differences in orientation, it is interesting to note that both models are able to predict and explain solution of the positive comparative problems (see Table 1, I - IV). This is not the case, however, with negatively stated or negative equative problems (Table 1, I' - IV') where the models make competing predictions. Clark proposes that only the linguistic principles of deep structure can explain solution of both the positive and negative forms.

In reply, the placement theorists have sought to expand their model to include solution of negative problem types. In their later investigations of these problems, Huttenlocher, Higgins, Milligan, and Kauffman (1970) have tentatively proposed that subjects use one of three symbolic reorganizations to change a negative equative problem into a positive comparative form. They suggest that subjects reorganize "A is not as bad as B" to "B is worse than A" so that the subject and object ("A" and "B") are reversed. Huttenlocher et al. (1970) and Huttenlocher and Higgins (1971) further contend that this operation is consistent with their solution model.

Although this reorganization principle has also been employed in an attempt to synthesize the linguistic and placement theories (Johnson-Laird, 1972), no experimental research to date has proved this theory or been credited for putting the controversy to rest.

It is also worth noting that many and varied experi-

mental methodologies have been used to resolve this feud and to test the hypotheses concerning the encoding process as well as the role of different relational terms. For example, subjects presented with a problem can follow instructions to solve the problem either in their heads, on paper, or by actually arranging the terms on a board as if they were physical objects. Premises can be split up and solution latency timed from either the beginning or end of a statement until the question is answered. However, little research at the time of this experiment's conception concerned the location of the question with respect to the premise or the amount of exposure given to a subject for each problem. Almost invariably the question has immediately followed the premises. Although Johnson-Laird (1972) had recommended placing the question before the premise in order to examine how a subject might work backwards from the solution of a three-term serial problem, little work on this has been recorded. Potts and Scholz (1975, Note 1), however, had demonstrated that the amount of study time did affect encoding and solution of three-term problems. They found that subjects' performance under the condition where the premises and question were presented simultaneously was explained by Clark's linguistic principles. On the other hand, when subjects were permitted unlimited viewing of the premises before signaling for, and receiving the question, these same principles weakened in predictive strength.

Given the unresolved controversy over three-term

problems, the simpler two-term problem was chosen for study here. It seemed plausable that an investigation of what may be considered the building block of all higher order serial problems would: 1) clearly elucidate what cognitive demands were made on the problem solver, and 2) lay the groundwork for future study of the encoding of three-term problems. Although these two-term serial problems have rarely been studied in and of themselves, existing research has shown that they, like the three-term problems, vary systematically in structure and solution speed (Clark, 1969b; Huttenlocher & Strauss, 1968). Problem structure, therefore, was chosen as an area of investigation as was the location of the question in relation to the premise.

In addition, a different method was used in order to study the effect of time on problem solution (Kaplan, Note 2). Subjects were allowed fixed exposure of the first part of the problem, ranging from 0 to 2 seconds, before the second part appeared on a viewing screen. This method was suggested by Posner, Bois, Eichelman, and Taylor's (1969) work on visual encoding and permitted exploration of at least three issues: 1) How would varying exposure of the premise before presentation of the question affect the encoding of a problem's structure? 2) Would certain linguistic principles weaken in predictive strength as pre-exposure of the first part of the problem increased or would the power of these principles remain unchanged over time? 3) Given the structure of a problem, how would the location of the

question prepare or interfere with a subject's processing of a premise? It was hypothesized that answers to these questions would highlight the importance of encoding time and provide a groundwork for future research with three-term problems.

II

Review of Prior Research

In this chapter, a review of the linguistic and then placement models will be presented. A review then follows of the serial problem and related literature pertaining to the variables under investigation. These variables include: 1) The presentation of positive vs. negative information; 2) Congruence or compatibility of the information in deep structure underlying the sentence or premise and question; 3) The linguistic nature of the adjective; 4) The location of the question with respect to the premise(s); 5) The amount of exposure permitted for each problem.

The Linguistic Approach to Solving Serial Problems

Drawing from his linguistic background, Clark (1969a, 1969b) proposes three linguistic principles to account for how a person stores, searches, and retrieves abstract information relevant to serial problem solving. He theorizes that a subject interprets and stores information from the premises and question in terms of deep grammatical structure. Upon completion of this encoding, a subject then searches his/her memory for information that is congruent in deep structure with the information asked for in the question. Clark's main thesis is that the linguistic factors affecting the storage and subsequent retrieval of information determine how quickly such a problem is solved.

Clark's first linguistic principle was the primacy of functional relations. In the same way as the standard logical subject, verb, and object of a sentence, these functional relations were said to underlie a sentence and be stored in a "more readily available form" after comprehension as base strings. Clark employed this principle to explain why the premises, "Sam is better than Joe" and "Joe is worse than Sam," although appearing to mean the same thing, did not impart the same information. He hypothesized that subjects encoded the problem, "Sam is better than Joe. Who is better?" in terms of the base strings, "Sam is good+ Joe is good. Who is good++?" Encoding of the problem "Joe is worse than Sam. Who is worse?" however, took a different form. Here subjects encoded the base strings, "Joe is bad+ Sam is bad. Who is bad++?" and realized that Sam and Joe were bad more readily than that Joe was more extreme in badness than Sam.

According to this analysis, various positively and negatively stated problems contained underlying similarities. Employing the notation of Table 1, two- and three-term serial problems I and II' were said to have the same underlying base strings of "good," while problems II and I' contained "bad." Likewise, three-term problem forms III' and IV were said to be similar in terms of their base strings as were problems IV' and III.

The second linguistic principle was congruence. Clark hypothesized that a subject searched memory to retrieve

information congruent to the information in the problem question. In the linguistic approach, this congruence was based on the functional relations (base strings) of both the premises and question in deep structure. If the two pieces of information were incongruent or resulted in the pairing of a "good" string with a "bad" string, the subject reformulated the question and was then able to complete the search for congruence. Since Clark always presented the premises and question simultaneously to his subjects, he contended that the premise information was processed, comprehended, and stored before the question. As a result, information from the question was reformulated and the premise information was left untouched.

The third linguistic principle was lexical marking. This principle concerned the different properties attributed to antonymous adjectives as a result of their linguistic asymmetry. Clark believed that members of polar adjective pairs were either considered linguistically unmarked (e.g., "better," "taller") or marked (e.g., "worse," "shorter") and that there were various advantages for the unmarked member over the marked form. For one, the unmarked or "positive" adjective was felt to serve as the name of the full scale being described. Clark also proposed that this adjective could be used in both the nominal and contrastive sense, and could neutralize in comparative statements. (For example, "A is better than B" can mean that "A" and "B" are only being compared evaluatively.) On the other hand, the marked

adjective (e.g., "worse" or "shorter") did not double as the scale name. The marked adjective could only be used to make a contrast. And, unlike the positive or unmarked member, this marked form implicated both terms as being at the ends of an extreme (e.g., "A is worse than B" implies that "A" and "B" are both bad.) Due to this analysis, Clark proposed that the neutral sense of the unmarked adjective was encoded in a simpler way than the marked adjective. This encoding also facilitated the storage and retrieval of the information in serial problems which contained the unmarked rather than the marked form, and applied to positive comparative (e.g., "better/worse") as well as negative equative (e.g., "not as good/bad as") forms.

Clark's only attempt at explaining how information from the two premises of a three-term serial problem became combined was by a fourth principle of "compression." He proposed that a subject compressed the information in each premise to a less cumbersome form in order to save memory space. The premise, "Joe is better than Sam," therefore, became "Joe+," the term attributed the most extreme quality of the scale. Confusion developed in the course of solving a problem, however, if Joe was not repeated in the second premise. In this case, the alignment of information was not immediate and the subject then had to backtrack to retrieve the full version of the first premise to figure out Sam's relationship to Joe.

As to the actual nature of this encoding process, Clark believed that the information retained from exposure to a serial problem was something highly abstract. He contended that whereas the more concrete attributes of a sentence (e.g., word phrases, surface structure) appeared to be done away with soon after a person heard them, this underlying abstract interpretation remained both for immediate use and for use at a "much later time." Given his linguistic orientation, Clark believed that a subject's solution strategy was analogous to "processing language without awareness" so that the underlying base strings were most accessible. It was the accessibility of this information and subsequent match for congruence that allowed for a quick solution to the serial problem.

In his original experiments, using randomized blocks of 32 three-term serial problems (displayed on cards), Clark found evidence to support his linguistic theory that certain problems would be easier to store and retrieve than others. He found that positive comparative problems were easier than negative equative problems as were linguistically congruent over incongruent problems and problems with the unmarked over marked adjective form. Subjects' solution time and error data, measured respectively, from a signal to turn up a problem card which was silently read until an answer was produced aloud (Clark, 1969b), and by circling an answer within a 10 second limit (Clark, 1969a) also confirmed his predictions based on deep structure. Still

employing the notation of Table 1, he found that three-term (and two-term, Clark, 1969b) problems I and II' were easier than II and I' due to lexical marking, deep structural similarity, and congruence; and that, regardless of the adjectives used, three-term problems III and IV' were easier to solve than III' and IV due to the principles of functional relations, congruence, and compression.

In sum, Clark proposes that there are certain linguistic principles based on deep grammatical structure that underlie the ability to understand language and solve serial problems. From his research with two- and three-term serial problems, he establishes a model which emphasizes the importance of the encoding and comparison operations involved in the solution process. His results indicate that positive sentences and problems are encoded in a simpler way than negative forms and that congruent information is easier to retrieve than incongruent information. He views an unmarked or neutral adjective as being linguistically simpler than a marked adjective and contends that an unmarked adjective is encoded and retrieved more quickly than its marked counterpart.

Object Placement Theories as Applied to the Solution of Positive and Negative Serial Problems

Huttenlocher (1968) proposes that subjects solve positive three-term serial problems as if they were actually placing objects in an array. She contends that both serial problems and placement tasks are easiest when the grammatical

or logical subject of the second premise or statement is moveable rather than immovable. Although her work with negatively stated problems and tasks produces mixed results, Huttenlocher maintains that the analogy between solution methods can still be upheld. She proposes that subjects reorganize certain implicitly negative problems so that the subject and object of a premise are reversed. Clark, on the other hand, argues that her methodology for testing the validity of this subject-object principle is faulty.

