Lin, P.J. (1998). A Study of Children's Comprehension Processes by Retelling After Solving Technique. 國立新竹師院學報，第十一期，187-227。

A Study of Children's Comprehension Processes By Using "Retelling After Solving" Technique

Pi-Jen Lin
Department of Mathematics and Science Education

ABSTRACT

This study seeks to understand children's comprehension processes when solving two-step compare problems. Comprehension processes are indicated by response time, deciding appropriate operations, and reconstructing a problem's situation based on cues provided.

An experiment was to designed to test the theoretical hypothesis that children prefer to solve the problems when problem information is consistent with the order of operations required. Two independent variables, problem structure and the order of operations, were examined in light of their effect on children's comprehension processes.

Thirty, fourth-graders, were involved in this study. Each subject was asked to solve eight problems and then retell them. Eight problems consisting of three types of problem structures. A CL problem had two relational sentences with consistent language structure. A PIL problem had inconsistent language structure in one of the two relational sentences and a TIL problem contained two relational sentences with inconsistent language structure.

This study provided empirical evidence for the Consistency Hypothesis (Lewis and Mayer, 1987). Results suggested that the number of relational sentences with inconsistent structure increased the degree of difficulty with two-step compare problems. In a PIL problem, the sequence of relational sentences presented was not a significant variable; however the inconsistency of problem structure was. Results indicated that TIL problems were frequently retold as CL problems, whereas CL problems rarely elicited retelling with PIL or TIL problems.

Two theoretical models, translation model and logical-mathematics model , were developed to interpret the findings of the study.

Key words: two-step, compare, comprehension processes, retelling.

Lin, P.J. (1998). A Study of Children's Comprehension Processes by Retelling After Solving Technique. 國立新竹師院學報，第十一期，187-227。

A Study of Children's Comprehension Processes By Using "Retelling After Solving" Technique

I. BACKGROUND

The purpose of this study is to understand children's cognitive processes in solving two-step word problems. Problem comprehension involves the translation of each sentence of the problem into an internal representation and the integration of the information to form a coherent structure.

In mathematics, learning to solve word problems seems to be particularly difficult for children. Previous studies have paid much attention to children's difficulties in solving one-step compare word problems (Riley, Greeno, & Heller, 1983; Lewis & Mayer, 1987; Verschaffel, De Corte, & Pauwels, 1992; Stern & Lehrndorfer, 1993; Verschaffel, 1994). The consistent finding is that problem solvers' difficulties are likely to occur when the structure of the presented information does not correspond to the order of operation required in the problem. In other words, problem solvers prefer to solve a one-step compare problem with consistent language structure rather than with inconsistent language structure (Lewis & Mayer, 1987; Verschaffel, De Corte, & Pauwels, 1992; Stern & Lehrndorfer, 1993; Verschaffel, 1994). This study attempts to extend the studies of one-step compare problems to two-step compare problems to determine whether problem solvers have a preference for the correspondence between the sequence of problem representation and the order of operations required in the problem.

It is generally accepted that children have greater difficulties in solving two-step word problems than one-step word problems (Quintero, 1983). The two-step problems involved in this research are characterized by the fact that it requires two necessary operations for the execution of the computations. This study examines two-step compare problems for the following reasons: First, children's difficulties in solving two-step problems are greater than in one-step problems (Quintero, 1983) because in two-step arithmetic problems, besides choosing the operations to apply, one has to plan and organize the order in which to apply the operations and to decide to which pair of numbers each operation should be applied. Second, previous studies paid much attention to one-step compare problems and two-step proportional problems (Quintero, 1983; Behr & Post, 1981), but previous studies rarely involved two-step compare additive problems. Third, an implication gained from a study by Verchaffel, De Corte, and Pauwels (1992) is that two-step compare problems, which demand more complicated comprehension processes, might be the most appropriate problems to be used to determine whether children have a preference for consistent language.

Most two-step problems dealt with in previous studies for examining the Consistency Hypothesis (Lewis & Mayer, 1987) one relational sentence, while each two-step word problem involved in this study contains two relational sentences. A two-step compare problem is operationally defined as a problem containing four sentences: one assignment, two relational sentences, and one question. Here, the relational sentences can include additive marked or unmarked terms, such as *"more than" and "less than"*. The mathematical expressions of the two-step problems involved in this study must be expressed by the combination of subtraction with addition (– & +).

To understand children's comprehension processes, an experiment was designed to examine the effects of two independent variables, structure of the problems and order of operations on three dependent variables: response time, correctness of choosing necessary operations, and correctness of retelling. Problem structures were categorized by the number of relational sentences having inconsistent language structure. There were three different structures of the problems contained in the experiment. A CL problem (Consistent Language Structure) had two relational sentences with consistent language structure. A PIL problem (Partially Consistent Language Structure) had inconsistent language structure in one of the two relational sentences. A TIL problem (Totally Inconsistent Language Structure) contained two relational sentences with inconsistent language structure.

Problem 1 expressed as (4 + 5) – 6 is an example of CL problems involved in this experiment. The subject of the assignment sentence is the object of the relational sentence 1. *"More than"* is the relational term of the relational sentence 1 and addition is required, so the relational sentence 1 is a sentence with a consistent language structure. The unknown variable (*Laura*) in question sentence is the subject of the relational sentence 2 and *"less than"* is its relational term. It can be executed with subtraction, so the relational sentence 2 has a consistent language structure. As defined, Problem 1 contains two relational sentences with consistent language structure.

Problem 1:

Joe has 4 marbles.
Juan has 5 marbles ***more than*** **Joe**.
Laura has 6 marbles ***less than*** Juan.
How many marbles does *Laura* have?

Verbal protocol is an appropriate technique for examining children's cognitive processes with respect to elementary arithmetic word problems (Ericssion & Simon, 1980; 1984). Retelling protocols can provide reliable and valid data to reveal children's representation processes (De Corte & Verschaffel, 1987). To reflect children's internal representations, retelling protocols were used in this research as a data collecting technique. However, it is important to note that according to memory theory, the most basic form of retrieval, in which a child remembers a familiar stimulus upon being presented with it again, does not necessarily reflect children understanding of compare problems but can merely reflect how much they remember the superficial structure of the compare problem.

The levels-of-processing theory predicts that the more deeply or carefully new information is analyzed by the learner, the more likely it will be recalled in the future (Craik and Lockhart, 1972). In retrieving the information encoded through the process of elaborate thinking, more accuracy is achieved in "retelling after solving" a problem than by "retelling before solving" it. Based on this analysis of levels-of-processing memory theory, when a child is asked to solve a problem and then retell it, he or she remembers it longer than if he or she is asked to retell the problem before solving it. The basic assumption underlying the use of this technique is that retelling a word problem is not a matter of memorizing and reproducing the problem text verbatim, but a constructive activity in which the subject rebuilds the problem starting from the mental representation that he or she has generated after reading the problem and choosing an arithmetic operation to solve it. In this research, children were given a problem first and asked to solve and represent it with a mathematical expression. Their retelling followed their solving and, therefore, more likely reflected the children's internal representation processes.

Because a retrospective report based on information in long term memory requires an additional process of retrieval, it will display some kinds of errors and incompleteness of data. The errors and incompleteness of retelling data in the present research might be minimized by the process of memorizing (depth of processing) and the process of forgetting (cued recall). Cued recall involves the provision of a specific external cue in some way related to the target information and is more likely to retrieve additional accurate information. The emphasis of this study is not on the accuracy of recalling a problem by memorizing verbatim the superficial structure of the problem, but rather on the accuracy of restating a problem with the same meaning but different structure. This research provided the cues with numerals, names of the people, and mathematical expression of the problem.

To minimize the possibility of memory and reflect children's comprehension processes, this research takes advantage of the idea of the depth of processing theory. Using the "retelling after solving" technique can minimize the possibility of the retellings only reiterating the memory traces of the verbatim information of the problems rather than reflecting the children's mental representations. Furthermore, using "retelling after solving" instead of "retelling prior to solving" (Verschaffel, 1994; Verschaffel & Pauwels, 1992) this research can avoid overestimating the number of children for which problems with inconsistent language structures are not solved by transforming these into problems with consistent language structures.

To examine whether children's preference for the correspondence of problem representation with the required operations, the following hypotheses will be tested.

(1) The structures of the problems that children reconstruct differ according to whether the problems require the operations (–,+) or the operations (+,–).

(2) The reversal errors of operation for problems requiring the operations (–,+) differ from the reversal errors for problems requiring the operations (+,–).

(3) The response time of children solving compare problems with operations (–,+) is not longer than their response time solving compare problems with the operations

(+,−).

II. METHODOLOGY

Subjects

Thirty, fourth-graders selected from three elementary schools in Hsin-Chu, whose mathematical achievement was beyond the average fourth-grade level, participated in the experiment. This research did not intend to consider the subjects as an independent variable, thus the distribution, urban or suburban schools, was not taken into account. Rather, the variables, such as the language structure of the problems tested, are related to the nature of mathematics. These were to be considered as the most important foci. Ten students were recommended by their homeroom teachers from two classes in each school for an equal number of representation per class and school.

The choice of this particular age and ability group was based on several reasons. Previous research has convincingly shown that children from the early grades of elementary school perform very weakly with inconsistent language structure of one-step compare problems (Briars & Larkin, 1984; De Corte & Verschaffel, 1985; Morales *et al*. 1985; Riley *et al*., 1983; Riley & Greeno, 1988; Verschaffel, 1984). Because reversal errors, in the sense of comprehension processes, are a focus of this research, if such young children were chosen, they would not be helpful in retelling protocols. Moreover, such young children frequently make errors with one of the given numbers when answering word problems because of misinterpretation of problem statements. One sample of this error is when the statement "Joe has 3 marbles more than Laura" is described as "Joe has 3 marbles." (Briars & Larkin, 1984; Cummins *et al*., 1988; De Corte & Verschaffel, 1985; Riley *et al*., 1983; Verschaffel, 1984). Or, such low-ability children frequently use unrelated structures when retelling a given problem just solved, such as when "6 − 3 + 2" is incorrectly stated as [1]

"Joe has 6 marbles. Juan has 3 marbles less than Joe.	instead of	"Joe has 6 marbles. Juan has 3 marbles less than Joe.

Juan has 2 marbles.
How many marbles does Laura have?"

Juan has 2 marbles more than Laura.
How many marbles does Laura have?"

The material which children have learned in school is another consideration. An analysis of Taiwan's elementary mathematics textbooks [2] indicated that children do not learn the notion of multiplication initiated by repeated addition for example "3 + 3 + 3 + 3" is expressed as "3 × 4", until the end of the second grade. Third-grade children are skilled in learning grouping and rate problems which can be expressed as "3, four times". Scalar problems referring to the comparative relationship of two objects through the concept of "A is *n* times as many as B" are increased in upper grade levels. The distributive property of multiplication over addition is not explicitly taught until the fourth grade. Namely, the notion of "( )" does not appear before the mathematics curriculum for fourth graders.

