THE UNIVERSITY OF CALGARY

The Effect of Problem Text on Solving Difference-Finding Word Problems

by

Ning Fan

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

CALGARY, ALBERTA

FEBRUARY, 1993

ISBN 0-315-83142-1

Name NING FAN

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

Educational Psychology (problem solving) | 0525 | U·M·I

SUBJECT TERM | SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523
Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
General 0679
Ancient 0289
Linguistics 0290
Modern 0291
Literature
General 0401
Classical 0294
Comparative 0295
Medieval 0297
Modern 0298
African 0316
American 0591
Asian 0305
Canadian (English) 0352
Canadian (French) 0355
English 0593
Germanic 0311
Latin American 0312
Middle Eastern 0315
Romance 0313
Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
General 0318
Biblical Studies 0321
Clergy 0319
History of 0320
Philosophy of 0322
Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
Archaeology 0324
Cultural 0326
Physical 0327
Business Administration
General 0310
Accounting 0272
Banking 0770
Management 0454
Marketing 0338
Canadian Studies 0385
Economics
General 0501
Agricultural 0503
Commerce-Business 0505
Finance 0508
History 0509
Labor 0510
Theory 0511
Folklore 0358
Geography 0366
Gerontology 0351
History
General 0578
Ancient 0579
Medieval 0581
Modern 0582
Black 0328
African 0331
Asia, Australia and Oceania 0332
Canadian 0334
European 0335
Latin American 0336
Middle Eastern 0333
United States 0337
History of Science 0585
Law 0398
Political Science
General 0615
International Law and Relations 0616
Public Administration 0617
Recreation 0814
Social Work 0452
Sociology
General 0626
Criminology and Penology 0627
Demography 0938
Ethnic and Racial Studies 0631
Individual and Family Studies 0628
Industrial and Labor Relations 0629
Public and Social Welfare 0630
Social Structure and Development 0700
Theory and Methods 0344
Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
Agriculture
General 0473
Agronomy 0285
Animal Culture and Nutrition 0475
Animal Pathology 0476
Food Science and Technology 0359
Forestry and Wildlife 0478
Plant Culture 0479
Plant Pathology 0480
Plant Physiology 0817
Range Management 0777
Wood Technology 0746
Biology
General 0306
Anatomy 0287
Biostatistics 0308
Botany 0309
Cell 0379
Ecology 0329
Entomology 0353
Genetics 0369
Limnology 0793
Microbiology 0410
Molecular 0307
Neuroscience 0317
Oceanography 0416
Physiology 0433
Radiation 0821
Veterinary Science 0778
Zoology 0472
Biophysics
General 0786
Medical 0760

EARTH SCIENCES
Biogeochemistry 0425
Geochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
General 0566
Audiology 0300
Chemotherapy 0992
Dentistry 0567
Education 0350
Hospital Management 0769
Human Development 0758
Immunology 0982
Medicine and Surgery 0564
Mental Health 0347
Nursing 0569
Nutrition 0570
Obstetrics and Gynecology 0380
Occupational Health and Therapy 0354
Ophthalmology 0381
Pathology 0571
Pharmacology 0419
Pharmacy 0572
Physical Therapy 0382
Public Health 0573
Radiology 0574
Recreation 0575
Speech Pathology 0460
Toxicology 0383
Home Economics 0386

PHYSICAL SCIENCES
Pure Sciences
Chemistry
General 0485
Agricultural 0749
Analytical 0486
Biochemistry 0487
Inorganic 0488
Nuclear 0738
Organic 0490
Pharmaceutical 0491
Physical 0494
Polymer 0495
Radiation 0754
Mathematics 0405
Physics
General 0605
Acoustics 0986
Astronomy and Astrophysics 0606
Atmospheric Science 0608
Atomic 0748
Electronics and Electricity 0607
Elementary Particles and High Energy 0798
Fluid and Plasma 0759
Molecular 0609
Nuclear 0610
Optics 0752
Radiation 0756
Solid State 0611
Statistics 0463

Applied Sciences
Applied Mechanics 0346
Computer Science 0984
Engineering
General 0537
Aerospace 0538
Agricultural 0539
Automotive 0540
Biomedical 0541
Chemical 0542
Civil 0543
Electronics and Electrical 0544
Heat and Thermodynamics 0348
Hydraulic 0545
Industrial 0546
Marine 0547
Materials Science 0794
Mechanical 0548
Metallurgy 0743
Mining 0551
Nuclear 0552
Packaging 0549
Petroleum 0765
Sanitary and Municipal 0554
System Science 0790
Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

Nom ____

Dissertation Abstracts International est organisé en catégories de sujets. Veuillez s.v.p. choisir le sujet qui décrit le mieux votre thèse et inscrivez le code numérique approprié dans l'espace réservé ci-dessous.