In her work with object placement tasks, Huttenlocher found that subjects achieved an easier solution to problem task instructions which said, "Block A is on top of Block B, put Block C below B" than to "Block B is below Block A, put Block B on top of Block C" (Huttenlocher, Eisenberg, & Strauss, 1965; Huttenlocher & Strauss, 1968). She attributed this result to the preferred grammatical-subject status of the "C" block or term in the second sentence. She contended that "C" was more easily mobile as the subject rather than the object of an active statement or request, and that when it was the object, an added operation or a reversal of terms was needed before the "C" term could be placed.

Huttenlocher (1968) further argued that there was a correspondence between solution of placement tasks and serial problems in the use of this subject-object principle. Instead of working with actual blocks which are placed on a board, a serial problem solver was said to move mental

representations of the items along a spatial array located in a temporary storage area or "cognitive space." Huttenlocher viewed this process as being different from Clark's memory store for deep structure and believed that the representations were written symbols or words.

Although Clark finds that negative equative reasoning problems are generally easier when the third item is described as the grammatical-object rather than the subject, Huttenlocher (Huttenlocher et al., 1970) claimed that an analogy between performance on negative placement tasks and negative equative serial problems was still plausible. Huttenlocher et al. (1970) found that depending upon the adjective contained in the problem or placement task, results could conform to either Clark's model or to their own. For example, whereas performance by adults on three-term reasoning problems and placement tasks using the adjective pair "good/bad" conformed to Clark's solution model, adults' and children's performance on three-term reasoning problems and placement tasks using the adjective pair "high/low," respectively, conformed to Huttenlocher et al.'s subject-object model. Upon closer examination, the authors found that only the adjective "bad" in both the comparative and negative problems produced significant subject-object mobile reversals under both the reasoning and placement conditions.

Due to this result, the authors proposed that problem solvers used one of three symbolic reorganizations to

change the negative to a positive form. They also argued that the inherent properties of an adjective affected the reorganization and explained why there were "subject-object" reversals with the "not as bad as" problems. They theorized that in solving an "A is not as bad as B" problem, subjects bypassed the first symbolic reorganization (A is not bad compared to B) in order to avoid a double negative. Likewise, they believed that subjects bypassed the second reorganization (A is good compared to B) because a paraphrase of the evaluative phrase "not as bad as" was not interchangeable with one containing its opposite "good." On the other hand, Huttenlocher et al. proposed that the third reorganization from "A is not as bad as B" to "B is worse than A," was permissible. This reorganization would result in the grammatical subject and object of the premise being reversed and lend support to their subject-object principle.

Clark (1972b), however, argued that the placement procedure used by Huttenlocher et al. (1970) was confounded with a "reasoning" task and, therefore, was an unrelated test of the imagery theory. Clark argued that Huttenlocher et al. operated with a "mixed placement task" because a subject heard the first premise instruction; placed these two items in an array; and then placed the third item by the second instruction. With this procedure, a subject could either ignore the first premise or use the first and second premise together to discover the final ordering of the third item. On the other hand, the procedure which Clark

used and defined as a "pure placement task" insured that the subject ignored the first premise and dealt only with the second premise information. Here, a subject read an instruction that was the second premise; viewed an array which was already constructed from the first premise; and inserted the third item. Although timing proceeded from the final placement of the third object until placement was completed, Clark argued that these procedures produced different results. Unlike findings from his pure placement task, results from the mixed placement experiments were generally consistent with data from the "reasoning" tasks. He concluded that this factor spoke against there being a correspondence between findings from reasoning or serial problem solution and placement tasks as Huttenlocher et al. had claimed.

Clark (1972b) also proposed that his theory about adjective marking and neutralization was stronger in predictive value than the negative adjective theory of Huttenlocher et al. He cited that he had found a significant 42 msec. advantage of "long" over "short" in a sentence judgment task which lent support to the prediction by the negativity theory. (This states that implicitly negative adjectives such as "short" and "bad" are harder than "tall" and "good," respectively.) However, a more significant advantage of the neutral and non-negative adjective "longer" over "shorter," as predicted by his neutralization theory, was evident. Clark concluded that whereas the "negativity"

theory had relevance for models attempting to explain the solution of serial problems, the neutralization of unmarked adjectives was the more important factor.

In sum, Huttenlocher proposes that subjects mentally solve three-term serial problems as if they were actual placement tasks. By an examination of solution time and error data, she finds that both placement task and serial problem solution are easier if the third of "C" term is the logical or "moveable" subject of the second premise (e.g., "A is better than B, C is worse than B"). Although there is support for her theory with positive problems, her work with the negative equative produces mixed results. As a consequence, she hypothesizes that certain reorganiztions of the negative are made which depend on the nature of the adjective (i.e., if the adjective is implicitly negative). Clark, on the other hand, accuses the placement theorists of using a confounded methodology, and continues to find support for his linguistic principles.

Positive vs. Negative Wording as it Affects Comprehension and Problem Solution

There has been widespread agreement among linguists (e.g., Clark & Card, 1969; Mehler, 1962; Miller, 1962; Miller & McKean, 1964) that a negative sentence is somewhat more complex than a sentence expressed in an active affirmative or kernel form. Their contention is that the negative requires an additional feature or processing operation than the "psychologically primary" kernel sentence. Psycho-

linguists (e.g., Slobin, 1966) and cognitive experimenters (e.g., Greene, 1970a, 1970b); Johnson-Laird & Tridgell, 1972; Johnson-Laird & Wason, 1972), on the other hand, attribute the difficulty of the negative to semantic factors. They, too, however, find that negatively stated tasks result in longer processing times and more errors than positively worded forms.

In a task where college subjects listened to and later tried to recall the various sentence types with the aid of prompt words, Mehler (1962) found that recall performance for the affirmative kernel was significantly better than for the other sentence constructions (e.g., negatives). In examining subjects' errors, he found that there was a marked tendency for subjects to recall sentences as kernels and to simplify the non-kernel forms by omissions. As a result, Mehler proposed that subjects analyzed sentences syntactically and used a "schemata plus correction" method to aid in the recall of various sentence types. He concluded that the more complex sentences (e.g., negatives, passives) were encoded as the kernel with an additional tag to signify their appropriate transformation.

Clark and Card (1969) also proposed that complex sentences were linguistically encoded as the simpler forms with an additional tag or feature. Their hypothesis was that this added feature would be lost over time and that a bias for simpler constructions would be evident. In testing the prompted and verbatim recall performance of 64

subjects on various sentences, the authors found strong support for their predictions. There was both a significant tendency for subjects to reconstruct and recall: a) positive sentence forms more often than negative forms; b) posi- comparatives using the relation "better than" and negative equatives using the relation "isn't as good as" more often than negative comparatives using "isn't better than" and positive equatives using "is as good as"; and c) unmarked sentence forms more often than marked adjective sentence forms. (Carpenter (1974) has also found that subjects recall positive and unmarked adjective sentences significantly more often than negative and marked adjective forms.) The authors concluded that subjects clearly tried to preserve the sense of these "primitive constituent sentences" (base strings) rather than the meaning of the relations described in the sentences.

Slobin (1966) also investigated syntactic sentence forms but proposed a semantic explanation for the difficulty of the negative. In his experiment, 5 groups of 16 subjects (children and adults) verified sentences as true or false by comparing them to scenes represented in a picture. Four syntactic forms were used (kernel, passive, negative, and passive negative), and response latencies were recorded. Slobin found that the main effect (sentence form) was significant and that ordering of the problems from the easiest to most difficult was: the kernel, the passive, the negative, and the passive negative. Because the

negatives, on average, were more difficult than even the passives, Slobin concluded that there were semantic problems with the negative that outweighed the syntactic problems of the passive. Upon further examination of his data on the negatives, he also found that verification times to negative sentences were significantly longer than to positive sentences; errors occurred twice as often with the negatives than with the positives; and that true negatives took significantly longer to verify than false negatives (2.14 vs. 1.91 sec. average).

Greene (1970a, 1970b), too, investigated various sentence forms and found that there were certain semantic functions peculiar to the negative. In her experiment, response latencies were recorded as college subjects decided if pairs of sentences were similar or different in meaning. Greene found that when active and passive sentences were compared, response time was faster if the two sentences meant the same thing (e.g., "A exceeds B," "B is exceeded by A"). However, when active negative sentence pairs were compared, response time was faster when the sentences were different rather than similar in meaning (e.g., "A exceeds B," "A does not exceed B" vs. "A exceeds B," "B does not exceed A"). Greene proposed that the negative serves certain specific functions in language. She stated that the "natural semantic function of the negative" was to signal a change in meaning, and that the negative was easier to understand when it corrected a misconception, contradicted, refused, or

pointed out a difference (see also Wason & Johnson-Laird, 1972). Greene concluded that this semantic feature was unique to the negative since the function of the passive did not involve this implied reversal.

Literature on deductive reasoning also indicates that the negative interferes with optimal processing and performance. In studying the solution rates of subjects using pencil and paper methods to solve "Who done it?" problems, Schwartz and Fattaleh (1972) found that problem wording (positive vs. negative) influenced the effectiveness of solution formats (e.g., matrix, informal grouping). Although the matrix, or the most efficient solution mode (see also Polich & Schwartz, 1974), was used about equally often in affirmative and negative problems (56% and 63%, respectively), the authors reported that the solution rate advantage of the matrix over other modes was greater in the affirmative (80% vs. 46%) than in the negative problems (37% vs. 32%). A closer examination of the 144 negative problems, in general, showed 45 to involve a negative to positive transformation. This added operation, however, provided a greater opportunity for error since 29 of the 45 transformations led to an incorrect representation. Schwartz (1971) and Schwartz and Fattaleh also found that negatively worded problems were solved significantly less often than the affirmative forms.

In sum, the work with negatively stated sentences or tasks indicates that the negative is clearly more difficult

than the positive form. Some experimenters attribute this to what may be additional or different syntactic operations involved in processing the negative compared to the positive. Others relate this finding to the semantic nature of the negative and say that its chief purpose is to deny. In any case, negatively worded deductive problems, including the two- and three-term serial problems (Clark, 1969a, 1969b), are reported as being significantly more difficult than the positively worded forms.

Congruent vs. Incongruent Information in the Solution of Serial Problems

Clark contends that a three-term serial problem is solved more quickly if the information in deep structure underlying the premises and question is congruent rather than incongruent. The placement theorists, on the other hand, propose a temporary accessibility or recency principle to explain some of Clark's results. In response, Clark continues to argue that this linguistic model and principle of congruence is stronger in predictive strength than the model proposed by Huttenlocher.