Task

The structures of the two-step problems dealt with in the experiment were sorted into two categories. In accordance with Nesher and Hershkovitz's (1994) analysis, both SA and AS structures contain two additional implicitly latent components which derive from the given text. In the sense of problem representations, the distinction between the problems with SA (-,+) structure and those with AS (+,-) structure is simply that the relational sentence requiring subtraction precedes the relational sentence requiring addition in the problems with SA structure, whereas the problems with AS structure reverse this relational sentence sequence and require subtraction followed by addition. However, in the mathematical sense, the different sequences of

1 This retelling protocol is adapted from the finding of the pilot study of this study.

2 This edition of the national mathematics textbooks published in 1992-1993 was designed according to the old

curriculum announced in 1975. The subjects participating in this research were using this edition rather than the new edition for fourth graders which has been not published yet and which was designed in accordance with new curriculum standards announced in 1993.

the two relational sentences presented in a problem with SA structure and one with AS structure possess different mathematical meaning. The correct mathematical expression of SA structure is $(a - b) + c$ or $c + (a - b)$, whereas AS structure is mathematically expressed as $(a + b) - c$. Indeed, $(a - b) + c$ is not equal to $(a + b) - c$. The SA and AS structures of CL problems requiring (– & +) are described in Table 1.

Table 1 SA and AS Structures of CL problems requiring (– & +)

	Schema of Problems	Examples
SA Structure (–,+) is required	**(Assignment Sentence)** A [agent] has a'x . **(Relational Sentence 1)** B [agent] has b'x less than A. **(Relational Sentence 2)** C [agent] has c'x more than B. **(Question Sentence)** How many x's does C have?	Joe has 6 marbles. Juan has 3 marbles less than Joe. Laura has 2 marbles more than Juan. How many marbles does Laura have?
AS Structure (+,–) is required	**(Assignment Sentence)** A [agent] has a'x . **(Relational Sentence 1)** B [agent] has b'x more than A. **(Relational Sentence 2)** C [agent] has c'x less than B. **(Question Sentence)** How many x's does C have?	Joe has 6 marbles. Juan has 3 marbles more than Joe. Laura has 2 marbles less than Juan. How many marbles does Laura have?

The examples of CL, PIL, and TIL problems, expressed as $6 - 3 + 2$, are described in Table 2.

Table 2 Examples of CL, PIL, and TIL Problems

CL	Joe has 6 marbles. Juan has 3 marbles less than Joe. Laura has 2 marbles more than Juan. How many marbles does Laura have?	
PIL	Joe has 6 marbles. Juan has 3 marbles less than Joe. Juan has 2 marbles less than Laura. How many marbles does Laura have?	Joe has 6 marbles. Joe has 3 marbles more than Juan. Laura has 2 marbles more than Juan. How many marbles does Laura have?
TIL	Joe has 6 marbles. Joe has 3 marbles more than Juan. Juan has 2 marbles more than Laura. How many marbles does Laura have?	

The control variables of each problem were the context, the number of sentences, the number of words per sentence, the placement of relational sentences, and the low difficulty of the required calculation. The quantities involved in each problem were considered discrete. The contexts contained in the problems with SA structure were dollars, ages, marbles, and apples. The contexts involved in the problems with AS structures were cats, candies, ages, and pencils.

Another controlled variable were the use of a pronoun, such as "Jon" replaced by "he". This task variable may provide an alternative interpretation of the consistency effect (Verschaffel, De Corte, and Pauwels, 1992; Lewis and Mayer, 1987). Particularly, pronouns cause difficulties not only for the young and inexperienced, but also for experienced subjects confronted with reading tasks that require complex processing (Ehrlich & Rayner, 1983; Oakhill & Yuill, 1986). Therefore, the use of pronouns contributed to the consistency effect. To avoid the probability of an alternative explanation, this study never used pronouns and thus did not take the notion of pronouns into account. The internal consistency of the items indicated by the coefficient of K-R 21 (Kuder-Richardson formula 21) was .86 (Mehrens & Lehmann 1991, P. 256).

Each of the eight problems shown in Table 3 was printed on a 3"× 5" index card. The back side of each card contained three given numbers and names of the people in the problem. The order of these three given numbers and the names of persons was randomly arranged on the card.

Table 3 Task

	SA (–,+)	AS (+,–)
CL	Pete has 6 dollars. Ann has 3 dollars less than Pete. Carol has 10 dollars more than Ann. How many dollars does Carol have? Charlie is 8 years old. Cathy is 2 years younger than Charlie. Ruth is 3 years older than Cathy. How old is Ruth.	David has 2 cats. Fred has 4 cats more than David. Jamie has 3 cats less than Fred. How many cats does Jamie have? Steve has 2 candies. Mark has 6 candies more than Steve. Roger has 2 candies less than Mark. How many candies does Roger have?
PIL	Joe has 6 marbles. Juan has 3 marbles less than Joe. Juan has 2 marbles less than Laura.	Tom is 8 years old. Tom is 4 years younger than Nancy. Pat is 2 years younger than Nancy.

	How many marbles does Laura have?	How old is Pat?
TIL	Diane has 6 apples. Diane has 2 apples more than Maria. Maria has 4 apples less than Jon. How many apples does Jon have?	Jen has 4 pencils. Jen has 2 pencils less than Sue. Sue has 2 pencils more than Toby. How many pencils does Toby have?

Procedures

Each participant was tested, followed by an interview. The procedure for accomplishing the task consisted of two stages. The first stage was the execution of solving the problems. This stage reflected the children comprehension of the translation from the word problem into the operational symbols. Writing a matching mathematical expression reveals a child's thinking process during the selection and the execution of the formal arithmetical operation. Each subject was administered the task individually and was given directions before solving the problems. The order of presenting the problems to each subject was different as each problem was randomly selected from the eight problems and randomly assigned to each subject. The subjects were asked to read each problem, solve it, and write the answer on the answer sheet. There was no time limit; they could read the problem on the card at their own pace and reread it as many times as they wanted before and during the solution process. The response time was defined as the time between the presentation of the card containing the word problem and the moment the answer given was recorded. During the solution, no help was provided by the interviewer. All students were supervised by the researcher.

The second stage involved retelling the statement of the problems. This stage reflects the transformation from mathematical representation to problem representation. The retelling data acquired from individual interviews is used to obtain information about problem solving in which the subject constructed an internal representation of the problem. After solving the eight problems, each subject was asked to retell the problems. Each subject was provided cues, including the three given numbers, the names of the people involved in the problem on the card, and the mathematical expression of the problem written down by the subject on the answer sheet; then, the subject was asked to retell the problem. Cues are helpful for retrieving information correctly from long-term memory. Each subject took proximately forty-five minutes to complete the two stages of the task.

During retelling, additional kinds of help were provided by the interviewer under the following circumstances:

(1) when the subject could not go on retelling because he or she had forgotten the order of the names of the people,

(2) when the subject could not go on retelling because he or she had forgotten the context of the problem, or

(3) when the subject didn't understand the means of retelling, in which case the retelling of a given example was demonstrated.

The subjects were administered the tasks in a quiet room in their school during school hours. Each interview session was recorded on audio tape and video tape. These assured a more accurate transcription of the subject’s thinking process than note taking. The subject was informed about the tape recorder, but was assured that no one but the researcher would hear the tape. The tape recorder was unobtrusive, and no child seemed uncomfortable with its presence.

Procedures of Data Analysis

Using the available data for all 30 subjects and for the eight problems, the following qualitative data was obtained.

(1) Correctness of operations

Children’s answers based on the written solutions for each of the eight problems were scored according to the following categories.

- Correct operation: The subject either gave a correct mathematical expression or made a technical error. Each correct mathematical expression was scored one point. The reason for not distinguishing technical errors from correct answers is that the technical aspect is not dealt with in the comprehension process of a word problem. The mathematical expressions were either as a single equation with parentheses, e.g., (2 + 4) - 3 = 3, or as combination of two equations without parentheses, e.g., 2 + 4 = 6, 6 - 3 = 3.

- Reversal error: A representation error in which the numerical operations for either the first or second steps, or both steps were reversed. The subject answered incorrectly as a result of one or two wrong operations, that is, either (-, -) instead of (-,+) or (-,+) instead of (×,+). A coding system was established to represent reversal errors in the analysis. For example, the code R_+ indicate a reversal error on addition.
- Other error: These answers did not belong in one of the two previous categories because of an incomprehensible response or no answer.

(2) Response time

Response time is defined as the time between the presentation of the card containing the word problem and the moment the answer was recorded.

(3) Correctness of the retelling protocols

Each retelling protocol was scored as either correct or incorrect. A correct retelling was scored (1) if the problem, was stated with the two relational sentences being consistent in subjects and relational terms with the original problem, (2) with a changed subject and object and a changed relational term in one of two relational sentences, (3) with a changed subject as well as object and changed relational terms in both relational sentences. The reason for these decisions is that both comparative expressions "A has x objects more than person B" and " B has x objects less than A" are semantically equivalent.

All other retelling protocols were scored as incorrect, so the category of incorrect retellings contains problems retold with semantic and mathematical characteristics that are different from those of the given problem. The children's various retellings of a problem involved in the experiment are shown in Table 4. The retellings of the rest of the problems administered in this research are presented in Appendix 7 of Lin's thesis (Lin, 1997).

Table 4 Children's Retellings With and Without C/PI/TI Inversion for CL Problem Requiring (–,+)

Pete has 6 dollars. Ann has 3 dollars less than Pete. Carol has 10 dollars more than Ann. How many dollars does Carol have?	
Correct Retellings	**Incorrect Retellings**
Without C/PI/TI Inversion	
Pete has 6 dollars. Ann has 3 dollars less than Pete. Carol has 10 dollars more than Ann. How many dollars does Carol have?	Pete has 6 dollars. Ann has 3 dollars less than Pete. Carol has 10 dollars less than Ann. How many dollars does Carol have?
	Pete has 6 dollars. Ann has 3 dollars more than Pete. Carol has 10 dollars more than Ann. How many dollars does Carol have?
	Pete has 6 dollars. Ann has 3 dollars more than Pete. Carol has 10 dollars less than Ann. How many dollars does Carol have?
With C/PI Inversion	
Pete has 6 dollars. Ann has 3 dollars less than Pete. Ann has 10 dollars less than Carol. How many dollars does Carol have?	Pete has 6 dollars. Ann has 3 dollars less than Pete. Ann has 10 dollars more than Carol. How many dollars does Carol have?
Pete has 6 dollars. Pete has 3 dollars more than Ann. Carol has 10 dollars more than Ann. How many dollars does Carol have?	Pete has 6 dollars. Pete has 3 dollars less than Ann. Carol has 10 dollars more than Ann. How many dollars does Carol have?
With C/TI Inversion	
Pete has 6 dollars. Pete has 3 dollars more than Ann. Ann has 10 dollars less than Carol. How many dollars does Carol have?	Pete has 6 dollars. Pete has 3 dollars less than Ann. Carol has 10 dollars more than Ann. How many dollars does Carol have?
	Pete has 6 dollars. Pete has 3 dollars more than Ann. Carol has 10 dollars more than Ann. How many dollars does Carol have?
	Pete has 6 dollars. Pete has 3 dollars less than Ann. Ann has 10 dollars less than Carol. How many dollars does Carol have?