______________________________ SUJET

☐☐☐☐ CODE DE SUJET

U·M·I

Catégories par sujets

HUMANITÉS ET SCIENCES SOCIALES

COMMUNICATIONS ET LES ARTS

Architecture 0729
Beaux-arts 0357
Bibliothéconomie 0399
Cinéma 0900
Communication verbale 0459
Communications 0708
Danse 0378
Histoire de l'art 0377
Journalisme 0391
Musique 0413
Sciences de l'information 0723
Théâtre 0465

ÉDUCATION

Généralités 515
Administration 0514
Art 0273
Collèges communautaires 0275
Commerce 0688
Économie domestique 0278
Éducation permanente 0516
Éducation préscolaire 0518
Éducation sanitaire 0680
Enseignement agricole 0517
Enseignement bilingue et multiculturel 0282
Enseignement industriel 0521
Enseignement primaire 0524
Enseignement professionnel 0747
Enseignement religieux 0527
Enseignement secondaire 0533
Enseignement spécial 0529
Enseignement supérieur 0745
Évaluation 0288
Finances 0277
Formation des enseignants 0530
Histoire de l'éducation 0520
Langues et littérature 0279
Lecture 0535
Mathématiques 0280
Musique 0522
Orientation et consultation 0519
Philosophie de l'éducation 0998
Physique 0523
Programmes d'études et enseignement 0727
Psychologie 0525
Sciences 0714
Sciences sociales 0534
Sociologie de l'éducation 0340
Technologie 0710

LANGUE, LITTÉRATURE ET LINGUISTIQUE

Langues
- Généralités 0679
- Anciennes 0289
- Linguistique 0290
- Modernes 0291

Littérature
- Généralités 0401
- Anciennes 0294
- Comparée 0295
- Mediévale 0297
- Moderne 0298
- Africaine 0316
- Américaine 0591
- Anglaise 0593
- Asiatique 0305
- Canadienne (Anglaise) 0352
- Canadienne (Française) 0355
- Germanique 0311
- Latino-américaine 0312
- Moyen-orientale 0315
- Romane 0313
- Slave et est-européenne 0314

PHILOSOPHIE, RELIGION ET THÉOLOGIE

Philosophie 0422
Religion
- Généralités 0318
- Clergé 0319
- Études bibliques 0321
- Histoire des religions 0320
- Philosophie de la religion 0322

Théologie 0469

SCIENCES SOCIALES

Anthropologie
- Archéologie 0324
- Culturelle 0326
- Physique 0327

Droit 0398
Économie
- Généralités 0501
- Commerce-Affaires 0505
- Économie agricole 0503
- Économie du travail 0510
- Finances 0508
- Histoire 0509
- Théorie 0511

Études américaines 0323
Études canadiennes 0385
Études féministes 0453
Folklore 0358
Géographie 0366
Gérontologie 0351
Gestion des affaires
- Généralités 0310
- Administration 0454
- Banques 0770
- Comptabilité 0272
- Marketing 0338

Histoire
- Histoire générale 0578
- Ancienne 0579
- Médiévale 0581
- Moderne 0582
- Histoire des noirs 0328
- Africaine 0331
- Canadienne 0334
- États-Unis 0337
- Européenne 0335
- Moyen-orientale 0333
- Latino-américaine 0336
- Asie, Australie et Océanie 0332

Histoire des sciences 0585
Loisirs 0814
Planification urbaine et régionale 0999
Science politique
- Généralités 0615
- Administration publique 0617
- Droit et relations internationales 0616

Sociologie
- Généralités 0626
- Aide et bien-àtre social 0630
- Criminologie et établissements pénitentiaires 0627
- Démographie 0938
- Études de l' individu et de la famille 0628
- Études des relations interethniques et des relations raciales 0631
- Structure et développement social 0700
- Théorie et méthodes 0344
- Travail et relations industrielles 0629

Transports 0709
Travail social 0452

SCIENCES ET INGÉNIERIE

SCIENCES BIOLOGIQUES

Agriculture
- Généralités 0473
- Agronomie. 0285
- Alimentation et technologie alimentaire 0359
- Culture 0479
- Élevage et alimentation 0475
- Exploitation des péturages 0777
- Pathologie animale 0476
- Pathologie végétale 0480
- Physiologie végétale 0817
- Sylviculture et faune 0478
- Technologie du bois 0746