Huttenlocher and Higgins (1972) spoke out against Clark's principle of congruence, and, instead, proposed a "temporary accessibility" theory which could account for some of Clark's results with three-term serial problems. This principle stated that regardless of where the answer was located in the premise, a problem would be easiest when the adjective in the second statement matched

the adjective in the question. Huttenlocher and Higgins felt that because operations such as obtaining the semantic interpretation of a sentence consumed time, the word just "looked up" for its meaning would be more accessible. As a result, the match between this temporarily more accessible information and what was asked for in the question would take less time than a match between the question and the information previously looked up. They suggested that problem solution was influenced by a recency effect especially since the adjective in the first premise was not considered the relevant variable for solution. Unfortunately, they posed this theory after the body of their placement research was completed and only used it to explain Clark's 1969 (b) results with 13 subjects.

In a later discussion of models pertaining to the solution of positive comparative three-term serial problems, Clark (1971) felt that, despite its failings, the linguistic model was still superior in predictive value to the imagery model. After a reexamination of earlier findings had been made, Clark found that: 1) the linguistic model was able to predict how many as well as which errors would occur in solving the positive comparative forms, and 2) the principle of congruence was supported by solution time as well as error measurements. On the other hand, the imagery theory received relatively less support. This model could only predict the number of solution errors that would occur on such tasks, and had only solution time data to support

its principle of "temporary accessibility." Although Clark agreed that the linguistic model was incomplete because it did not adequately specify how the information given in the premises was eventually combined, he did not defer to the imagists. He concluded that while there was a weakness in his model, which even his principle of compression did not adequately explain, the fact that imagery could fill the "gap" had not been demonstrated.

Clark (1972a) also found evidence to support his theories by examining subjects' performance on a placement task. He proposed that when adults are instructed to insert an object into a physical array in a pure placement task, they must implicitly answer "Where should it be?" before they carry out a solution. In doing so, they did not seek the grammatical subject (as in Huttenlocher's studies) nor the most extreme member of the ordering (as in De Soto, London, & Handel's studies). Instead, their performance was influenced by the deep structural state of congruence between the sentence and the question relating to it.

Results from this placement task experiment, where subjects were presented with a display containing a one premise sentence on the left and an item on the right and told to indicate where the item should go, again, confirmed the linguistic theory. Clark found that congruence was the underlying principle for solution, and that positive comparative instructions were significantly easier than negative equative instructions (by an average of 843 msec.). Solution

by instructions with unmarked adjectives "high" or "good" were also significantly easier than those with marked adjectives "low" or "bad" (by an average of 229 msec.), and for all four types of sentences, subjects placed the "A" or first term significantly faster than the "B" term (by an average of 215 msec.). Error data from these two-term serial placement tasks were also in agreement with the latency patterns.

Relating this to model building, Clark proposed that the mean times for the placement task were accounted for by a simple additive model which required four parameters. Each of the following parameters was estimated from all the data by the least squares method: 1) a = 215 msec. was the time taken to "place B" over that for "place A"; 2) b = 229 msec. was the time taken by marked adjectives over that for unmarked adjectives; 3) c = 843 msec. was the time taken by negative equatives over comparatives; and 4) t_o = 2113 msec. was the base time taken up by operations not included in a, b, or c (or some combination of a, b, and c). Base time, therefore, depended on whether the condition involved "placing B," used a marked adjective, or was a negative equative instruction (or any other combination). Clark found that the model with the four parameters accounted for 99.4% of the variance amoung the actual means with their eight degrees of freedom. Similar results were also obtained with three-term serial placement tasks. Clark, again, concluded that the relative difficulty of the

various forms of negative equative problems differed from the relative difficulty of arranging real items according to corresponding negative equative instructions. These findings were, again, seen as evidence for the congruence principle and against the placement theory.

Keeney and Gaudino's (1973) study with a mixed problem type also provided further support for Clark's theory of congruence and lexical marking. Here, college students were asked to solve 32 taped three-term serial problems composed of one negative and one positive comparative premise (using "good/bad," "not as good/bad as," and vice versa). An analysis of the mean percentage of correct responses revealed that problems with unmarked adjectives were easier than corresponding problems with the marked adjective forms. For example, problems "A is better than B, C isn't as good as B" were solved more often than "B is worse than A, B isn't as bad as C"; and "B isn't as good as A, B is better than C" forms were solved more often than "A isn't as bad as B, C is worse than B." Likewise, Keeney and Gaudino found evidence that problems whose premises and questions were linguistically congruent were solved more often than problems that were incongruent. The authors further reported that subjects did not reorganize the marked negative equative problems, and found no support for the proposal that a match between the adjective of the second premise and the adjective in the question facilitated solution.

As a result of an experiment that tested whether the solution of positive and negative three-term serial problems was in part a function of subjects' verbal and spatial reasoning ability (measured by the Differential Aptitude Test), Zurawski (Note 3) also found evidence to support Clark's theory. When errors according to problem type were computed for 29 college subjects regardless of aptitude scores, the author found that the error sequence for negative problems did support Clark's principle of functional relations. Again employing the notation of Table 1, problems III and IV' with "taller/shorter" or "lighter/darker (haired)" were substantially easier to solve (97 errors) than problems IV and III' (141 errors), as were problems I and II' (87 errors) over problems II and I' (102 errors). (Error rates are totalled over problems regardless of adjective pair used.) The author also found that negative problems II' and IV' with the adjectives "taller/shorter" (39 errors) were easier than I' and III' (61 errors), as were problems II' and IV' with marked/marked adjectives "lighter/darker (haired)" (43 errors) over problems I' and III' (68 errors). On the other hand, since there was no difference in error rate between problem types using different adjective pairs, the specific adjective markings were seen as being unrelated to problem difficulty.

In review, Clark (also Keeney & Gaudino, 1973; Zurawski, Note 3) continues to find evidence to support the linguistic theory. In particular, congruence between the

underlying linguistic information or base strings of a problem is found to be a principle relatively strong in predictive value. Huttenlocher and Higgins' principle of temporary accessibility, on the other hand, while relevant to the solution of three-term serial problems, is seen as providing less power.

The Nature of the Adjective as it Relates to Serial Problem Solving

Researchers from both the spatial and linguistic camps believe that the adjective contained in a serial problem influences solution performance. The spatial theorists (e.g., De Soto, London, & Handel) hypothesize that subjects construct an imagined spatial array and that there is a natural preference for ordering problems based on the adjectives contained in the problem. Their initial work with positive comparative statements and three-term serial problems lends support to this theory. The linguists (e.g., Clark), on the other hand, propose a simpler encoding for adjectives which are linguistically unmarked rather than marked. Their model receives support from work with both positive comparative and negative equative problem types.

In an early investigation of spatial assignments, De Soto, London, and Handel (1965) tested how subjects imputed spatial directions to the adjectives "better," "worse," "lighter," and "darker." The authors asked 50 subjects to place the names from two-term relational or comparative statements (e.g., "A is better/worse than B"

and "A has lighter/darker hair than B") on paper with predrawn axes. Presentation order of the adjective in relation to its opposite was counterbalanced among subjects.

De Soto et al. found that subjects who received the statement with "better" as their first evaluative relation significantly placed the first name or better person on top of the paper (23 of 25 subjects) and proceeded downward. However, this finding did not hold true for problems with the adjective "worse." In this case, 19 of the 25 subjects who received "worse" as the first evaluative relation placed the first term on the bottom and proceeded from the bottom up. With regard to "lighter" and "darker" (hair), an altogether different pattern emerged. The authors found that these adjectives were not tied to one position on an array. That is, regardless of which statement was presented ("lighter" or "darker"), subjects tended to place the first name mentioned in the top or leftmost box. The authors concluded that evaluative relations such as "better" and "worse" were tied to a vertical array in "most people's thinking" but that the "lighter" and "darker" relations were not.

Handel, De Soto, and London (1968) also proposed that there were variations in how adjectives were presented on a physical array. Their contention was that adjectives such as "better" or "worse" were consistently tied to one position on an array (top or left and bottom or right, respectively), and that adjectives such as "lighter" or "darker"

(hair) were inconsistently placed (top or bottom, left or right). Carrying their 1965 finding into the world of the serial problem, the authors hypothesized that an adjective's spatial assignment could be useful in predicting the difficulty of various problem types. They argued that what was involved in problem solution was not a "fortuitous linguistic convention," but a glimpse at a person's inner construction.

Handel et al. also proposed that the solution of three-term serial problems which contained these spatially tied adjectives could be explained by two principles. These principles were directionality and end-anchoring and labelled "paralogical." For example, the authors hypothesized that problems which contained the same adjective and proceeded from a top to bottom direction would be easier than problems that went in the reverse (directionality), and that problems that were solved going from the end terms to the middle would be easier than problems solved from the middle to the ends (end-anchoring).

In the case of problems using adjectives which were inconsistently tied to a spatial array, the authors believed that the term in the first premise would determine the spatial assignment. Presented with the problem, "John has lighter hair than Bill, Bill has lighter hair than Tom," a subject would place "John" at the top (or left) and proceed downward (or to the right) in the array. A problem solver would encounter difficulty, however, if the adjective in the second premise was "darker" (hair). In this case, the

subject would have to transform the information in this premise to match the assignment he or she had already made by the first statement. As a result, problems I and II would be easier than III and IV.

Lastly, Handel et al. also investigated subjects' preferences for ordering a representation (top to bottom or left to right). The authors hypothesized that subjects who preferred to order in a top to bottom or left to right direction would perform better on the three-term serial problems that were ordered for solution in the same direction.

The 122 college students tested in this experiment solved 68 positive three-term serial problems using various adjective pairs (e.g., "better/worse," "earlier/later," "left/right"). Each problem was presented on a card with two premises, a question (e.g., "Is Tom faster than Bill?"), and three possible answers. Subjects were given 10 seconds to both view each problem and to circle the correct response. They were also given 12 problems to solve in a booklet with pre-drawn axes in order to assess their preference for ordering.

Based upon the percentage of correctly solved reasoning problems, the authors found evidence for both the principles of directionality and end-anchoring, and for the proposed solution strategies for problems with inconsistently tied adjective pairs. The hypothesis concerning the spatial axes assignment, however, was not borne out. The authors found that the pattern of premise combination difficulty did

not conform to that hypothesized for each type of spatial axis assignment. That is, although people using right to left solutions in their written arrays should have found problem II (e.g., "C is worse than B, B is worse than A") easier than problem I (e.g., "A is better than B, B is better than C"), they, in fact, did not. The authors concluded that subjects' spatial representations alone do not adequately tap the process of solving three-term problems.