Four examples of correct retellings and eight of incorrect retellings for problem with the operations (- & +) are listed in the left and right columns of Tables 4. The row dimensions of these tables are sorted into three categories: Zero changed relational sentences, only one changed relational sentence, and two changed relational sentences. C/PI/TI inversion is described in detail in next section.

(4) Consistent /Partially Inconsistent /Totally Inconsistent Inversion (C/PI/TI Inversion)

The inversions among consistent, partially inconsistent, and totally inconsistent language of the given problem are indicators of the internal representation of the subjects involved in this study. Each retelling protocol was identified as to whether it contained a consistent/partially inconsistent /totally inconsistent inversion (C/PI/TI inversion). A C/PI/TI inversion is characterized as a CL problem being retold in the form of a PIL or TIL problem, a PIL being retold as a CL, PIL, or TIL problem, or a TIL problem being retold as a CL problem or a PIL problem.

Table 5 Code Transcripts of Inversions for Each Problem Requiring (– & +)

Code	Indicators (Examples)	Code	Indicators (Examples)
0	***Original problem*** Pete has 6 dollars. Ann has 3 dollars less than Pete. Carol has 10 dollars more than Ann. How many dollars does Carol have?	4	***Change the multiplicative relational term*** Pete has 6 dollars. Ann has 3 dollars less than Pete. Carol has 10 dollars less than Ann. How many dollars does Carol have?
		5	*Change the additive relational term* Pete has 6 dollars. Ann has 3 dollars more than Pete. Carol has 10 dollars more than Ann. How many dollars does Carol have?
		6	***Change the multiplicative and additive relational terms*** Pete has 6 dollars. Ann has 3 dollars more than Pete. Carol has 10 dollars less than Ann. How many dollars does Carol have?
1	***Change the subjects and relational term of the relational sentence as addition required*** Pete has 6 dollars. Ann has 3 dollars less than Pete. Ann has 10 dollars less than Carol. How many dollars does Carol have?	7	***Interchange the subjects of the relational sentence as multiplication required*** Pete has 6 dollars. Ann has 3 dollars less than Pete. Ann has 10 dollars more than Carol. How many dollars does Carol have?
2	*Change the subjects and relational term of the relational sentence as multiplication required* Pete has 6 dollars. Pete has 3 dollars more than Ann. Carol has 10 dollars more than Ann. How many dollars does Carol have?	8	***Interchange the subjects of the relational sentence as addition required*** Pete has 6 dollars. Pete has 3 dollars more than Ann. Carol has 10 dollars less than Ann. How many dollars does Carol have?
		9	***Interchange the subjects of the relational sentences as addition and as multiplication required*** Pete has 6 dollars. Pete has 3 dollars more than Ann. Carol has 10 dollars less than Ann. How many dollars does Carol have?
3	***Change the subjects and relational terms of the two relational sentences*** Pete has 6 dollars. Pete has 3 dollars more than Ann. Ann has 10 dollars less than Carol. How many dollars does Carol have?	10	***Change one relational sentence as multiplication required and interchange the subjects of the relational sentence as addition required*** Pete has 6 dollars. Pete has 3 dollars more than Ann. Carol has 10 dollars more than Ann. How many dollars does Carol have?
		11	***Change one relational sentence as addition required and interchange the subjects of the relational sentence as multiplication required.*** Pete has 6 dollars. Pete has 3 dollars less than Ann. Ann has 10 dollars less than Carol. How many dollars does Carol have?
		E	***Miscellaneous***

Note: The numeral codes are nominal variables.

The C/PI/TI inversion is emphasized not only by correct retelling but also by incorrect retelling within each problem category. A correct retelling scored without C/PI/TI inversion means the retelling was the same as the original problem. A C/PI/TI inversion score for a correct retelling implies that the retelling contained at least one changed relational sentence. An incorrect retelling was scored with C/PI/TI inversion when it had at least one changed subject but no changed relational terms. An incorrect retelling without C/PI/TI inversion was scored when all subjects were unchanged but there was at least one relational term that was different from the given problem.

A coding system of inversions was developed; its scheme was determined by the alternation of the subjects and the relational terms of the given problem. There are six main elements in the two relational sentences of a problem (three per sentence): the subject, relational term, and object of relational sentence 1, and the subject, relational term, and object of relational sentence 2. The code transcripts of inversions of each problem, described in Table 5, and subjects' retellings were coded by these code transcripts.

Statistical Analysis

A 2×3 way repeated measure (zero between and two within variables) (Kirk, 1968; 1982; Winer, 1971; Howell, 1991, p.431-483) was used to test the overall effect of the two independent variables, the order of operations ((–,+) vs. (+,–)), and the structure of the problems (CL, PIL, TIL structures) on the three dependent variables, response time, correct operations, and correct retellings. To test the simple main effect of each dependent variable at each level of independent variable, Satterthwaite's formula (Howell, 1991) was used to find the degree of freedom for the pooled error. For example, the degree of freedom for the pooled error of MS_1 and MS_2 is

$$df = \frac{(a_1 MS_1 + a_2 MS_2)^2}{\frac{a_1^2 MS_1^2}{df_1} + \frac{a_2^2 MS_2^2}{df_2}}$$

df_1 and df_2 are the degrees of freedom of MS_1 and MS_2

Here, $a_1 MS_1 + a_2 MS_2 = df_1 \times MS_1 + df_2 \times MS_2$

$df_1 + df_2$

If the simple main effect was significant, then the post hoc Scheff'e— t-test was used to test the comparisons in pairs.

III. RESULTS AND ANALYSES

Comparison of Response Time Between CL, PIL, and TIL Problems

The 2×3 repeated measures ANOVA shows that the factor of a problem structure (CL, PIL, or TIL) had a significant effect on response time, $F(2,58) = 105.93$, $MSe = 8.93$, $p<.01$. There was no interaction with response time between the order of operations and the structure of problems. The main effect of the structure of problems was tested by t-test pairwise. The t-test shown in Table 6 indicated that CL problems were solved significantly faster than PIL and TIL problems; PIL problems were solved significantly faster than TIL problems.

Table 6 Follow-up T-test of Response Time for CL, PIL, and TIL Problems Requiring (– & +)

Means of Scores (seconds)			T value		
CL	PIL	TIL	CL vs . PIL	CL vs. TIL	PIL vs. TIL
10.34	14.01	18.26	6.67**	14.40**	7.73**

*p<.05 **p<.01 $t_{.975}(50) = 2.009$, $t_{.995}(50) = 2.678$

Comparison of Children's Choosing Operations Between CL, PIL, and TIL Problems

The correctness of choosing appropriate operations and reversal errors were used to indicate children's comprehension from problem representation transforming to symbolic representation.

1. Correctness of Operations

There was an interaction of executing operations between the structures of problems and the orders of operations. F(2, 58) = 3.69, p<.05. MSe = .20. This significance tells us that the effect of the orders of the operations depended on the level of the structures of the problems.

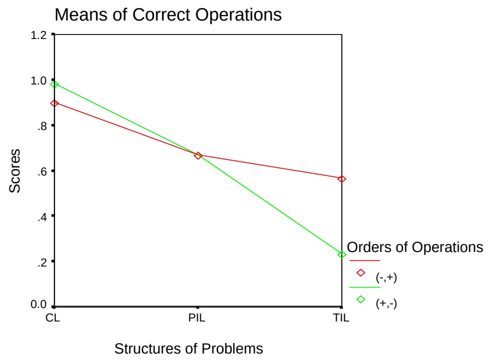


Figure 1　Comparison of the Correct Operations for CL, PIL, and TIL Problems Requiring (–,+) and Problems Requiring (+,–)

In Figure 1, the interaction of the order of operations with the structure of the problems demonstrates that when the operations (+,–) were required in the problems, the execution of CL problems scored significantly higher than PIL and TIL problems and PIL problems scored higher than TIL problems. Likewise, when the operations (–,+) were required in the problems, the performance on CL problems was better than on PIL and TIL problems and PIL problems scored higher than TIL problems, but the difference of scores was slight.

The interaction also shows that the scores of CL problems requiring the operations (–,+) were higher than the operations (+,–), whereas the scores of TIL problems requiring the operations (–,+) required were significantly lower than the operations (+,–). The simple main effects of the structure of problem under (-,+) and (+,-) show that the factor of a problem structure indeed affected children's performance on the problems requiring (–,+) and the problems requiring (+,–), F(2,121) = 20.22, p <.01 and F(2,121) = 30.56, p <.01.

Table 7　Follow-up T-test of Correct Operations for CL, PIL, and TIL Problems Requiring (– & +)

Orders of Operations	Means of Scores			T value		
	CL	PIL	TIL	CL vs. PIL	CL vs. TIL	PIL vs. TIL

(–,+)	.90	.67	.57	2.09*	3.00**	0.90
(+,–)	.98	.67	.23	2.82**	6.82**	4.00**

*p<.05 **p<.01 $t_{.975}(50) = 2.009$, $t_{.995}(50) = 2.678$

Table 7 indicates that for the problems requiring (–,+), the score of CL problems elicited was significantly higher than PIL and TIL problems, while the score of PIL problems elicited was not significantly higher than TIL problems. For the problems requiring the operations (+,–), there were significant differences in performance between CL and TIL, between PIL and TIL, and between CL and PIL problems.

An analysis reveals that there was a significant effect between (–,+) and (+,–) for the TIL problems but not for the CL and PIL problems. The means of scores of PIL problems requiring the operations (–,+) and the operations (+,–) were also the same. The score of TIL problems requiring (–,+) was significantly higher than TIL problems requiring (+,–). The means of these scores were .57 and .23 seconds, respectively, $F(1, 84) = 9.82$, p<.01.

2. Reversal Errors

Children's correct solutions were coded with three different representations and incorrect answers were sorted into three codes, reversal error on addition (R_+), reversal error on subtraction (R_-), and reversal error on both addition and subtraction (R_{+-}). The frequencies of reversal errors are accounted for in Table 8.

Table 8 Types of Reversal Errors on CL, PIL, and TIL Problems

Types of Reversal Errors	Percentage					
	CL		PIL		TIL	
	(–,+)	(+,–)	(–,+)	(+,–)	(–,+)	(+,–)
R_+: Reversal error on addition	2	2	10	12	9	3
R_-: Reversal error on subtraction	2	0	2	2	3	9
R_{+-}: Reversal error on the combination of addition and subtraction	0	0	2	0	14	28

Previous statistical analysis revealed that children made more reversal errors with PIL and TIL problems than CL problems. The reversal errors occurred at the time of the problem representation being translated into symbolic representation. More inconsistent structures in a problem contributed to the higher degree of difficulty in transformation from a problem representation to a symbolic representation.