Biologie
- Généralités 0306
- Anatomie 0287
- Biologie (Statistiques) 0308
- Biologie moléculaire 0307
- Botanique 0309
- Çellule 0379
- Écologie 0329
- Entomologie 0353
- Génétique 0369
- Limnologie 0793
- Microbiologie 0410
- Neurologie 0317
- Océanographie 0416
- Physiologie 0433
- Radiation 0821
- Science vétérinaire 0778
- Zoologie 0472

Biophysique
- Généralités 0786
- Medicale 0760

SCIENCES DE LA TERRE

Biogéochimie 0425
Géochimie 0996
Géodésie 0370
Géographie physique 0368
Géologie 0372
Géophysique 0373
Hydrologie 0388
Minéralogie 0411
Océanographie physique 0415
Paléobotanique 0345
Paléoécologie 0426
Paléontologie 0418
Paléozoologie 0985
Palynologie 0427

SCIENCES DE LA SANTÉ ET DE L'ENVIRONNEMENT

Économie domestique 0386
Sciences de l'environnement 0768
Sciences de la santé
- Généralités 0566
- Administration des hipitaux 0769
- Alimentation et nutrition 0570
- Audiologie 0300
- Chimiothérapie 0992
- Dentisterie 0567
- Développement humain 0758
- Enseignement 0350
- Immunologie 0982
- Loisirs 0575
- Médecine du travail et thérapie 0354
- Médecine et chirurgie 0564
- Obstétrique et gynécologie 0380
- Ophtalmologie 0381
- Orthophonie 0460
- Pathologie 0571
- Pharmacie 0572
- Pharmacologie 0419
- Physiothérapie 0382
- Radiologie 0574
- Santé mentale 0347
- Santé publique 0573
- Soins infirmiers 0569
- Toxicologie 0383

SCIENCES PHYSIQUES

Sciences Pures

Chimie
- Genéralités 0485
- Biochimie 487
- Chimie agricole 0749
- Chimie analytique 0486
- Chimie minérale 0488
- Chimie nucléaire 0738
- Chimie organique 0490
- Chimie pharmaceutique 0491
- Physique 0494
- PolymÇres 0495
- Radiation 0754

Mathématiques 0405
Physique
- Genéralités 0605
- Acoustique 0986
- Astronomie et astrophysique 0606
- Electronique et électricité 0607
- Fluides et plasma 0759
- Météorologie 0608
- Optique 0752
- Particules (Physique nucléaire) 0798
- Physique atomique 0748
- Physique de l'état solide 0611
- Physique moléculaire 0609
- Physique nucléaire 0610
- Radiation 0756

Statistiques 0463

Sciences Appliqués Et Technologie

Informatique 0984
Ingénierie
- Généralités 0537
- Agricole 0539
- Automobile 0540
- Biomédicale 0541
- Chaleur et ther modynamique 0348
- Conditionnement (Emballage) 0549
- Génie aérospatial 0538
- Génie chimique 0542
- Génie civil 0543
- Génie électronique et électrique 0544
- Génie industriel 0546
- Génie mécanique 0548
- Génie nucléaire 0552
- Ingénierie des systàmes 0790
- Mécanique navale 0547
- Métallurgie 0743
- Science des matériaux 0794
- Technique du pétrole 0765
- Technique minière 0551
- Techniques sanitaires et municipales 0554
- Technologie hydraulique 0545

Mécanique appliquée 0346
Géotechnologie 0428
Matières plastiques (Technologie) 0795
Recherche opérationnelle 0796
Textiles et tissus (Technologie) 0794

PSYCHOLOGIE

Généralités 0621
Personnalité 0625
Psychobiologie 0349
Psychologie clinique 0622
Psychologie du comportement 0384
Psychologie du développement 0620
Psychologie expérimentale 0623
Psychologie industrielle 0624
Psychologie physiologique 0989
Psychologie sociale 0451
Psychométrie 0632

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "The Effect of Problem Text on Solving Difference-Finding Word Problems" submitted by Ning Fan in partial fulfillment of the requirements for the degree of Master of Science.