Clark (1969b) argued that the crucial test against De Soto et al.'s theory of spatial imagery involved the adjectives "deeper" and "shallower." Clark chose this adjective pair because the spatial assignment and the lexical marking of each member led to contradictory predictions. Clark assumed that if the spatial model were correct, problems containing "deeper" would be harder to solve than problems containing "shallower." That is, problems with "deeper" would proceed from the bottom up and be harder than the top to bottom ("shallower") problems. If, however, the linguistic model were correct, the order of difficulty would be reversed. In this case, problems with the marked adjective "deeper" would be easier than problems with its lexically marked counterpart.

In order to test the predictions of these models, Clark (1969b) visually presented 21 adult subjects with positive comparative problems that dealt with people as having deeper or shallower wells. He found that his results were clearly against the spatial theory and in support of the principle

of lexical marking. He found that Type I problems using the unmarked adjective "deeper" were significantly faster to solve than Type II problems using the marked adjective "shallower" (e.g., "Jack has a deeper well than Dick, Dick has a deeper well than Tom" was significantly faster to solve than "Jack has a shallower well than Dick, Dick has a shallower well than Tom"). Clark concluded that the principle of adjectival marking was far superior in predictive value than the principle of spatial assignment.

Testing predictions from the research of both Clark and Handel et al., Jones (1970) explored the relationship between an adjective's lexical marking and its spatial assignment when presented in a three-term serial problem. She presented 10 (positive comparative and negative equative) three-term serial problems to 72 college students and allowed them as much time as needed for solution. The problems were broken into two parts and subjects were instructed to record the three names anywhere on a sheet of paper.

Jones found that an overwhelming majority (72%) of subjects preferred to arrange the three names or terms in a systematic order on a vertical or horizontal axis. (The remainder (28%) wrote out the names in the order presented.) She also found that the asymmetry of unmarked/marked adjectives affected how the names were hierarchically ordered. Examining the data of the 48 subjects who used a vertical axis to represent problems, Jones found that there were differences in how unmarked and marked adjective problems

were arranged. Whereas the "A" term in an "A is x than B" positive comparative problem was assigned top position in 100% of the cases when the adjective was unmarked (i.e., "better," "fatter," "thicker"), the "A" term was assigned to this position only 23% of the time when the adjective was marked (i.e., "worse"). In the case of differentially marked adjective pairs "lighter/darker" (marked/marked) and "lighter/heavier" (marked/unmarked), Jones found that the directional preference for "light" also depended on its semantic use. The "A" term was placed on top of the paper 71% of the time when the premise contained "light" (shade) and only 18% of the time when the premise contained "light" (weight).

Jones found that, in general, the consistent top to bottom spatial assignment of positive problems was no guarantee that this direction would hold for negative equative problems. For example, whereas the "A" term was placed on top 100% of the time by subjects making a vertical assignment, the "A" term in the "A is not as bad as B" problem was placed on top only 67% of the time. Jones indicated that subjects making a vertical ordering gave negatives a more arbitrary assignment and usually assigned the "A" term in the negative problems top position regardless of whether the term was described by an unmarked or marked adjective.

Since principles from both the spatial and linguistic theories can be used to predict subjects' spatial assignments, Jones ran a second experiment to test the predictive strength of these two models <u>re</u> solution performance.

Here, congruent three-term serial problems were presented on cards and solved within a 10 second limit by practiced subjects. Jones assumed that the spatial theorists would predict that problem III would be easier than problem IV because of its directionality (i.e., "A is not as bad as B" (top to bottom), "C is not as good as B" (bottom to top) vs. "A is not as good as B" (bottom to top), "C is not as bad as B" (top to bottom), respectively). The linguists, on the other hand, would predict no difference between solution speed because the underlying base strings of these forms are identical (each contain a "good" and a "bad" string). Jones found that her results did support the linguistic model (e.g., the mean solution times for these two problems were 10.5 and 10.4 secs., respectively), and concluded that the deep structural or linguistic theory could adequately account for the differences in serial problem solution.

In review, the spatial or imagery theorists believe that in the solution of deductive serial problems, subjects construct a spatial array in order to represent the information in a suitable format. In their experiments, they find that subjects demonstrate a preference to order terms in a top to bottom direction with the actual placement being determined by the spatial nature of the adjectives contained in the premises. Although initial work was done only with positively worded problems, Clark, and Jones' later research with negative equative problems indicates that the linguistic theory of adjectival marking is superior in predictive strength.

The Effect of Question Location and Time on the Solution of Three-Term Serial Problems

At the conception of the present experiment, the only theoretical attempt to synthesize the linguistic and imagery or placement models had been made by Johnson-Laird (1972). In offering suggestions for further research, he proposed that the effect of the question being placed before the premise should be explored in relation to serial problem solution. In a separate field of inquiry, Chase and Clark (1972) and Clark and Chase (1972) had found that encoding and comparison operations were different when subjects verified a sentence-picture and picture-sentence task. Potts and Scholz (1975), on the other hand, were interested in the effect of time on the solution process. Their experiment with positive comparative problems in which the question was presented after the premises showed that this variable was, indeed, important. Unlimited viewing time of the premises before presentation of the question led to a weakening in certain previously strong effects (e.g., congruence).

In the solution of three-term serial problems, Johnson-Laird assumed that initially subjects constructed a representation of the full three-terms (with or without imagery) and made do with the least amount of information as they worked back from a question. The author felt that once the subjects were practiced, they passed from this problem solving approach, analogous to the imagery theory, to one predicted by the more parsimonious linguistic model. In

terms of the negative problems, Johnson-Laird also suggested that if the negative equatives were transposed into comparatives (e.g., "A is not as good as B" into "B is better than A"), the major differences between the imagery and linguistic theories would be reconciled. He proposed that this reversal (Huttenlocher et al.'s third reorganization) would enable subjects to end-anchor and work down through the order in which the items were inserted into a spatial array, and also make the principles of end-anchoring and congruity coincide.

Johnson-Laird also attempted to explain why the imagery and linguistic theories predict the same results for positive comparative problems (even though subjects were supposed to be passing from an imagery representation to a linguistic strategy). The author contended that the above investigations were "almost certainly" based upon the performance of experienced subjects, and states, "No wonder that despite their divergent assumptions they have tended to converge upon the same empirical predictions. They are both likely to have miscalculated slightly, with the IMAGE theory failing to be sufficiently 'naive' and the LINGUISTIC theory failing to be sufficiently 'sophisticated'. Both are perhaps guilty of 'regression towards the mean', and our final task is to offer some suggestion on how they might be reformulated to correct this bias (p. 78)."

Johnson-Laird's proposal was to have the subjects' interpretation of a problem be guided by the nature of the question. The author, therefore, recommended the use of an

experimental procedure that would allow subjects to read a question before the two premises appeared. Johnson-Laird concluded that while it would be unnecessary for subjects to examine more than one premise in detail in a three-term problem, this procedure would reveal how subjects worked backwards from the conclusion of an inference to its premise.

Chase and Clark (1972) (and Clark & Chase, 1972) also performed an experiment which may be considered similar to a test for question placement. They presented different sentence-picture displays which conveyed positive or negative information in different orders and studied subjects' verification latencies. The format of their sentences was: "The (star/plus) (is/isn't) (above/below) the (plus/star)." The picture either showed a star above a plus sign or the opposite. The authors assumed that the sentences in the task were interpreted and encoded in terms of their deep structural propositions. For example, positive sentences were represented at the encoding stage more quickly than negative ones, as were unmarked sentences using "above" over marked sentences with "below." Likewise, pictures were thought to be ultimately represented in the same way as sentences. If a picture of a star above a plus sign appeared first (before the sentence), they hypothesized that it would be encoded as "star above plus." Conversely, if a sentence was viewed first, the encoding of the picture would be contingent on the sentence. Here, the picture would be encoded as "star above plus" if the sentence contained "above" and as "plus below star" if the sentence contained "below."

They proposed that various comparison operations were performed to translate the representations of the picture and sentence into a "symbolically equivalent" format before a match-mismatch judgment could be made. When a comparison was made of the sentence and picture, a subject compared the subject noun of the sentence (inner string) and the picture for synonymity in Operation 1. If there was a match, the subject went on to Operation 2. If there was no match, the subject changed the truth index associated with the sentence into the opposite value (Operation 1A). In Operation 2, the subject then compared the embedding strings of the sentence and picture to see if they matched (if neither contained a negative). If they did match, the subject responded. If they did not match (i.e., if the sentence contained "not" and the picture did not falsify the sentence), the subject changed the truth index by Operation 2A, and then responded.

The authors obtained times for these additive mental operations by estimation from the latencies for 12 subjects on 10 blocks of 16 sentences. They found that sentences with "above" took less time to encode than sentences with "below" (93 msec.). They also estimated the two mismatch (Falsification) and translation (Negation) operations to be 187 and 685 msec., respectively. They found that the "above/below" difference in encoding time plus the two mental operations accounted for 99.8% of the variance between the means of eight basic conditions.

Chase and Clark further manipulated the presentation

order of the stimuli. Students were asked to judge 16 sentence-picture displays by viewing: 1) first the sentence on the left and then the picture on the right, or 2) first the picture on the left and then the sentence on the right. In the latter condition, the authors assumed that subjects would attend to the picture first and encode the picture in the same way on every trial as: "star above plus." When the subjects attended to the sentence, they would encode it based on deep structural analysis. As a result, two independent codes would be established which the subject would have to make equivalent.

The authors found, as predicted, that viewing the picture before the sentence was more difficult in terms of time and errors than the reverse, and attributed this to an extra mental operation. They felt that when subjects viewed the sentence first, almost all the encoding of the picture was contingent upon how the sentence was encoded. When the subjects then viewed the picture, all that was required was to recognize a pattern and fill in the star or plus. On the other hand, when the picture was viewed first, subjects could not use this conditional encoding because two independent codes existed and had to be compared. An added operation was, therefore, needed in order to achieve an equivalence between the two codes (i.e., to insure that the subjects of the picture and sentence representations were identical). The authors concluded that both sentences and pictures have abstract mental representations which are

compared symbolically. However, they felt that deep structural meaning, rather than imagery, was the more basic representation for comparisons and durable memory.