Despite the fact that it was found that children's difficulties with PIL problems requiring (–,+) were distinct from the PIL problems requiring (+,–), they had a common reversal error. The common error was that when solving a PIL problem, whatever the order of operations that was required, reversal errors only appeared in relational sentences requiring addition (see Table 8).

(1) Why did children make more reversal errors on addition for PIL problems?

There are several possible interpretations. The sequence of the relational sentences presented in a problem might be considered a possible variable; however this experiment afforded strong evidence to exclude this possibility because the relational sentence which required addition in each of the two PIL problems was located either in the first relational sentence or in the second relational sentence. Still the results show more reversal errors on addition than on subtraction. For instance, Problem 2 had more reversal errors on addition than Problem 3. Problem 2 was expressed as 6 + 4 – 2 and Problem 3 was expressed as 6 – 3 + 2.

Problem 2:

Tom is 6 years old.
Tom is 4 years younger than Nancy.
Pat is 2 years younger than Nancy.
How old is Pat?

Problem 3:

Joe has 6 marbles.
Juan has 3 marbles less than Joe.
Juan has 2 marbles less than Laura.
How many marbles does Laura have?

This difficulty was because the inconsistent structure, underlined above, was presented in the first relational sentence of Problem 2 and the inconsistent structure was presented in the second relational sentence of Problem 3.

Another possible interpretation would look at whether the structure was consistent with the operation required in a problem. Namely, more reversal errors being made on addition than on subtraction for PIL problems was due to the inconsistent structure being included in the relational sentence which required addition. Examining the structure of the PIL problems in this experiment, we find that the inconsistent structure was only presented in the relational sentence requiring addition. Consequently, the problems involved in this experiment were not enough to provide sufficient evidence that children's preference for addition over subtraction contributed to more reversal errors on addition than subtraction.

In addition to the inconsistency of a problem structure, the use of key-word strategy is an alternate interpretation for more reversal errors on addition than on subtraction for PIL problems. For instance, some children incorrectly reversed the following problem on addition as 6 – 3 – 2 instead of the correct answer 6 – 3 + 2.

Joe has 6 marbles.
Juan has 3 marbles less than Joe.
Juan has 2 marbles less than Laura.
How many marbles does Laura have?

Analysis: Children's difficulty in transforming the relational sentence requiring addition into a mathematical expression was because of the inconsistent structure located in the relational sentence which required addition. It is hard for children using the mathematical expression "+ 2" to express "2 marbles less than" unless they comprehend the meaning of the problem. Children corresponded "less than" to "subtraction". Case 6, Lu, said, "3 less than 6 is 3 and 2 less than 3 is 1; more "less than" phrases means it is getting smaller, so 6 – 3 – 2 = 1."

(2) Why did children have more difficulties with TIL problems requiring (+,-) than TIL problems requiring (-,+)?

The higher degree of difficulty with TIL problems requiring the operations (+,–) than those requiring the operations (–,+) might be interpreted as being caused by the translation from two inconsistent structures contained in relational sentences into two consistent structures. Thus, many reversal errors on the combination of addition and subtraction occurred in TIL problems. Particularly, the number of errors for TIL problems requiring (+,–) was twice as much as TIL problems requiring (–,+), 28% versus 14%. In addition to the unsuccessful synchronous translations of two relational sentences with inconsistent structure from problem representation into symbolic representation, incomplete translation of one of two relational sentences was another source of children's difficulty with TIL problems.

Comparing the structures between the TIL problems requiring (–,+) and the problems requiring (+,–), the high frequencies of reversal errors both on addition for the problems which required (–,+) and on subtraction for the problems which required (+,–) was because of the fact that the inconsistent language structure of each of these problems was located in the second relational sentence. As an explanation, Problem 4 is expressed as 6 – 2 + 4 and Problem 5 is expressed as 4 + 2 – 2.

Problem 4:	Problem 5:
Diane has 6 apples. Diane has 2 apples more than Maria. Maria has 4 apples less than Jon. How many apples does Jon have?	Jen has 4 pencils. Jen has 2 pencils less than Sue. Sue has 2 more pencils than Toby. How many pencils does Toby have?

In each problem containing two relational sentences with inconsistent structure, children had more reversal errors on addition for Problem 3 and more reversal errors on subtraction for Problem 4 (underlined). The source of the reversal errors is that both the relational sentence requiring addition required in Problem 3 and the relational sentence requiring subtraction required in Problem 4 were located in the second relational sentence.

(3) Why did children have the most difficulties with TIL problems?

A model termed a "Translation model" was used to interpret the children's production of reversal errors. The translation model is patterned on Lewis and Mayer's (1987, P.368) translation model that was used to explain difficulties with one-step problems. The procedures of encoding each object (subjects or quantities) in each sentence as a code in the translation model and the schema of each problem involved in this experiment is explained in Figures 2 and 3.

Schema for Assignment Sentence: First Input Sentence **(SET A) = (value a)**
Schema for Relational Sentence 1: Second Input Sentence **(SET B) = (value b) relation (SET A)**
Schema for Relational Sentence 2: Third Input Sentence **(SET C) = (value c) relation (SET B)**
Schema for Output Equation **(SET C)= (value a) operator (value b) operator (value c)**

Figure 2 Schema of Each Problem

Procedure for Encoding Assignment Sentence 1. Find (name 1) and (number 1) 2. Assign (name 1) to (SET A) 3. Assign (number 1) to (value a)
Procedure for Encoding Relational Sentence 1 4. Find (name 2), (number 2), (comparative relation), and (name 3) 5. Assign (comparative relation) to (operator) 6. If (name 2) is (name 1) go to **rearrangement subprocedure for Rearranging Relational Sentence 1** 7. Assign (name 2) to (SET B) 8. Assign (number 2) to (value b) 9. Assign (name 3 is name 1) to (SET A)
Procedure for Creating symbolic Representation 1 10. Create output equation using current values of SET B, value a, operator, and value b.
Procedure for Encoding Relational Sentence 2 11. Find (name 4), (number 3), (comparative relation), and (name 5) 12. Assign (comparative relation) to (operator) 13. If (name 4) is (name 2) go to **rearrangement subprocedure for Rearranging Relational Sentence 2** 14. Assign (name 4) to (SET C) 15. Assign (number 3) to (value c) 16. Assign (name 5 is name 2) to (SET B)
Procedure for Creating Symbolic Representation 17. Create output equation using current values of SET C, value b, operator, and value c. Or, create output equation using current values of SET C, value a, operator, value b, operator, and value c.
Subprocedure for Rearranging Relational Sentence 1 R_1 Interchange (name 3) and (name 2) R_2 With P_1 %, reverse (operator) R_3 If (comparative relation) is a marked term R_4 With p_2 %, assign (comparative term) to (operator) R_5 Go back to Procedure for Encoding Relational Sentence 1
Subprocedure for Rearranging Relational Sentence 2 R_6 Interchange (name 5) and (name 4) R_7 With p_3 %, reverse (operator) R_8 If (comparative relation) is a marked term R_9 With p_4 %, assign (comparative term) to (operator) R_{10} Go back to Procedure for Encoding Relational Sentence 2

Figure 3 Translation model of procedures from problem representation into mathematical symbols (patterned from "Students' Miscomprehension of Relational Statements in Arithmetic Word Problems," by Lewis & Mayer, 1987, p. 368)

Steps 1 – 3 involve encoding the assignment sentence. The assignment schema is Jen = 4. Relational sentence 1 is encoded in steps 4 –9. In step 6, relational sentence 1 contains the first name Jen, the number 2, the relational term less than, and the second name Sue. In step 5, the relational term is translated into an operational symbol, i.e., "less than" is translated into a minus sign. Step 6 determines whether the first name in relational sentence 1 corresponds to the name of the assignment sentence. If the names do not match, children continue in steps 7-9. If the names do match, as in the example problem, the encoding of relational sentence 1 must be rearranged by jumping to the rearrangement subprocedure of relational sentence 1 (steps $R_1 - R_5$).

The development of the translation model of the present study, given in Figure 3, has simplified some procedures of one-step compare problems involving the work of Lewis and Mayer (1987, p.368) and has extended corresponding procedures from one relational sentence to other relational sentences. Namely, the model removed two procedures for encoding schema from Lewis and Mayer's translation model and added seven procedures for encoding relational sentence 2 and five sub-procedures for rearranging relational sentence 2. Besides this, the word "probability" described in Lewis and Mayer's model is replaced by "percentage" in this study.

The translation model describes the translation procedures from problem representations into mathematical expressions. This model is compatible with the text-processing model which emphasizes linguistic structure reconstruction strategy. The following TIL problem expressed as 4 + 2 – 2 is given as an illustration.

Jen has 4 pencils.
Jen has 2 pencils less than Sue.
Sue has 2 pencils more than Toby.

How many pencils does Toby have?

In step R_1 the ordering of the two names in relational sentence 1 is reversed so that *Sue* becomes the subject, and *Jen* becomes the object. In step R_2 the operator is reversed (i.e., the "–" is changed into a "+") but children may fail to carry out this reversal with the percentage $p_1 < .03$ (from Table 4-21, reversal errors). This percentage depends on how well the children comprehend this sentence. Step R_3 tests whether the relational term is a marked or unmarked term. According to Clark (1969), unmarked terms are stored in memory in a less complex and more accessible form than their opposites. "*more than*" is an unmarked term and "*less than*" is a marked term. As in this example problem, the term is marked (*less than*), so children are more likely ($0 < p_2 <.03$) to retain the original operator (minus). In step R_5 the rearrangements are complete, and the procedure returns to steps 7 – 9. In step 12, either the correct output equation (Sue = 4 + 2) or the reversed equation (Sue = 4 – 2) is created.

For problems in which relational sentence 1 contains inconsistent language structure using a marked term, as in the example problem, children run the risk of making a reversal error at step R_2 and at step R_4 at a rate of approximately .03. For problems in which relational sentence 1 contains inconsistent language structure using an unmarked term, children run the risk of making a reversal error only at step R_2, so the rate of creating a reversed equation in step 10 is p_1. For problems in which relational sentence 1 contains consistent language structure with or without marked terms, the rate of making a reversal error is 0 because children do not invoke the rearrangement sub-procedure.

Likewise, Steps 11 – 16 involve encoding relational sentence 2. In step 11 relational sentence 2 contains *Sue* (first name), *2, more than*, and *Toby* (second name). In step 12 the relational term is translated into an operational symbol, i.e., "*more* *than*" is translated into a plus sign. Step 13 determines whether the first name in relational sentence 2 differs from the second name of relational sentence 1. Since the names match, as in the example, the encoding of relational sentence 2 must be rearranged by jumping to the rearrangement sub-procedure of relational sentence 2 (steps $R_6 - R_{10}$).