Like Johnson-Laird (1972), Potts and Scholz (1975, Note 1) also wished to resolve the conflict between the placement (spatial or imagery) and linguistic theories. In an experiment that explored the element of time, they found that increased study time could affect the encoding and solution of positive three-term serial problems. Moreover, depending on the procedure used, support could be found for either theory. In their first experiment with 40 college students, 20 students were randomly presented with problems (on computer printouts) and timed from the simultaneous presentation of premises and question until solution (Clark's procedure). The remaining subjects were allowed to study the premises for as long as they wished before signaling that they were ready for the question. Once they signaled readiness, the problem question was presented and solution times were recorded. In general, Potts and Scholz found that the results from the first group supported Clark's model. Problems using premises with unmarked adjectives ("better") were significantly faster to solve than problems using the marked forms ("worse"), as were congruent over incongruent problem types. A different picture emerged, however, when data from the subjects given enough time to study the premises before answering the question was analyzed. Now, the authors found that the solution latency to the question "Who is best?" was shorter

than to "Who is worst?" regardless of which adjective was used in either of the premises. This was even true when the adjective in the two premises was "better" or "worse" ($p < .02$). With this unlimited time procedure, differences between problems using the adjective "better" and those using "worse" were just barely significant. Potts and Scholz, therefore, reported that when there is no time pressure to achieve a solution, subjects store marked and unmarked adjectives in the same form.

The authors concluded that whereas marked and unmarked adjectives differ at a preliminary level of processing, given sufficient study time, subjects reduce both marked and unmarked forms to a "common unmarked form." On the other hand, in the condition where additional study time was not allowed and timing began with the appearance of the stimuli, the authors found that solution latencies were significantly quicker when the adjective in the premise was congruent with the adjective in the question than when it was not. Potts and Scholz (1975) also concluded that neither Clark nor Huttenlocher are totally accurate in their models of how subjects solve the three-term serial problem. What, instead, appears to be critical to solution is the congruence between the adjective used to describe a term and the desired placement of that term in the ordering.

In sum, Johnson-Laird's theory of serial problem solution and Chase and Clark's work on sentence encoding indicate that location of the question in a serial problem would be

worthy of investigation. Given Huttenlocher and Higgins proposal that problem solution is enhanced when the adjective in the question and second premise match, Johnson-Laird believes that investigating a question before premise condition would reveal how subjects encode a serial problem in reverse. Results from Chase and Clark's experiments with sentence verification tasks also show that encoding operations for sentence/picture and picture/sentence displays differ as a function of which display is attended to first. With respect to display time, Potts and Scholz emphasize that unlimited study time does affect subjects' encoding process. In particular, the strength of the congruence principle declines with increased exposure of the problem before presentation of the question. Results from the procedure which simultaneously presents the total problem, on the other hand, tend to support these same linguistic principles. Whether serial problem solution can totally be explained by the linguistic model or the placement model, however, is not resolved.

Conclusions

In summary, then, the area of controvery between theories has been the three-term serial problem. The linguists propose three principles to account for both positively and negatively stated problems' solution. Derived from the study of deep grammatical structure, these principles pertain to functional relations, adjectival marking, and congruence. The linguistic camp also believes that the negative is composed of one more feature or operation than the positive or kernel statement.

The linguistic theorists consistently find that solution is easier and faster for positive problems as compared with negative problems, for congruent problems as compared with incongruent problems, and for unmarked adjective problems as compared with marked adjective problems.

The placement (spatial or imagery) theorists, on the other hand, hypothesize that a spatial array is involved in solution of positive three-term problems and that the spatial assignment of certain adjectives affects how subjects order these terms in "cognitive space." The placement theorists also propose a correspondence between how subjects solve the serial reasoning problems and placement tasks. They even carry this hypothesis into the realm of the negative equative forms, and further contend that subjects make certain symbolic reorganizations to understand the negative. This process is said to be affected by the nature of the adjective dimension and to involve a subject-object reversal if the adjective implies a negative (e.g., "A is not as bad as B" to "B is worse than A"). Historically, there is support for their theory with positive comparative problems. However, less evidence for their hypotheses is produced when the negative equative form is examined.

III

Present Study

Since the previously described controversy has not adequately been resolved, an investigation of the two-term serial problem seemed a better way to assess the cognitive demands made on the problem solver than the convention three-term serial problem. A review of the literature shows that the two-term problem has rarely been studied in and of itself, although it is the building block of all higher order serial problems.

The first order of events, then, is to explain the structure of the two-term serial problem. Since the body of evidence presented so far tends to give the greatest support to Clark's linguistic model, the following structural variables were investigated: 1) positive or negative statement of the problem (POSNEG); 2) compatibility or congruence of the underlying base strings (COMPAT); and 3) the unmarked over the marked adjective type (ADJECT).

Following Johnson-Laird's (1972) suggestion, the effect of question location on problem solution was also examined to see if problem solution would be facilitated or inhibited if the question appeared before (QPRE) or after (QPOST) the premise. Moreover, Potts and Scholz's (1975) finding that the amount of study time affected three-term serial problem solution also seemed applicable to and worth investigating in the simpler one premise form. Following a presentation

time procedure used by Posner et al. (1969) in their study of letter encoding, exposure time of the premise before presentation of the question as well as question before appearance of the premise was tested under four durations (0, .5, 1, and 2 seconds). When two pieces of information were shown directly in sequence (hereby designated as the zero delay), it was hypothesized that no difference between the two question placement conditions would exist. However, when subjects were given additional time to read, encode, process, or in some way prepare the premise for the question, it was felt that differences between the two conditions would become apparent. This was defined as the time variable (TIMES).

The relative contribution of the three structural variables was also assessed under the different exposures for each question location condition. Four time durations were used in the question before and question after premise conditions (0, .5, 1, and 2 seconds). However, a longer fifth time (4 seconds) was also introduced in the question after premise condition. It was assumed that data from this longer duration would show a general weakening in the strength of the variables and give some indication that subjects were now encoding the premise in preparation for either question. Conversely, since subjects were only allowed to view and encode the question in the question before premise condition (before the premise appeared), it was felt that the effects of the variables would remain constant over time.

Hypotheses

In sum, the hypotheses concern the structure of the problem, the effect of the question placement, and the pre-exposure duration of the question or premise. As to the structure of the problem, and in terms of solution times and errors, it was assumed that: (a) positive comparative problems would be easier to solve than negative equatives; (b) congruent problems would be easier than incongruent problems; and (c) problems containing the unmarked adjective "better" or "not as good as" would be easier than those with the marked adjective "worse" or "not as bad as." It was also hypothesized that the problems which seem to imply a double negative ("not as bad as") would be the hardest problems to solve.

With regard to question placement, it was hypothesized that question location would have no significant effect on the ease of solving problems at the zero delay. That is, when two pieces of information appeared in immediate sequence, there would be no significant advantage of the premise first/question last condition (QPOST) over the question first/premise last condition (QPRE). However, with longer durations of the first piece of the problem, it was hypothesized that the placement of the question would have a differential effect depending upon which piece of information appeared first. It was, therefore, assumed that: 1) problem solution times and variable effects would remain fairly constant over time in the QPRE condition; and 2) longer durations of the pre-

mise before presentation of the question would lead to shorter solution times and a decrease in the effect of the variables in the QPOST condition.

It was also predicted that in the question after premise condition, the rank ordering of the problem types on ease of solution would change over time. That is, whereas problems that are both positive and congruent and contain the unmarked adjective form would reach a minimum solution time almost immediately within the shorter exposure durations, problems with the opposite characteristics would reach the same asymptote or baseline solution time within longer durations of the premise. This effect would result in less stable correlations between problems as they were solved under the shorter and longer duration conditions. Conversely, within the question before premise condition, rank orderings across all durations would remain highly stable.

It was also assumed that subjects would report individual differences in method of problem solution (Störring in Woodworth & Schlosberg, 1954). During a preliminary investigation, subjects were found to report solution by either a "meaning" or "structural" approach (i.e., subjects relied on the structural properties of the different problem types to achieve solution). It was, therefore, hypothesized that there would be a significant difference in solution times on problems representing the various variables between these two groups.

IV

Method

The Two-Term Serial Problems

Two-term serial problems were chosen to represent the three linguistic variables under study. Four basic premise types were used ("A is better than B"; "A is worse than B"; "A is not as good as B"; and "A is not as bad as B") and examined with respect to their wording (positive/negative), the congruence between the adjective in the premise and in the question, and the nature of the adjectival expression. These factors are defined as POSNEG, COMPAT, and ADJECT, respectively. Table 2 presents the eight possible arrangemetns coded according to these factors in the question after premise condition (QPOST). Additionally, all eight arrangements appeared with the question before the premise (QPRE). Each premise appeared equally often with Sam and Joe as the subject, thereby resulting in each of these names being correct or incorrect half the time. The coding of this factor, named SUBOBJ, was positive or negative depending on whether Sam or Joe was the subject of the premise.

The interval between the presentation of the question and premise (QPRE) was either 0, .5, 1, or 2 seconds. The interval between the presentation of the premise and the question (QPOST) was either 0, .5, 1, 2, or 4 seconds. These nine conditions of question location and display delay were crossed with each of the two forms (Sam or Joe

Table 2

Eight Basic Problem Types

Coded on Three Factors

		POS-NEG	COM-PAT[1]	AD-JECT[2]
(Sam/Joe) is better than (Joe/Sam)	Who is best?	+	+	+
(Sam/Joe) is worse than (Joe/Sam)	Who is worst?	+	+	-
(Sam/Joe) is better than (Joe/Sam)	Who is worst?	+	-	+
(Sam/Joe) is worse than (Joe/Sam)	Who is best?	+	-	-
(Sam/Joe) is not as good as (Joe/Sam)	Who is best?	-	+	+
(Sam/Joe) is not as bad as (Joe/Sam)	Who is worst?	-	+	-
(Sam/Joe) is not as good as (Joe/Sam)	Who is worst?	-	-	+
(Sam/Joe) is not as bad as (Joe/Sam)	Who is best?	-	-	-

[1]Where "+" indicates base string congruence and "-" indicates incongruence.

[2]Where "+" indicates unmarked adjective and "-" indicates marked.

as subject) of the eight problem types in Table 2 to yield 144 stimuli. This pool was divided into two randomizations (A and B) subject only to the constraints that no problem type (Table 2) ever occurred twice in a row or in the same location in the two randomizations. Thereafter, each randomization was subdivided into a first and second half (i.e., A1, A2, B1, B2) and programmed into a PDP-8E computer. Any subject received one of four equally likely test orders: A1A2, A2A1, B1B2, or B2B1.

Subjects

Subjects were volunteer undergraduate and graduate psychology students affiliated with the City College of The City University of New York. There were 32 subjects in total, including 16 women and 16 men. One male was later dropped from the study due to poor performance level. Twenty seven of the 31 subjects were right handed and each subject was aware that the experiment involved problem solving.