In step R_6 the ordering of the two names in relational sentence 2 is reversed so that *Toby* becomes the subject, and *Sue* becomes the object. In step R_7 the operator is reversed (i.e., the "+" is changed into a "–") but children may fail to carry out this reversal at the rate p_1 (< .09). This rate is decided by how well the children comprehend this sentence. Step R_8 tests whether the relational term is marked or unmarked. As in this example problem, the term is marked, so the operation is more likely retained with the rate p_2 (< .09). In step R_{10}, the rearrangements are complete, and the procedure returns to step 14 – 16. In step 17, either the correct output equation (Toby = Sue – 2 or Toby = 4 + 2 – 2) or the reversed equation (Toby = Sue + 2 or Toby = 4 – 2 + 2 or 4 + 2 + 2) is created.

For problems containing a relational sentence 2 with inconsistent language structure using a marked term, children run the risk of making a reversal error at step R_7 and at step R_9. The rate of creating a reversed equation in step 17 is approximately .09. For problems with marked relational terms in two relational sentences, children run the risk of making a reversal error at step R_2, R_4, R_7, and R_9, so the risk of creating a reversed equation is .28. For problems containing a relational sentence 2 with inconsistent language structure using an unmarked term, children run the risk of making a reversal error only at step R_7, as in the example problem, so the rate of creating a reversed equation in step 17 is somewhat less than .09. For problems with marked terms in both relational sentences, children run the risk of making a reversal error at step R_2, R_4, and R_7, so the risk of creating a reversed equation is $p_1 + p_2 + p_3$. For problems containing relational sentences with consistent language structure with or without the marked terms, the percentage of making a reversal error is 0 because children did not invoke the rearrangement sub-procedures.

The translation model is used to provide an interpretation of children's greater degree of difficulty with TIL problems than PIL and CL problems. The TIL problems require two sub-procedures for rearranging relational sentence 1 and relational sentence 2, and PIL problems require one sub-procedure for rearranging either relational sentence 1 or relational sentence 2. This model implies that children run the greater risk of making a reversal error during the rearrangement of two sub-procedures only when solving a TIL problem [the rate is combined P_1 at step R_2 with P_3 at step R_7]. Moreover, the rate of reversal errors is greater if the TIL problem contains a marked term. Indeed, as this example shows, the risk of reversal errors is situated at steps R_2, R_4, and R_7. In general, the risk of reversal errors for PIL problems involved in this experiment was situated either at steps R_2 and R_4 or at steps R_7 and R_9. CL problems, however, do not contain any risk at all.

The translation model provides better interpretation of the higher number of reversal errors in the second relational sentence in a TIL problem than the first relational sentence. The rate of reversal errors on the first relational sentence is situated at steps R_2 and R_4, and the rate of reversal errors on the second relational sentence is situated at steps R_7 and R_9. In this experiment, the rate of reversal errors on the second relational sentence is .09 and on the first relational sentence is .03, shown in Table 8.. The rate of reversal errors on the second relational sentence is greater than those on the first relational sentence. The interpretation of the translation model is compatible with the memory theory of cognition. The rearrangement the second relational sentence demands more procedures and more complex working memory. Increased memory demands could hinder children's rearrangement and thereby increase the percentage of risk that they would make errors during the translation of the second relational sentence.

Comparison of Children's Retellings Between CL, PIL, and TIL Problems

1. Correctness of Retellings

In order to understand children's comprehension processes for translating a problem representation into a symbolic representation, children were asked to retell the problem solved. Children's retellings were sorted into correct and incorrect categories. If the meaning of the given problem was not changed, it would be sorted into the correct category; otherwise it would be sorted into the incorrect category.

It is found that there was an interaction between the structures of problems and the orders of operations: $F(2,58) = 3.69$, MSe = .09, $p<.01$. This interaction tells us that the effect of the problem structure depended on the effect of the order of operations. Figure 4 reveals that the mean of CL problems requiring (+,–) was a little higher than CL problems requiring (–,+), while the mean of PIL problems requiring (+,–) was much lower than that of PIL problems requiring (–,+).

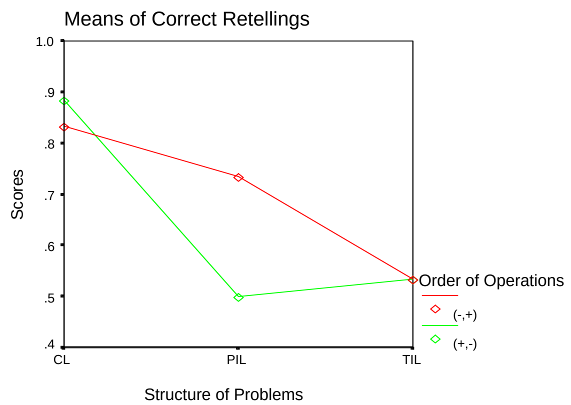


Figure 4 Comparison of Correct Retellings for CL, PIL, and TIL Problems Requiring (–,+) and (+,–)

The data of Table 9 indicates that when children were asked to retell problems requiring the operations (+,–), their correct retellings were significantly better for CL problems than for PIL and TIL problems. When the operations (–,+) were required in a problem, CL problems were better than TIL problems. It was found that PIL problems were slightly, but not significantly, easier than TIL problems.

Table 9 Follow-up T-test for CL, PIL, and TIL Problems Requiring (+ & –)

Correct Retellings	Means of Scores			T value		
	CL	PIL	TIL	CL v. PIL	CL v. TIL	PIL v. TIL
(–,+)	.83	.73	.53	.91	2.72**	1.81
(+,–)	.88	.50	.53	2.27*	3.18**	0.27

*p<.05 **p<.01 $t_{.975}(50) = 2.009$, $t_{.995}(50) = 2.678$

It is found that there was no significant difference between solving CL problems requiring (–,+) and CL problems requiring (+,–). However, the score of PIL problems requiring (+,–) was significantly lower than that of PIL problems requiring (–,+), $F(1,94) = 8.10$, $MSe = .10$, p<.01. The score of TIL problems requiring (–,+) was the same as TIL problems requiring (+,–).

In sum, the order of operations required in CL and TIL problems was not a significant variable affecting children's retellings, whereas the structure of problems was a significant factor of correct retellings. The order of operations only affected the PIL problems. Moreover, more correct retellings were elicited for problems which required (–,+) than for problems which required (+,–). The interview data contributes to the explanation of why there were more incorrect retellings on the problems requiring (+,–) than the problems requiring (–,+).

To analyze the interview data, the categories have been rearranged to facilitate comparison. These do not follow the order of the data on the gathering sheets but are grouped to show the subjects (S) as well as the eight compare problems requiring (– & +) in relation to responses. Each retelling code is based on the code transcripts of inversion. The thirty strings of inversions C/PI/TI are catalogued in Table 10.

The data of Table 10 demonstrates that there were more correct retelling protocols for CL problems (codes 0, 1, 2, and 3) than PIL and TIL problems. Children frequently correctly inverted PIL problems into CL problems (code 1). Thirty percent of children retelling correctly preferred to invert TIL problems into CL problems (code 3).

The following retelling example serves as an illustration that an incorrect retelling might originate in the phase of transformation from problem representation

Table 10 Categories of Children's Retellings for Problems Requiring (– & +)

	(C)	**(TI)**	**(TI)**	**(C)**	**(C)**	**(PI)**	**(PI)**	**(C)**
1	0	3	3	0	0	1	1	0
2	1	3	3	0	0	1	1	0
3	0	1	3	0	0	1	7	0
4	0	9	11	0	0	1	0	0
5	0	9	9	0	8	1	11	7
6	0	3	9	0	0	7	1	0
7	0	10	3	0	8	1	1	0
8	0	3	12	1	0	12	1	0
9	3	0	0	10	0	0	1	11
10	0	3	9	0	E	7	7	0
11	0	3	3	0	0	1	1	4
12	0	3	3	0	0	1	1	0
13	0	3	3	0	0	7	1	0
14	0	3	3	0	8	E	1	4
15	0	0	1	0	1	0	1	0
16	0	9	9	0	0	1	10	0
17	0	3	3	0	2	1	1	0
18	0	9	9	0	0	7	1	2
19	0	11	3	0	0	1	1	0
20	0	1	3	0	0	9	7	0
21	0	9	9	0	0	11	11	0
22	0	10	3	0	0	1	7	0
23	0	3	3	0	E	10	1	0
24	0	3	9	0	0	10	1	0
25	5	10	3	0	0	7	7	0
26	0	9	9	0	0	7	1	0
27	0	3	3	0	E	1	1	0
28	5	9	10	0	0	E	1	0
29	2	8	12	3	1	0	0	11
30	0	11	11	E	E	9	1	0

Correct retellings: Code 0, 1,2, and 3
Incorrect retellings: Code 4-12 and E

to operational symbols.

Tom is 8 years old. Tom is 4 years younger than Nancy. Pat is 2 years younger than Nancy. How old is Pat?	was frequently retold as	Tom is 8 years old. Nancy is 4 years older than Tom. Pat is 2 years older than Nancy. How old is Pat?

The first relational sentence was recorded as "8 – 4 = 4" instead of "8 + 4 = 12". After that, the mathematics representation "8 – 4" was described as "Nancy is 4 years younger than Tom." The retelling based on the children's solutions provides convincing support for children's preference for consistent language structure. However, the retelling was not compatible with the meaning of the original problem.

In addition to the incorrect retellings in which children preferred to state problems with consistent language structure, the elicited correct retellings also consisted of consistent language structure. The following TIL problem requiring (+,–) was retold correctly.

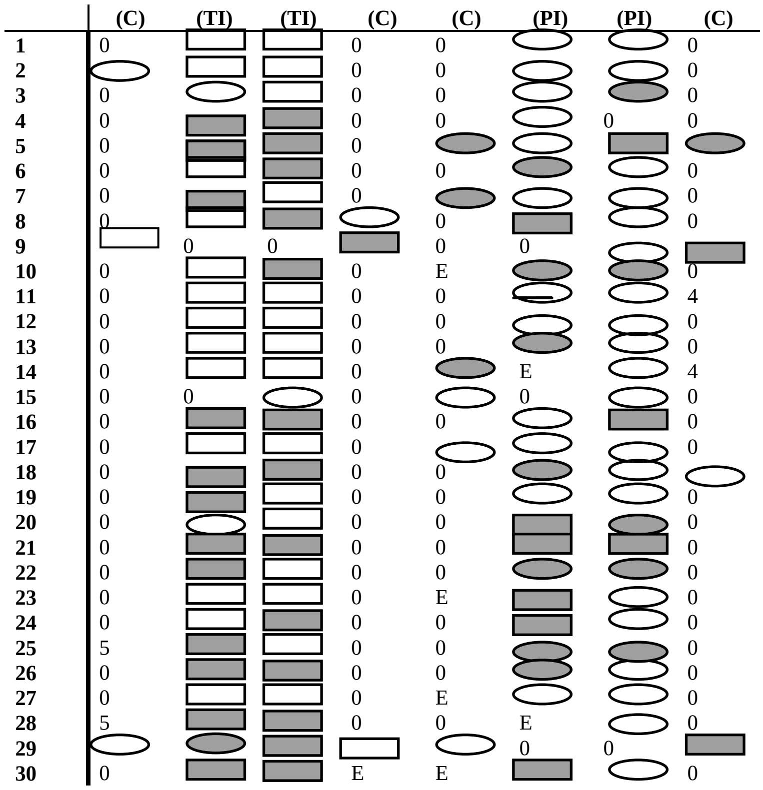
TIL Problem		**Solution**		**Correct Retelling**
Diane has 6 apples. Diane has 2 apples more than Maria. Maria has 4 apples less than Jon. How many apples does Jon have?	→	6–2+4 = 8	→	Diane has 6 apples. Maria has 2 apples less than Diane. Jon has 4 apples more than Maria. How many apples does Jon have?

Successful retellings have to succeed throughout two phases of transformations: problem representation to mathematical expression and mathematical expression to problem representation. Correct retelling reflects children's understanding of the structure of the text rather than their use of key-word strategies. After reading the problem, children succeeded in giving the correct mathematical expression 6 – 2 + 4 = 8. However, when asked to restate the problem, children preferred to transform "6 – 2 + 4" to consistent language structure. As the above example illustrated, consistent language structure was preferred by children for the transformation from mathematical expression to problem representation.

2. Patterns of Incorrect Retellings

As mentioned in the discussion of incorrect retelling patterns, when children were asked to retell a problem, the number of changed relational sentences has to do with the structure of the problems. The interviewee's retellings are displayed in Table 11. The use of an ellipse ⬭ , represents a retelling with one changed relational sentence. The use of a rectangle , represents a retelling with two changed relational sentences. Shaded symbols represent incorrect retellings. Conversely, unshaded symbols represent correct retellings.

Table 11 Pattern of Retellings for Problems Requiring (– & +)



0 : Correct retelling without changed relational sentences
4–6 : Incorrect retelling with unchanged relational sentence
: Correct retelling with one changed relational sentence
: Incorrect retelling with one changed relational sentence
: Correct retelling with two changed relational sentences
: Incorrect retelling with two changed relational sentences
E : Miscellaneous

According to the data in Table 11, the elicited retellings with two unchanged relational sentences (code 0) frequently occurred in CL problems. The elicited retellings with one changed relational sentence frequently occurred in the PIL problems (ellipses). The elicited retellings with two changed relational sentences frequently occurred in the problems with TIL structure (rectangles). The frequencies of correct and incorrect retellings with unchanged, one changed, and two changed relational sentences for each problem requiring (– & +) are calculated in Table 12.

Table 12 Frequencies of Unchanged, One Changed, and Two Changed Relational Sentences for Each Problem Requiring (– & +)

Relational Sentences	CL		TI		TI		CL		CL		PI		PI		CL		Sum of Incorrect Retellings
	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	
Unchanged	2	25	0	2	0	1	0	26	0	20	0	3	0	2	2	24	4
One Changed	0	2	1	2	0	1	0	1	3	3	10	13	8	20	1	1	23
Two Changed	0	1	12	13	13	15	1	1	1	0	2	0	0	0	2	0	31
Sum	2	28	13	17	13	17	0	27	4	23	12	16	8	22	5	25	58

Besides miscellaneous errors (E), there were 58 incorrect retellings including 4 unchanged relational sentences, 23 one-changed relational sentences, and two-changed relational sentences. Relying on this data, over 50% of incorrect retellings were due to difficulties with changing two relational sentences at a time. These incorrect retellings with two changed relational sentences usually occurred in TIL problems (25 out of 31). Of the incorrect retellings with one changed relational sentence, PIL problems were more frequent (18 out of 23).

3. Consistency Between Solutions and Retellings

The matrix displayed in Table 13 shows the consistency of solutions with the retelling protocols. One dimension contains retellings with correct and incorrect responses. The other dimension is the solutions with correct and incorrect responses.

Table 13 indicates that a large number of the children (89 out of 99) solved and retold the CL problems correctly without changing the structure of CL problems when solving them. Moreover, three cases where the children correctly answered CL problems but failed to retell them indicates that key-word strategy played a less important role for the required operations (– & +) in the present study.

Table 13 Frequencies of Retellings with Respect to Solutions for Problems Requiring (– & +)

Solutions	Correct Retellings						Incorrect Retellings							
	Unchanged		One Changed		Two Changed		Unchanged		One Changed		Two Changed		Miscellaneous	
	(–,+)	(+,–)	(–,+)	(+,–)	(–,+)	(+,–)	(–,+)	(+,–)	(–,+)	(+,–)	(–,+)	(+,–)	(–,+)	(+,–)
CL														
Incorrect	2	1	0	0	0	0	4	0	1	0	0	0	0	0
Correct	44	45	3	4	1	1	3	0	1	3	2	1	0	4
PIL														
Incorrect	2	0	2	3	0	0	0	0	6	7	0	0	0	0
Correct	0	3	18	9	0	0	0	0	2	3	0	2	0	3
TIL														
Incorrect	0	2	0	2	2	11	0	0	0	1	12	7	0	0
Correct	1	0	1	0	14	1	0	0	0	0	3	3	0	0
Totally														
Incorrect	4	3	2	5	3	11	4	0	7	8	12	7	0	1
Correct	45	48	22	13	15	2	3	0	2	6	5	6	0	7

The 30 out of 60 occurrences of one changed relational sentence in retelling protocols for those who solved and retold the PIL problems correctly suggests that these solutions were indeed obtained by mentally transforming the difficult PIL structure into an easier CL structure rather than by key-word strategy. The 13 occurrences of one changed relational sentence for PIL problems and the 19 occurrences of two-changed relational sentences for TIL problems in the erroneous retelling protocols of the children who answered the problems incorrectly with a reversal error indicate that these reversal errors typically resulted from an incomplete transformation from given PIL and TIL problems into CL problems. The 19 erroneous retelling protocols of those who solved correctly suggest that these solutions were obtained by mentally transforming PIL and TIL problems into CL problems, while the erroneous retellings resulted from difficulty of retrieving information from long-term memory (3 occasions) or resulted from incomplete transformation of one or two relational sentences (16 occasions) from symbolic representation to problem representation.

The high percentage (91%, 90 out of 99) of those who solved and retold the

CL problems requiring (− & +) correctly, with the original structure in the retelling protocols suggests that children did not change the structure of CL problems when solving them. Moreover, the low proportion (9 out of 120) of children who correctly answered CL problems but failed to retell them correctly provides more evidence for confirming that key-word strategy played a less important role when the operations (− & +) were required in the problems in the present study. Only 3% (4 out of 120) of children's retellings might have been due to difficulty in retrieving the information from long-term memory; therefore, these retellings did not belong to two-step compare problems.

The high proportion (27/30) of children's retellings with one changed relational sentence in the retelling protocols of those who both correctly solved and retold the PIL problems suggests that these solutions were indeed obtained by mentally transforming the difficult PIL structure into an easier CL structure rather than obtained by key-word strategy. In the erroneous retelling protocols of those who answered the problems incorrectly because of a reversal error, both the 22% (13/60) of children's retellings with one changed relational sentence for PIL problems and the 33% (20/60) of children's retellings with two-changed relational sentences for TIL problems indicate that these reversal errors typically resulted from an incomplete transformation from given PIL and TIL problems into CL problems. Overall, the proportion (33/240) of erroneous retelling protocols of those who solved correctly suggests that the correct solutions were obtained by mentally transforming PIL and TIL problems to CL problems, while the erroneous retellings resulted from difficulty in retrieving information from long-term memory (7) or resulted from incomplete transformation of one or two relational sentences (23) from mathematical expressions to problem representations.

V. CONCLUSIONS, DISCUSSION, AND IMPLICATIONS

1. Conclusions

The present study provides additional empirical evidence of children preference for solving two-step compare problems with consistent language structure. The central finding of the research is that when children are confronted with a problem with partially or totally inconsistent language structure, they will mentally transform the given problem into a problem with consistent language structure before and after choosing the appropriate arithmetic operation. However, this mental transformation is cognitively demanding. This cognitive demands are explained by translation model developed in this research and shown in Figure 3 Part IV. To re-iterate each R_i, i = 1, 2, .. 10, is the subprocedure for rearranging relational sentences. Each p_i , i = 1, 2, 3, 4, is the percentage in making reversal errors of rearranging relational sentences. The percentage of reversal errors in rearranging problems with totally inconsistent language structure occurred either at steps R_2, R_4, and R_7 or R_2, R_7, and R_9, and is equal to $p_1 + p_2 + p_3$ or $p_1 + p_3 + p_4$, whereas the percentage of reversal errors for problems with partially inconsistent language structure either at steps R_2 and R_4, or at steps R_7 and R_9 and is equal to $p_1 + p_2$ or $p_3 + p_4$. The percentage of reversal errors in CL problems is approximately 0 because children did not invoke the rearrangement sub-procedures. Because of the increased and more complicated demands placed on working memory by two-step compare problems, the mental transformation may be only partially completely. In this case, a problem representation is built incorrectly; therefore, the translation results in a reversal error.

This research provides further evidence of the different demands placed on the working memory in problems with partially inconsistent language structure and those with totally inconsistent language structure. Although previous studies provided data in line with the Consistency Hypothesis, this research yields more empirical evidence of the hypothesis.

Six results of the present study were found to support the hypothesis of this research that children prefer the information of a problem to correspond to the order of necessary operations.

First, the degree of difficulty with two-step compare problems is increased by the number of relational sentences with inconsistent language structure. This study found that the degree of difficulty from the hardest to the easiest for children is TIL, PIL, and CL problems. "Difficulty" in this research was indicated by the response time, choice of appropriate arithmetic operations, and correct retellings when solving a problem. Results suggested that children had greater difficulty with determining the appropriate operations required in TIL problems than PIL and CL problems and have more difficulties in retelling TIL problems than PIL and CL problems. This result confirms Verschaffel, De Corte, and Pauwels's conclusion that children's preference for the problems with consistent language structure is more likely to be articulated when they are given tasks with a higher level of difficulty (1992).

Two theoretical models, a logic-mathematics model (part-whole schema) and a translation model (text-processing), were developed in this research and the working memory of the information-processing approach was used to interpret the findings. These three theories do not conflict with each other but the emphases of each are distinct.