Apparatus

PDP-8E Computer. The PDP-8E computer was programmed to emit each problem on a Tektronix 4010-1 Telex screen from a blinking dot in the upper left hand corner of the screen. The computer was programmed to output the first part of the problem (which remained on the screen), execute the appropriate exposure, and then output the second part of the problem (question/premise or premise/question) on the same line. The computer recorded the response latency in centi-

seconds from the appearance of the second part of the problem on the Telex up until the subject responded on the Telex typewriter. If the subject did not respond to the problem, the computer was programmed to wait 10 seconds and then erase the entire screen. If the subject responded within that period, the computer recorded the time, erased the screen, waited 5 seconds, and emitted the next problem. After completion of each half randomization (72 stimuli), the computer typed out the words, "YOU MAY REST NOW." Attached to a Telex screen, this computer emits characters at an approximate rate of 1,200 per second in a sweep from left to right. The characters appear in white on a green background.

Telex Screen. The Tektronix 4010-1 Telex screen is similar to a 9" long x 6 3/4" high (22.86 cm x 17.14 cm) television screen and appears above a typewriter keyboard. Initially, a blinking dot in the upper left hand corner indicates that the screen's power is switched on. In this experiment, the dot also indicated to the subject that a new trial would begin. When the subject responded to a problem, he/she pressed either the "S" key for Sam or the "J" key for Joe on the Telex keyboard. This character and the solution time were stored in the PDP-8E computer and retrieved after each experimental session.

Procedure

Each subject was requested to sit in front of the

apparatus. The subject was told that this was a problem solving experiment on the relationship between two people, Sam and Joe. When Sam was the answer, the subject was instructed to press the "S" key and when Joe was the answer, to press the "J" key. Although accuracy was stressed, a subject was told not to worry if an answer was missed or a wrong key hit. The experimenter suggested that a subject's fingers be kept on the "S" and "J" keys in readiness to respond.

The computer tape for the first half of a randomization was manually loaded into the computer and the experiment began. There were no practice trials. When this block of problems was completed, the subject was given a 3 minute rest during which time the second half-randomization was loaded into the machine. At the completion of this block of problems, the subject was asked about the strategy used to solve the problems (i.e., if he/she relied on meaning or the structural characteristics of the problems to obtain an answer). The duration of the experiment was approximately 30 minutes. The computer was then cued to emit the response latencies and answers on a sheet of paper.

V

Results

Adjustments to Data

The first subject tested, a right handed male, was omitted from the data analysis due to a relatively large percentage of errors (22%) on the task. The remaining 31 subjects had error rates of 10% or less on the 144 problems, and across all problems for all subjects, only 4% of the responses were scored as incorrect. Although missing cases (where a subject did not complete a problem in 10 seconds) were included in this ratio, these were rare and occurred only 14 times across all subjects and all problems. However, it was evident that a subject's initial encounter with a set of problems either led to a missing observation (4 of the 14 cases) or an extremely long solution time. Therefore, in order to prepare the data for analysis, two adjustments were made. First, for every subject, an overall mean correct solution time was computed for each succeeding quartile of the data (36 stimuli). This mean then replaced the times for any stimulus in the quartile on which there had been an error or a miss. Secondly, to adjust for the "warmup effect" or long solution times upon encounter with the problems, for any subject, the latencies for each of the first three problems in a test order were replaced by the mean latencies for those problems in the other three test orders, where invariably the stimuli in question

occurred late in the series.

The adjusted raw data for 31 subjects were then analyzed on an IBM computer by means of a Data-Text analysis of variance program for repeated measures (Armor & Couch, 1972). All analyses were performed both on raw latencies as well as on log transforms of the latencies. In no case did the latter analyses differ from the former, whence all results are reported in linear units. A 2x2x2x2x4 analysis (POSNEG x COMPAT x ADJECT x SUBOBJ x TIMES--0, .5, 1, and 2 seconds) was performed separately on the question before premise and question after premise conditions. (To achieve compatibility between the QPRE and QPOST analyses, the 4 second QPOST condition was not included.) For both these question placement conditions, the POSNEG x COMPAT interaction (PO x CO) and all the main effects (POSNEG, COMPAT, ADJECT, and TIMES) were significant at $p < .001$. These findings confirm the hypotheses relating to problem structure which are that positive comparative problems are faster than negative equatives, congruent problems are faster than incongruent problems, and problems with the unmarked adjective are faster than problems containing the marked adjective. Sam-first problems were not significantly faster than Joe-first problems (SUBOBJ, $p > .05$), and subsequent analyses were performed with this variable averaged over problems.

Solution Times

Question Placement and Problem Structure. In order to assess the contribution of the variables when the SUBOBJ effect was averaged over problems, a 2x2x2x4 analysis (POSNEG x COMPAT x ADJECT x TIMES--0, .5, 1, and 2 seconds) was performed again separately for QPRE and QPOST conditions. Comparing the question placement conditions, the contribution of the main effects was found to vary. As shown by F-magnitudes in Table 3 and percentage of explained variance in Table 4, POSNEG, COMPAT, and PO x CO have a greater influence on solution times in QPRE than they do in QPOST, whereas the reverse is true for ADJECT and TIMES. Additionally, only in QPOST do we see significant interactions between time and problem structure. Finally, only in QPRE is there evidence of an additional structural interaction (POSNEG X ADJECT), albeit a weak one.

All posthoc analyses were performed using the Scheffe test at the $p < .01$ level. In the question before premise interactions, results pertaining to the structure of the problem were as follows: 1) In the PO x CO interaction, all means were significantly different from one another so that in rank order from the fastest to slowest they were: positive-congruent (1.66 sec.), positive-incongruent (2.52), negative-congruent (2.77), and the negative-incongruent problem (3.01). 2) In the POSNEG x ADJECT interaction, the "double negative" form of the problem ("not as bad as") was indeed the slowest problem type (3.03 sec.), as predicted.

Table 3

Results of F-Tests for Main Effects and Significant Interactions in Two Question Placement Conditions

	QPRE (0,.5,1,2 sec.)	QPOST (0,.5,1,2 sec.)
POSNEG	141.867	124.259
COMPAT	127.002	86.709
ADJECT	19.562	53.060
TIMES	46.792	260.646
PO x CO	48.834	23.269
POSNEG x ADJECT	4.834	
POSNEG x TIMES		5.003
COMPAT x TIMES		7.947
ADJECT x TIMES		6.161
POSNEG x ADJECT x TIMES		3.290

Note. Degrees of freedom are 3, 30 for TIMES and all TIMES interactions, and 1, 30 for all remaining tests.

Table 4

Percentage of Total Sum of Squares for Main Effects and Significant Interactions in Two Question Placement Conditions

	QPRE	QPOST
POSNEG	13.71***	7.85***
COMPAT	6.41***	3.60***
ADJECT	.74***	.89***
TIMES	6.27***	22.72***
PO x CO	2.09***	1.42***
POSNEG x ADJECT	.15*	
POSNEG x TIMES		.39**
COMPAT x TIMES		.54***
ADJECT x TIMES		.36***
POSNEG x ADJECT x TIMES		.24*

Note. Four delays were analyzed in each question condition.

*$p < .05$.

**$p < .01$.

***$p < .001$.

The positive unmarked (2.04) and marked (2.14) adjective forms were the fastest, while the "not as good as" form (2.76) was intermediate in speed.

An examination of the interaction relating to problem structure in the question after premise condition showed that within the PO x CO interaction, the positive congruent problems were the fastest (1.79 sec.), and the positive incongruent (2.44) and negative congruent (2.63) forms were intermediate in speed. The slowest problem types were the negative-incongruent (2.77) problem forms, but they did not differ significantly from the negative congruent forms.

Display Delay. As seen in Table 5, the question before premise condition, as predicted, was less differentiated than the question after premise condition in the posthoc analyses of the TIMES effect. When the question preceded the premise, the problems solved under the 2, 1, and .5 second durations were not significantly different from one another, but they did differ significantly from problems solved under the zero delay. Conversely, when the premise preceded the question, the problems solved under the 2, 1, .5, and zero durations were significantly different from one another. Here, problem solution time increased with shorter delays between premise and question. Moreover, as predicted, a separate analysis of variance comparing the zero delay problems in the QPRE and QPOST conditions proved insignificant (F (1, 30) = 2.406, $p > .05$). That is,

Table 5

Mean Solution Times for Problems Solved Under Different Display Delays in Two Question Placement Conditions

Exposure delay in seconds	QPRE	QPOST
0	2.94	3.05
.5	2.45	2.67
1	2.31	2.18
2	2.25	1.72
4	*	1.56[a]

Note. Means are in seconds.

[a]Mean provided for visual comparison only.

*Not tested in this condition.

problems which were presented as premise/question and question/premise in immediate succession were not significantly different from one another.

Table 6 presents the data required to explicate the significant time interactions in the question after premise condition (POSNEG x TIMES, COMPAT x TIMES, ADJECT x TIMES). Scheffe tests revealed that under the 2 second delay, the positive, the congruent/incongruent, and the unmarked/marked adjective problems were fastest in solution time, whereas the negative, incongruent, and the unmarked/marked problems, respectively, under the zero delay were the slowest.

The varying advantage of positivity, congruence, and the unmarked adjective over time, however, was more clearly demonstrated with five separate 2x2x2 analyses of variance in the QPOST condition. Here, POSNEG, COMPAT, and ADJECT were tested under each of the 0, .5, 1, 2, and 4 second display conditions in the question after premise condition. Results of these analyses are shown in Table 7 in the form of percentages of explained variance. These analyses show that the variable effects behaved differently depending on the amount of time allowed for pre-exposure of the premise before presentation of the question. Whereas the advantage of a problem being positive and congruent peaked at the .5 second duration, the benefit of an adjective being unmarked was realized only in the 1 and 2 second delays. Unlike the main effects, the PO x CO interaction remained fairly constant across these same durations.

Table 6

Mean Solution Times for Problems Solved According to Variable Dimensions in Question After Premise Condition

	Display Delay in Seconds				
	0	.5	1	2	4[a]
Variable Dimension					
Positive comparative	2.73	2.29	1.91	1.52	1.46
Negative equative	3.37	3.06	2.44	1.93	1.66
Congruent	2.80	2.40	1.98	1.65	1.54
Incongruent	3.30	2.95	2.37	1.80	1.58
Unmarked adjective	3.01	2.62	1.98	1.61	1.52
Marked adjective	3.08	2.73	2.37	1.84	1.60

Note. Means are in seconds.

[a]Means provided for visual comparison only.

Table 7

Percentage of Total Sum of Squares for Main Effects and Significant Interactions in Separate Time Conditions

Delay in sec.	QPRE[a]	QPOST				
	-	0	.5	1	2	4
POSNEG	19.74***	12.17***	13.53***	8.19***	6.95***	1.94*
COMPAT	9.23***	7.50***	7.03***	4.38***	.93*	.10
ADJECT	1.07***	.16	.23	4.51***	2.08***	.36
PO x CO	3.01***	2.18**	1.69**	1.74*	1.87**	.09
POSNEG x ADJECT	.21*		.89*			.69*
Total Explained Sum of Squares	33.35	22.70	23.66	19.80	12.20	3.49

[a]The numbers for this condition were obtained from a separate analysis of variance where the 0, .5, 1, and 2 second display delays were summed and averaged over the QPRE condition, and are provided for visual comparison only.

*$p < .05$.

**$p < .01$.

***$p < .001$.

The inclusion of the 4 second QPOST delay, indeed, demonstrated that the effects of the variables changed as a function of time. In this condition, the POSNEG effect was the only main effect that remained significant. Moreover, this significance was weaker ($p < .05$) than under the shorter delays ($p < .001$). The POSNEG x ADJECT interaction was also significant at $p < .05$. All other main effects in the 4 second duration ceased to account for a significant share of the variance.

Error Rate

There were only 176 errors in 4464 responses (31 subjects x 144 problems). This small data base did not lend itself to statistical testing, yet comparisons of partitions of this sum were informative and tended to confirm the latency analyses. As seen in Table 8, positive comparative problems were easier than negative equative problems, congruent problems were easier than incongruent problems, and unmarked adjective problems were solved with less errors than marked adjective problems. Comparing QPRE and QPOST conditions over the display times they shared, we observe that congruent and unmarked problems were solved with a fairly equal number of errors, whereas negative equative, incongruent, and marked adjective problems were solved with fewer errors in the QPOST than QPRE condition. Positive comparative problems, however, were somewhat easier in the QPRE condition. Generally, all problems solved under the

Table 8

Error Rate on Problems as a Function of Variable Dimension

	Total	QPRE (0,.5,1,2 sec.)	QPOST (0,.5,1,2 sec.)	QPOST (4 sec.)
Variable Dimension				
Positive comparative	54	20	26	8
Negative equative	122	64	52	6
Congruent	67	29	31	7
Incongruent	109	55	47	7
Unmarked adjective	68	30	33	5
Marked adjective	108	54	45	9

QPOST condition were slightly easier than those under the QPRE condition (78 vs. 84 errors, respectively).

Ordering of Problems

The hypothesis that the ease of ordering the eight problems would change as a function of the time variable in the question after premise condition was not confirmed. Spearman's correlational test for nonparametric data was used to assess changes in the ordering of how quickly each problem was solved under both question conditions (nine times in all). As seen in Table 9, the ordering of problems on ease of solution remained significantly stable across all times and for both question locations, faltering only with the 4 second delay in the QPOST condition.

Individual Differences

The hypothesis that differences would emerge in the problem variables between subjects who reported solving the two-term serial problems by meaning or structure also received some support. In response to the question about solution strategies, 17 subjects indicated they relied on meaning (M), 8 subjects indicated they relied on the problems' structural properties (S), and 6 subjects responded that they employed both methods to achieve solution (MS).

Five possible comparisons of the M and S, and MS subjects were made: 1) M vs. MS; 2) M vs. S and MS; 3) S vs. M; 4) S vs. MS; and 5) S vs. M and MS. Although t-tests comparing the mean solution times for each of these

Table 9

Correlation Matrix for Problems Rank Ordered On Ease of Solution Across All Display Delays in Two Question Placement Conditions

Delay in sec.	QPRE 0	QPRE .5	QPRE 1	QPRE 2	QPOST 0	QPOST .5	QPOST 1	QPOST 2	QPOST 4
QPRE									
0	-	.905***	.976***	.905***	.905***	.833**	.738*	.762*	.762*
.5			.952***	1.000***	.881**	.905***	.881**	.881**	.619
1				.952***	.952***	.905***	.883**	.810**	.667*
2					.881**	.905***	.881**	.881**	.619
QPOST									
0						.833**	.883**	.714*	.571
.5							.810**	.786**	.357
1								.929***	.548
2									.667*
4									-

*$p < .05$.

**$p < .01$.

***$p < .001$.

pairs on the main effects and PO x CO interaction in the QPRE and QPOST conditions were performed, no significant differences were found. Included in this analysis, however, was a contrast of the range of effects of two of the main structural variables. This was measured as the diagonal of the PO x CO interaction (i.e., mean solution time of negative-incongruent problems minus the mean solution time of the positive-congruent problems) as shown in Table 10. For this range of effects, a significant t-test was obtained for the following comparisons: S vs. M; S vs. MS; and S vs. M and MS in both QPRE and QPOST conditions. The results of the t-tests for structure vs. any meaning (i.e., S vs. M and MS) are here recorded. Differences in the diagonal for QPRE were significant at $p = .03$ (t (29) = 2.30), and differences in the diagonal for QPOST were significant at $p = .02$ (t (29) = 2.55).

Additionally, it was found that female subjects were less influenced by the PO x CO variables, with mean diagonal effects of 2.24 sec. and 3.20 sec. in QPRE and 1.69 sec. and 2.27 sec. in QPOST, respectively. For QPRE, t (29) = 2.76, $p = .01$; for QPOST, t (29) = 2.29, $p = .03$. Both findings with regard to method and sex should be viewed with caution, however, since in our sample the two effects are confounded by the fact that all but one of the structure-only users were females. Indeed, when the 9 females and 14 males who reported using meaning or meaning plus structure were intercompared, there were no significant differences in PO x CO

Table 10

Mean PO x CO Effects for

Three Groups of Subjects

		Solution Method		
		Meaning N = 17	Meaning + Structure N = 6	Structure N = 8
(A) Negative-incongruent problems				
	QPRE	6.33	5.95	5.43
	QPOST	5.81	5.67	4.90
(B) Positive-congruent problems				
	QPRE	3.41	2.90	3.42
	QPOST	3.73	3.29	3.47
(A) - (B)				
	QPRE	2.92	3.05	2.01
	QPOST	2.08	2.38	1.43

Note. Means are in seconds.

diagonal effects (2.53 sec. and 3.22 sec. in QPRE and 1.84 sec. and 2.37 sec. in QPOST, respectively). Thus, the two preceding findings seem wholly due to the seven females who reported relying on the structural properties of a problem for solution. This is confirmed by the fact that when those 7 are compared against the remaining 23 subjects, the results were significant in both the QPRE and QPOST conditions. In sum, then, some aspects of sex and solution method are both relevant to problem structure but with this small sample the two cannot be teased apart.

A discussion of these results as they pertain to the processing and solution of two-term serial problems now follows.

VI

Discussion

Since there is such controversy over the three-term serial problem, the simpler two-term serial problem was investigated in order to examine the strategies and encoding processes that are involved in solution. Aside from the general findings, which are in strong support of Clark's linguistic principles, the data reveal that subjects' encoding strategies do change as a function of problem type, question location, and pre-exposure delay. In particular, performance is facilitated with initial pre-exposure of the question before premise; longer pre-exposures of the premise before question; and when a positive problem is presented.

In this experiment, both solution time and error data confirmed the major hypotheses relating to problem structure and lent support to the principles of congruence and adjectival marking as proposed by Clark. Likewise, the use of two question placement conditions demonstrated that depending on how subjects were cued, these factors (COMPAT, ADJECT) contribute differentially to the solution process. Since COMPAT (as well as the PO x CO interaction) had a relatively greater effect in the question before than question after condition, one can conclude that the question in some way acts as a prime for the problem solvers in their search through the premise for the response. And, that knowing

the question in advance somehow reduces the encoding of the premise, especially when the underlying base strings between the two parts of the problem match.

The use of different exposure delays also indicated how the relative contributions of the main effects changed with time. In the question before premise condition, the value of the question appearing first (e.g., priming) occurred only in the problems solved under the .5 second duration of the question before premise. When the durations were increased (1 and 2 seconds), however, no further benefit in solution time resulted. Conversely, in the question after premise condition, the magnitude of effects varied and interacted with and across all times.

The above findings indicate that subjects' encoding of certain solution variables changed within 2 seconds after presentation of a problem. It appears that in the QPRE condition, subjects discovered a heuristic that aided their encoding of a premise and search for an answer. This strategy for solution, however, was a constant factor and no further benefit was derived with longer exposures of the question before presentation of the premise. A clearly different picture became apparent in the question after premise condition. Not only did the benefits of the significant main effects behave differently under the various delays, but the relative effects of the variables decreased as a function of increased exposure of the premise before presentation of the question. The behavior of the adjective

effect was especially interesting. It was found that whereas the significant benefits of the POSNEG and COMPAT factors had already peaked by .5 seconds, the ADJECT effect only became significant under the longer 1 and 2 second durations. This indicates that, in this experiment, the adjective dimension either: 1) required relatively more processing time than both the logical "not" and the linguistic match for congruence, or that 2) subjects relied on the structural factors to solve problems until there was enough time to read, digest, and process the meaning of the premise (adjective) before the question appeared (i.e., in the 1 and 2 second delays). With enough time for processing, then, the nature of the adjective (whether it is unmarked or marked) became a strongly relevant factor for solution.

The finding that ADJECT only became significant at later QPOST exposures also speaks against Potts and Scholz's (1975, Note 1) theory that with unlimited processing time, all adjectives are stored in the unmarked form. Although they found that answering "Who is best?" was significantly quicker than "Who is worst?" for problems containing "better" or "worse" when additional study of the two premises was allowed, in this experiment: 1) the unmarked adjectives were significantly easier than the marked adjectives at 1 and 2 seconds (QPOST) and 2) the difficulty of an incongruent "A is worse/not as bad as B. Who is best?" problem held throughout.

It is even possible, given the difficulty of solving the marked negative, the "double negative," or implicitly negative problem (e.g., see Huttenlocher et al., 1970; Wason & Johnson-Laird, 1972), that at short durations subjects may even prefer to rely on congruence or other time saving heuristics before they tackle the full processing of the adjective. In attempting solution, subjects might even rapidly scan the problem in an effort to pick out the phrase that would produce a close structural and visual match. On the other hand, at longer durations of the premise, the subject is literally "forced" to encode the adjective before the question appears. In this case, the problem solver may have to rely on the meaning and marking of the adjective to achieve solution. This might speak for a structure to meaning switch in the encoding process for a given problem rather than, as Johnson-Laird has proposed, a result of practice over the duration of the experiment.