The logic-mathematics model emphasizes the mathematical relationship in part-whole schema of three quantities contained in two-step compare problems. Children had greater degree of difficulty with TIL problems than PIL and CL problems in that solving TIL problems required more help from mathematical transformation strategies. According to the interpretation of proponents of the logic-mathematics model (Briars and Larkin, 1984; Riley and Greeno, 1988; Riley *et al*., 1983; Resnick and Greeno, 1990), some problems can be modeled externally and thus require only counting-based procedures, whereas other problems demand the transformation of the problem text into part-whole relations. Riley and Greeno (1988) postulated three levels of performance that reflect different levels of cognitive development. According to their categories of level, the problems with consistent language structure involved in this research can be solved on level 2. Level 2 allows understanding of quantitative information of the problems that is not presented externally but requires understanding by making inferences about the relationships between the quantities (Stern, 1993). The problems containing two relational sentences with one or two inconsistent language structures can be solved on level 3. At level 3, the part-whole schema is presented and can be combined with knowledge about numerical operations such as the comparison of sets (Stern, 1993, p.9). This involves the comparison of three quantities: small set, large set, and difference set. The solutions for problems with inconsistent language structure dealt with in this research require the help of a mathematical transformation strategy such as "small set = large set - difference set", "large set = small set + difference set", or "difference set = large set – small set", depending on the relational sentences whether the set in the assignment sentence is the small set or the large set and the set in the question sentence is the small set or the large set.

As an example, Figure 4 shows that the primary and least demanding of mathematical transformation strategies for CL problems requires two equations to express the relationship of the quantities involved in the problem. An additional mathematical transformation strategy is required in PIL problems, and two additional mathematical transformation strategies are required in TIL problems. The common requirement of the additional mathematical transformation strategies for PIL and TIL problems is the primary requirement of the CL problems. The relationship between relational sentence 1 and relational sentence 2 is built on small set A of relational sentence 1 being identical with small set E of relational sentence 2.

CL problem:	
Joe has 6 marbles. →	large set A (Joe) = 6
Juan has 3 marbles less than Joe. **(relational sentence 1)** →	small set B (Juan) (?) = large set A (Joe) (6) – difference set C (3)
Laura has 2 marbles more than Juan. **(relational sentence 2)** →	large set D (Laura) (?) = small set E (Juan) + difference set F (2)
How many marbles does Laura have?	The relationship between relational sentence 1 and 2 is connected by the step that small set B of relational sentence 1 is the same as small set E of relational sentence 2.

PIL problem:	
Joe has 6 marbles. →	large set A (Joe)= 6
Juan has 3 marbles less than Joe. **(relational sentence 1)** →	small set B (Juan) (?) = large set A (Joe) (6) – difference set C (3)
Juan has 2 marbles less than Laura. **(relational sentence 2)** →	small set E (Juan) = large set D (Laura) (?) – difference set F (2) is required to transform to: large set D (Laura) (?) = small set (Juan) E + difference set F (2)
How many marbles does Laura have?	The relationship between relational sentence 1 and 2 is connected by the step that small set B of relational sentence 1 is the same as small set E of relational sentence 2.

TIL problem:	
Joe has 6 marbles. →	large set A (Joe) = 6
Joe has 3 marbles more than Juan. **(relational sentence 1)** →	large set A (Joe) (6) = small set B (Juan) (?) + difference set C (3) is required to transform to: small set B (Juan) (?) = large set A (Joe) (6) – difference set C (3)
Juan has 2 marbles less than Laura. **(relational sentence 2)** →	small set E (Juan) (?) = large set D (Laura) – difference set F (2) is required to transform to: large set D (Laura) = small set E (Juan) (?) + difference set F (2)
How many marbles does Laura have?	The relationship between relational sentence 1 and 2 is connected by the step that small set B of relational sentence 1 is the same as small set E of relational sentence 2.

Figure 4: The Logic-Mathematics Model for Interpreting the Difficulty with PIL and TIL problems

The translation model developed in Figure 4 of Part IV underscores the importance of text processing. According to the proponents of the text-processing model, children's difficulties with word problems arise from a lack of textual understanding, which prevents them from making connections with relevant mathematical knowledge. Performing the TIL problems involved in this research requires the help of linguistic restructuring strategy. Referring to the work of Lewis and Mayer (1987), this translation is the usual way of solving problems with inconsistent language structure even for adults and is preferred because subjects attempt to match the language used in the relational statements to describe the situation. Equally, according to the translation model, children run a different risk of making a reversal error during the process of rearrangement when solving CL, PIL, and TIL problems. The analysis of the translation model suggests that the percentage of reversal errors is increased from CL to PIL to TIL problems. The percentage of reversal errors in a TIL problem involving this research at steps R_2, R_4, and R_7 or R_2, R_7, and R_9 is equal to $p_1 + p_2 + p_3$ or $p_1 + p_3 + p_4$, p_i , i = 1, 2,3,4, is the percentage of reversal errors. The percentage of reversal errors of a PIL problem at steps R_2 and R_4 or R_7 and R_9 is equal to either $p_1 + p_2$ or $p_3 + p_4$. There is no percentage of reversal errors required in a CL problem when it is solved. Since $p_1 + p_2 + p_3$ is greater than $p_1 + p_2$ or $p_1 + p_3 + p_4$ is greater than $p_3 + p_4$, children had more possibilities for reversal errors with TIL than PIL and CL problems, and with PIL than CL problems as well. The internal processes involved in the translation model are compatible with the interpretation of the information-processing approach.

From the perspective of information-processing, the demands on working memory are increased while children encode more complicated information in a problem. Increased memory demands could hinder children's success in rearranging the information and thereby increase the probability they will make errors during the procedure of transformation. The TIL problems dealt with in this research possess much more complicated language structure than the PIL and CL problems, thus children solving the TIL problems had a greater percentage of reversal errors.

Many previous studies have documented the effect of language consistency on the relative difficulty of one-step compare problems (Lewis & Mayer, 1987; Verschaffel, De Corte, & Pauwels, 1992; Stern & Lehrndorfer, 1992; Verschaffel, 1994) and two-step compare problems that involved merely one relational sentence (Lewis & Mayer, 1987; Verschaffel, De Corte, & Pauwels, 1992). Although those studies support the Consistency Hypothesis, this research differs from the previous studies in that it used two relational sentences in each problem and provides strong evidence that the effect of language consistency is a critical variable in deciding whether children 10-11 years old are able to solve two-step compare problems.

From the aspect of Clark's (1969) marked and unmarked terms, children are more likely to miscomprehend the relational terms in a problem with inconsistent language structure and therefore commit more reversal errors when the relational term is a marked term (e.g., *less* or *younger* in an addition sentence; or *1/n as many* in a multiplication sentence) than when the relational term is an unmarked term (such as *more* or *older* in subtraction sentence; *or n times as many as* in a division sentence). According to Clark, the meaning of unmarked terms like good, long, and more are stored in memory in a less complex form than that of marked terms (1969). The finding of this research indicates a high percentage of reversal errors in TIL problems in which all relational terms involved are marked and a small percentage of reversal errors in CL problems which merely contain unmarked terms. This finding that children are more influenced by marked terms than by unmarked terms is consistent with Clark's theory.

Conversely, when children were asked to construct a problem situation, in the sense of children’s retellings dealt with this research, the research did not provide adequate evidence to support Lewis and Mayer’s (1987) study that children have a strong preference for unmarked additive terms such as *more* or *older* rather than for marked additive terms such as *less* or *younger*. The difference in this research is that children were asked to retell the problems based on their solution cue, thereby the relational terms used are based mostly on the operations contained in mathematical expressions. In the method of Lewis and Mayer’s study, two known quantities were provided for children to describe the relationship of the quantities using comparative terms.

A second finding in this study is that, for a TIL problem, the sequence of the relational sentences with inconsistent language structure presented in the problem is a crucial factor in determining whether children are able to solve it. When a TIL problem required the combination of two operations, addition and subtraction, children made more reversal errors originating from the second relational sentence than originating from the first sentence. For instance, the following problem can be expressed as 12 – 6 + 2:

There are 12 fourth graders.
There are 6 more fourth graders than sixth graders.
There are 2 less sixth graders than fifth graders.
How many fifth graders are there?

In accordance with the findings of this research, reversal errors of addition originating from the second relational sentence will be predicted to occur more often than reversal errors of subtraction initiated from the first relational sentence.

However, for a PIL problem, the results suggested reversal errors made in the second relational sentence were not more common than reversal errors made in the first relational sentence. The reversal errors always stemmed from the inconsistent language structure which was contained in the second relational sentence, not the first, which contained consistent language structure. Thereby, this research provides further evidence to support that children made more reversal errors on the relational sentence with inconsistent language structure than the relational sentence with consistent language structure. This research verifies not only the evidence that a problem structure is a factor in developing children's ability to solve addition and subtraction problems for one-step problems (Carpenter and Moser, 1979; Quintero, 1980; Lewis and Mayer, 1987, Shalin, 1985; Verschaffel, 1994) but also that structure is a factor for two-step compare problems. The result of more reversal errors in the combination of two operations than in that of one operation suggests that two-step compare problems indeed demand more complicated comprehension processes than one-step compare problems. This result not only supports Quintero's study (1983) but also provides the evidence to confirm the implication from Verchaffel, De Corte, and Pauwels's work (1992) that two-step compare problems might be the most appropriate problems for examining whether children have a preference for consistent language structure.

In short, the number of and the sequence of the relational sentences with inconsistent language structure presented in a problem were two crucial variables for TIL problems, whereas the sequence of the relational sentences presented in PIL problems was not a significant variable. The inconsistency of problem structure was a significant variable rather than the sequence of the relational sentences presented in a PIL problem.

A third finding is that the encoding of the problem representation for TIL problems requires a longer time than PIL and CL problems. PIL problems require a longer response time than CL problems. It was observed during the process of retelling that the time of processing a relational sentence with an inconsistent language structure was longer than processing a relational sentence with consistent language structure in a PIL problem. This finding, which replicates the results of several previous studies (Lewis & Mayer, 1987; Verschaffel *et al.*, 1992; Verschaffel, 1994), indicates that additional processing for the PIL or TIL problems is required. However, results suggested that this additional processing did not guarantee success in solving the problem.

On the basis of the assumptions of the study, children's choice of an appropriate operation are based on the problem representation, and a retelling protocol based on their solutions reflects children's internal processes. One of the final three findings indicated that the greater number of reversal errors and the longer response time for TIL and PIL problems than for CL problems are initiated by the difficulties of processing from a problem representation to a mathematical expression. This conclusion is also in line with Lewis and Mayer's Consistency Hypothesis (1987).

Also, the results of the present study strongly support children's preference for problems in which the problem information corresponds to the order of necessary operations; that is, the problems with consistent language structure. This hypothesis, in addition, is provided by the aspect of the percentage of successfully selecting appropriate operations. It is also supported by the aspect of retellings found in this research that CL problems were retold correctly more frequent than PIL and TIL problems. Also, TIL problems and PIL problems were retold frequently as CL problems, whereas a CL problem elicited little or no retelling with the structure of a TIL or PIL problem. Even though this research dealt with fourth-graders and had them retelling after solving two-step problems while Verschaffel's study dealt with fifth-graders and had them retelling prior to solving one-step problems, the result of this research is coherent with Verschaffel's findings (1994).