Although the main effects, including ADJECT, begin to lose their significance when more time is allowed for premise preparation at 4 seconds, POSNEG and the POSNEG x ADJECT interaction remain weakly significant at $p < .05$. Whereas this weakening of effects over the fixed presentation times (display delays) indicates that subjects prepare the premise in such a way that a response to either question is readily available, blocking to the cognitive processing of the negative is still evident. Given the assumption that four seconds should be ample time to encode and process a

two-term serial problem, it may be that amount of exposure time does not effect solution of the negative combined with a marked adjective ("not as bad"). What then might be needed to explain this finding is a combination of linguistic (adjective type), logical ("not as x as"), as well as psychological factors ("cognitive blocking").

Further support for the structural properties of the negative having an enduring effect throughout the length of the experiment is the relative stability of the problems' difficulty or ease over eight of the nine exposure durations and question placement conditions. In the ninth, or 4 second delay, there is a slight hint of deviation from this trend. These results indicate that whereas the main effects lose their strength over time in the question after premise condition, the inherent structure of the problem continues to exert an influence over how quickly each problem is solved even when four seconds is allowed for problem solvers to pre-study the premise. Unfortunately, because of significant interactions among the problem dimensions in the QPOST condition, there was no clear way to map out time savings across problem types as a function of duration. Moreover, the irregularily that occurred when this was attempted spoke against a simple linear model.

The existence of individual differences between subjects reporting that they solved problems by meaning or by a heuristic involving the structural properties of the problems lends support to the work of Quinton and Fellows (1975).

They found some significant differences in solution times of subjects reporting having used either a meaning or perceptual strategy in the solution of positive three-term problems. Since solution of the two-term problem provides even less of a test for the imagery, spatialization, or structural manipulation approach (a middle term is not consolidated as with three-term problems), it is surprising that introspective reports alone were able to differentiate subjects on the diagonal effect of the PO x CO interaction. This result indicated that subjects who reported having used structure were less influenced by the experimental variables of POSNEG and COMPAT than subjects who reported having used meaning. This finding is also viewed as additional evidence for individual differences in problem solving and especially for serial problem solution. (See Shaver, Pierson, & Lang, 1974; Sternberg, Note 4; Zurawski, Note 3 for studies of individual differences on three-term serial problems as measured by verbal and spatial aptitude tests.)

Of equal interest, is the unexpected finding that, in this experiment, women were relatively better problem solvers than men in terms of speed of solution. This disagrees with Shaver et al. (1974) who have reported that while college women do not differ significantly from men in overall performance on serial problems, they do score lower on spatial tests as a function of some problem with motivation. Although there was no spatial aptitude test administered here, both the women and men were undergraduate

and graduate psychology majors. This selection of readily available volunteers may have biased the sample in favor of more spatially oriented or highly motivated individuals. On the other hand, the experimenter was female and the idea that women would lag behind men was not even entertained. Whether sex differences or motivational factors are related to the sex of the experimenters on such tasks would be an area for further research. It seems worth investigating why level of motivation would affect a spatial ability test and not a spatial task. The fact that the present findings with regard to method and sex were both probably due to the behavior of females using structural solution methods also argues for future studies pre-screening subjects for their proclivities. Counterbalancing the sex of subjects and experimenters might also be introduced before sex difference results were published, to control for any possible overt or hidden sex bias.

In general, results from this experiment suggest that display delay is an important variable in solution of serial problems. The results indicate that further research with two- and three-term serial problems using a similar procedure could examine at what point in the solution process various effects, as defined by the opposing models, gain or recede in magnitude. The author believes that the most recent exploration of the three-term serial problem's solution (Sternberg, 1978; Sternberg, Guyote, & Turner, 1978; Sternberg, Note 4) which includes a simultaneous

presentation of question/premise and premise/question as well as the unlimited time procedure of Potts and Scholz (1975) should also study the fixed display delays of question and premise(s) and vice versa. In this way, support for a mixed solution model (either Sternberg's linguistic to spatial) or the author's structural to meaning model could be tested when subjects had fixed times to encode the component parts of serial problems.

In conclusion, the study of performance on the two-term serial problem did reveal that subjects employed different heuristics or encoding strategies for solution. These strategies were found to depend upon the structural properties of the problem, the location of the question, and the time allowed for pre-viewing of one part of the problem. There was evidence for cognitive blocking to the negative problems, an initial priming effect when the question was presented before the premise, and an encoding process which prepared the subject for either question when more time was allowed for preparation of the premise. Results indicated that subjects passed from a "structural" to a "meaning" approach within the time allowed for pre-coding, and that the procedure of varying display delay has relevance for future studies of serial problem solution.

References

Armor, D. J., & Couch, A. S. Data-text primer. An introduction to computerized social science analysis. New York: Free Press, 1972.

Carpenter, P. A. On the comprehension, storage, and retrieval of comparative sentences. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 401 - 411.

Chase, W. G., & Clark, H. H. Mental operations in the comparison of sentences and pictures. In L. W. Gregg (Ed.), Cognition in learning and memory. New York: John Wiley & Sons, 1972.

Clark, H. H. Influence of language on solving three-term series problems. Journal of Experimental Psychology, 1969, 82(2), 205 - 215. (a)

Clark, H. H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 76(4), 387 - 404. (b)

Clark, H. H. More about "adjectives, comparatives, and syllogisms": A reply to Huttenlocher and Higgins. Psychological Review, 1971, 78(6), 505 - 514.

Clark, H. H. Difficulties people have in answering the question "Where is it?" Journal of Verbal Learning and Verbal Behavior, 1972, 11, 265 - 277. (a)

Clark, H. H. On the evidence concerning J. Huttenlocher and E. T. Higgins' theory of reasoning: A second reply. Psychological Review, 1972, 79(5), 428 - 432. (b)

Clark, H. H., & Card, S. K. Role of semantics in remembering comparative sentences. Journal of Experimental Psychology, 1969, 82(3), 545 - 553.

Clark, H. H., & Chase W. G. On the process of comparing sentences against pictures. Cognitive Psychology, 1972, 3(3), 472 - 517.

De Soto, C. B., London, M., & Handel, S. Social reasoning and spatial paralogic. Journal of Personality and Social Psychology, 1965, 2(4), 513 - 521.

Greene, J. M. The semantic functions of negatives and passives. British Journal of Psychology, 1970, 61(1), 17 - 22. (a)

Greene, J. M. Syntactic form and semantic function. Quarterly Journal of Experimental Psychology, 1970, 22, 14 - 27. (b)

Handel, S., De Soto, C. B., & London, M. Reasoning and spatial representation. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 351 - 357.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning. Psychological Review, 1968, 75(6), 550 - 560.

Huttenlocher, J., Eisenberg, K., & Strauss, S. Comprehension: Relation between perceived actor and logical subject. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 527 - 530.

Huttenlocher, J., & Higgins, E. T. On reasoning, congruence, and other matters. Psychological Review, 1972, 79(5), 420 - 428.

Huttenlocher, J., Higgins, E. T., Milligan, C., & Kauffman, B. The mystery of the "negative equative" construction. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 334 - 341.

Huttenlocher, J., & Strauss, S. Comprehension and a statement's relation to the situation it describes. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 300 - 304.

Johnson-Laird, P. N. The three-term series problem. Cognition, 1972, 1(1), 57 - 82.

Johnson-Laird, P. N., & Tridgell, J. M. When negation is easier than affirmation. Quarterly Journal of Experimental Psychology, 1972, 24, 87 - 91.

Jones, S. Visual and verbal processes in problem-solving. Cognitive Psychology, 1970, 1, 201 - 214.

Keeney, T. J., & Gaudino, D. L. Solution of comparative and negative-equative three-term series problems. Journal of Experimental Psychology, 1973, 101(1), 193 - 196.

Mehler, J. Some effects of grammatical transformations on the recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 346 - 351.

Miller, G. A. Some psychological studies of grammar. American Psychologist, 1962, 17, 748 - 762.

Miller, G. A., McKean, K. O. A chronometric study of some relations between sentences. Quarterly Journal of Experimental Psychology, 1964, 16, 297 - 308.

Polich, J. M., & Schwartz, S. H. The effect of problem size on representation in deductive problem solving. Memory & Cognition, 1974, 2(4), 683 - 686.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. Retention of visual and name codes of single letters. Journal of Experimental Psychology Monograph, 1969, 79(1), 1 - 16.

Potts, G. R., & Scholz, K. W. The internal representation of a three-term series problem. Journal of Verbal Learning and Verbal Behavior, 1975, 14(5), 439 - 452.

Quinton, G., & Fellows, B. J. 'Perceptual' strategies in the solving of three-term series problems. British Journal of Psychology, 1975, 66(1), 69 - 78.

Schwartz, S. H. Modes of representation and problem solving: Well evolved is half solved. Journal of Experimental Psychology, 1971, 91(2), 347 - 350.

Schwartz, S. H., & Fattaleh, D. L. Representation in deductive problem solving: The matrix. Journal of Experimental Psychology, 1972, 95(2), 343 - 348.

Shaver, P., Pierson, L., & Lang, S. Conveying evidence for the functional significance of imagery in problem solving. Cognition, 1974, 3(4), 359 - 375.

Slobin, D. I. Grammatical transformations and sentence comprehension in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219 - 227.

Sternberg, R. J. A proposed resolution of curious conflicts in the literature on linear syllogisms. Attention and performance VIII, 1978.

Sternberg, R. J., Guyote, M. J., & Turner, M. E. Deductive reasoning. Proceedings of the Conference on Aptitude, Learning and Instruction: Cognitive Process Analysis. San Diego: March, 1978.

Wason, P. C., & Johnson-Laird, P. N. Psychology of reasoning: Structure and content. Cambridge, Mass.: Harvard University Press, 1972.

Woodworth, R. S., & Schlosberg, H. Experimental psychology (2nd ed.). New York: Holt, Rinehard, & Winston, 1954.

Reference Notes

1. Potts, G. R., & Scholz, K. W. The internal representation of a three-term series problem. Paper presented at the meeting of the Psychonomic Society, Boston, November, 1974.

2. Kaplan, I. Personal communication, 1975.

3. Zurawski, V. B. R. Visualization, verbalization, and three-term serial problem solving. Paper presented at the meeting of the American Psychological Association, Montreal, August, 1973.

4. Sternberg, R. J. Representation and process in transitive inference. Manuscript submitted for publication, 1977.