This research provides additional empirical evidence to support the Consistency Hypothesis from the aspect of the number of changed relational sentences. Several studies, including this one, support the Consistency Hypothesis from the evidence of more frequencies of problems with consistent language structure being retold correctly than problems with inconsistent language structure and from the evidence of preference for retelling problems with inconsistent structure as problems with consistent language structure (Verschaffel, 1994; Verschaffel, De Corte, and Pauwels, 1992; Lewis and Mayer, 1987). The number of changed relational sentences provides evidence for the hypothesis of this research that children prefer problems in which the problem information corresponds to the order of operations required. The results showed that the elicited retelling protocols containing unchanged relational sentences usually occurred in problems with CL structure. Those with one changed relational sentence frequently occurred in the problems with PIL structure and containing two changed relational sentences frequently occurred in TIL problems. Moreover, the protocols containing two changed relational sentences were more often incorrect retellings.

2. Discussion

(1) Are the children's choices of appropriate operations due to their ability to understand the semantic representation of the essential elements and relationships in the problem?

Reversal errors in two-step compare problems are due to difficulties in understanding and representing the relational terms in the problem or originated from the use of key-word strategy (Fuson *et al.*, 1992; Nesher & Teubal, 1975; Stern, 1993; Verschaffel, De Corte, & Pauwels, 1992). Children applying this strategy do not intend to understand and represent the problem statement but simply look for key words in the problem statement. For example, the words "more" and "older" are associated with addition, the words "less" and "younger" associated with subtraction. The key-word strategy sometimes yields success, while sometimes it results in reversal errors. In the present study, key-word strategy was frequently applied in TIL problems and made reversal errors, whereas it played a less important role in CL and PIL problems. In the context of the present research, the correct answers on all CL problems could result from applying a key-word strategy or understanding the meaning of the problems. If children indeed used key-word strategy rather than understood the relational information in the problem, incorrect retellings with unchanged relational terms and interchanged subjects with objects would probably have occurred. Apparently, the results of this research do not support such evidence. It is concluded that a large majority of the children in this research did not apply superficial key-word strategy when solving two-step compare problems with consistent language structure or with partially consistent language structure. Using two-step compare problems, this research verifies the assumption in Verschaffel's study, namely, children's choice of the appropriate arithmetic operation is based on a semantic representation of the essential elements and relationships in the problem rather than on the use of key-word strategy (Verschaffel, 1994). For example, Jenny

solution is 4 – 2 = 2, 2 + 2 = 4. Her retelling is "*Jen has 4 pencils. Jen has 2 pencils more than Sue. Toby has 2 pencils more than Sue. How many pencils does Toby have?*" If the key-word strategy determines the case's success in solving this problem, the sentence "*Jen has 2 pencils more than Sue*" should be retold as "*Sue has 2 pencils less than Jen*." associated with "4 – 2 = 2". However, when two-step compare problems with inconsistent language structure were too complicated to understand the relational terms, key-word strategy played an important role and lead to an incorrect answer sometimes.

(2) Are the retelling protocols and the written number sentence actually able to reflect children's understanding and mental transformation in solving a two-step compare problem?

The indicator of the mental transformation of solving a problem was characterized by the presence of a TI/C or PI/C inversion in the retelling protocols. The absence of PI/C or TI/C would imply that no such mental transformation took place.

This research uses "retelling the problem after solving it" technique which differs from previous studies using the "retelling the problem prior to solving it" technique (Verscheffel, 1994; Verscheffel & Pauwels, 1992). When children recall the problem text after having solved the problem, recall should become reconstructive, with the number sentence as the basis for the reconstruction. Therefore, the large number of PI/C and TI/C inversions in the related protocols of the PIL and TIL problems demonstrates children's preference for relational sentences with consistent language structure. This implies that in all these cases the PIL or TIL problem was effectively solved by transforming it into a CL problem. This implication avoids the limitation of Verschaffel's study (1994) using the "retelling prior to solving it" technique which may contribute to an overestimation of the number of cases in which a PIL or a TIL problem was not solved by transforming it into a CL problem. "Retelling after solving" can avoid the possibility that retelling with a PI/C or TI/C results from applying prior knowledge about how compare problems are typically formulated or from simply using memory traces of how the problem was actually represented.

On the other hand, the "retelling after solving" technique does reveal all cases in which problems with relational sentences with one or two inconsistent language structures have been solved by transforming it into a problem with consistent language structure. Thus, children's retellings no longer rely on the memory traces of the verbatim problem information, minimizing the limitation of those studies using the "retelling prior to solving" technique in which retellings with TI/C or PI/C are simply due to children remembering verbatim problem information rather than transforming it into a problem with consistent langauge structure (De Corte & Verschaffel, 1987; Stern, 1990).

3. Implications

(1) Instructional Perspective

Although this research did not directly address instructional issues, some guidelines for teaching how to represent and solve PIL and TIL problems can be given. First, instruction should pay attention to children's development of awareness that the comparative relationship between two sets can be expressed in two parallel ways with the terms *more* and *less*, *younger* and *older*. For instance, "Tom has 5 dollars more than Joel" has the same semantic meaning as "Joel has 5 dollars less than Tom". Second, results suggested that children had less difficulty with CL problems than PIL and TIL problems. Theories of learning indicate that learning should be initiated by the easiest learning materials moving to more difficult materials. Therefore, instruction should present CL, PIL, and TIL problems in sequence, with first priority given to CL problems.

(2) Research Perspective

This research utilized a set of problems to understand children's difficulties. The number sentence of each problem in the task required the combination of subtraction and addition. Results suggested that consistency of language structure was significant for children's performances of the sequence of the two operations required in a mathematical expression. Further research could investigate whether problems requiring the combination of operations, such as addition and addition, subtraction and subtraction, and multiplication and multiplication, also have similar results.

Further research is recommended to investigate how much children performance is influenced by the sequence of relational sentences not being consistent with the required operations. For instance, comparing the difference of children responses to the following two problems expressed as 6 – 3 + 4.

Tom has 6 dollars.	Tom has 6 dollars.
Joel has 3 dollars less than Tom.	Lisa has 4 dollars more than Joel.
Lisa has 4 dollars more than Joel.	Joel has 3 dollars less than Tom.
How many dollars does Lisa have?	How many dollars does Lisa have?

VI. REFERENCES

Behr, M. J., & Post, T. R. (1981). The effect of visual perceptual distracters on children's logical—mathematical thinking in rational-number situations. In T. Post & M. P. Roberts (Eds.), Proceedings of the Third Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (pp.8-16). Minneapolis: University of Minnesota.

Briars, D. J., & Larkin, J. H. (1984). An integrated model of skill in solving elementary word problems. Cognition and Instruction, 1, 245-296.

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

Craik, E. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Cummins, D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. Cognitive Psychology, 20, 405-438.

De Corte, E., & Verschaffel, L.(1985). Beginning first graders' initial representation of arithmetic word problems. Journal of Mathematical Behavior, 4, 3-21.

De Corte, E., & Verschaffel, L. (1987). Using retelling data to study children's word problem solving. In D. Rogers & J. Sloboda (Eds.), Cognitive Processes in Mathematics (pp.42-59). Oxford, England: Clarendon Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87, 215-51.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis. Verbal reports as data*. MIT Press, Cambridge, MA.

Fuson, K. C., Carroll, W. M., & Landis, J. (1992). Developmental sequences of conceptual structures for children's solution of addition and subtraction word problems. Paper presented at a workshop on "word problems in algebra and arithmetic: From language comprehension to problem solving. " Max-Planck- Institute for psychologische Forschung Mumchen, Germany.

Howell, D. C. (1991). Statistical Methods for Psychology (3rd Ed.). Duxbury.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review, 95*(2), p. 163-182*.

Kirk, R. E.(1968). Experimental Design: Procedures for the Behavioral Sciences. Belmont, CA: Brooks/Cole.

Kirk, R. E. (1982). Experimental Design: Procedures for the Behavioral Sciences (2nd Ed.) CA: Wadsworth.

Lewis, A., & Mayer, R. (1987). Students' miscomprehension of relational statements in arithmetic word problems. Journal of Educational Psychology, 79, 363-371.

Lin, P. J. (1997). Children's cognitive processes in solving two-step compare problems. Dissertation of university of Minnesota (Unpublished).

Mehrens, W. A. & Lehmann, I. J. (1991). Measurement and Evaluation in Education and Psychology. Florida: Holt, Rinehart and Winston.

Morales, R. V., Shute, V. J., & Pellegrino, J. W. (1985). Developmental differences in understanding and solving simple word problems. Cognition and Instruction, 2, 41-57.

Nesher, P. & Hershkovitz, S. (1994). The role of schemes in two-step problems: Analysis and research findings. *Educational Studies in Mathematics*, 26, pp. 1-23.

Nesher, P., & Teubal, E., (1975). Verbal cues as an interfering factor in verbal problem solving. *Educational Studies in Mathematics*, 6, 41-51.

Quintero, A. H.(1980). The role of semantic understanding in solving multiplication word problems (Doctoral dissertation, Massachusetts Institute of Technology. Dissertation Abstraction Institutional, 1980, 41,2522A.

Quintero, A. H. (1983). Conceptual understanding in solving two-step word problems with a ratio. Journal for Research in Mathematics Education,.14(2), 102-112.

Resnick, L. B., & Greeno,J. G. (1990) Conceptual Growth of Number and Quantity. Unpublished manuscript. University of Pittsburgh.

Lin, P.J. (1998). A Study of Children's Comprehension Processes by Retelling After Solving Technique. 國立新竹師院學報，第十一期，187-227。

Riley, M. S., Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. Cognition and Instruction, 5, 49-101.

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic, In H. P. Ginsberg (Eds.) The development of mathematical thinking (pp. 153-196). San Diego, CA: Academic Press.

Stern, E. (1993). What makes certain arithmetic word problems involving the comparison of sets so hard for children? Journal of Educational Psychology, 85, 7-23.

Stern, E., & Leherndorfer, A.(1993). The role of situational context in solving word problems. Cognitive Development, 2, 259-268.

Verschaffel, L. (1984). First graders' representations and solutions of elementary arithmetic word problems. Dissertation of Center for Instructional Psychology and Technology, University of Leuven, Belgium.

Verschaffel, L. (1994) Using retelling data to study elementary school children's representations and solutions of compare problems. Journal for Research in Mathematics Education, 25(2), 141-166.

Verschaffel, L., De Corte, E., & Pauwels, A. (1992). Solving compare problems: An eye-movement test of Lewis and Mayer's consistency hypothesis. Journal of Educational Psychology, 84, 85-94.

Winer, B. J. (1971). Statistical Principles in Experimental Design (2nd Ed.). New York: McGraw-Hill.