John H Mueller

Dr. J. H. Mueller, Supervisor

Department of Educational Psychology

Dr. A. McKeough

Department of Educational Psychology

A. Marini

Dr. A. E. Marini

Department of Teacher Education and Supervision

February 24, 1993

(Date)

ABSTRACT

This study investigated the effect of problem text on arithmetic word problem solving involving difference finding. Data on correct solutions and strategies for three types of such word problem, namely COMPARE, EQUALIZE, and WON'T GET, were collected from first-grade students and analyzed by repeated-measures analyses of variance. The correct solutions for the EQUALIZE and WON'T GET problems were found to be significantly higher than for the COMPARE problems. The dependency of strategy use on the problem text was also found, specifically, the EQUALIZE problems were most frequently solved by using an ADD-ON strategy, and the WON'T GET problems by a MATCH strategy, which reflected the construction of a coordination between two mental number lines as the problem representation. There was no one strategy used significantly more than others for the COMPARE problems, suggesting that the problem text of COMPARE did not facilitate the coordination.

ACKNOWLEDGEMENTS

The writer wishes to express his gratitude and sincere thanks to the people who made the completion of this thesis possible: Dr. J. Mueller for his constant guidance and encouragement; Dr. A. McKeough for her encouragement and help with theoretical issues; Dr. A. Marini for his support through the years; Dr. M. Pyryt and Dr. T. Fung for their support on statistical design and analyses; and the administrations, teachers, and students of participating schools for their cooperation. The writer would also like to thank his wife, Xujing Zhu, for her constant support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The focus of this study is the effect of problem texts on simple arithmetic word problem solving, specifically problems with the same mathematical content and structure but having different wording. Studies have found the change of problem wording influences the relative difficulty of the problem (Hudson, 1983; Carpenter, 1985; De Corte, Verschaffel, & De Win, 1985; Cummins, Kintsch, Reusser, & Weimer, 1988; Cummins, 1991; Okamoto, 1992). However, the explanation of how problem texts affect the performance of problem solving remains unclear. Some authors suggested that it could be attributed to the alteration of problem semantic structure (Riley, Greeno, & Heller, 1983; De Corte et al., 1985). Other authors focus on the mathematical knowledge required to represent the semantic relations (Riley & Greeno, 1988; Resnick, 1989; Gelman & Greeno, 1989). Still others oppose the mathematical viewpoint, and proclaim a linguistic development view emphasizing the effect of the problem text comprehension (Cummins, 1991; Hudson, 1983).

Because of the different perspectives of the researchers and the various types of problem covered by the studies, the current situation in this field perfectly fits what T. P. Carpenter (1985) described: "It is clear that differences in wording contribute to a problem's difficulty, but it is not at all clear exactly how. ...beyond knowing that certain wordings are more difficult, we have a much less precise picture of how differences in wording influence children's solutions." (p. 26)

The existing controversy, on the other hand, reflects that word problem solving has become the common interest of researchers from different perspectives. It has become relevant to such areas as problem-solving processes, mathematical knowledge required for problem solving, text processing, linguistic knowledge required for problem solving, and cognitive development of problem solving. Therefore, the joint effort of researchers from various perspectives has provided the opportunity to study the complexity of word problem solving.

The problems involved in research of this field are simple one-step arithmetic addition and subtraction problems. On the one hand, the simplicity provides a way to model the problem solving processes, and to infer the knowledge required to solve the problems; on the other hand, the processes and knowledge involved in problem solving show the complexity of the research questions raised from these simple problems.

The interest of the present study was derived from an issue with the difficulty in solving the word problems involving comparison. Young children up to first grade find it difficult to solve problems stated this way: "Tom has 8 marbles; Joe has 5 marbles; How many more marbles does Tom have than Joe?". However, Hudson (1983) found that the problem became much easier when the text was changed to "There are 5 birds; There are 3 worms; Suppose the birds race over and each tries to get a worm; How many birds won't get a worm?". What made this difference? What is the effect of the problem texts? This phenonenon has been related to almost every fundamental aspect of word problem solving, such as the conceptual understanding of word problems, the construction of problem representations, the required linguistic knowledge, the required mathematical knowledge, and the development of children's word problem solving ability. Research on this topic is extensive, but, as mentioned above, controversy exists.

What is the effect of problem text? How does problem text affect the problem solving process? The present study's research problem derives from the controversy. Also, from the previous studies, the present study adopted a fundamental framework for answering the questions.

The theoretical framework regarding the word problem solving process adopted by the present study assumes that (1) successful solution to a problem relies on the conceptual understanding of the problem, that is, building a coherent mental conceptual representation of the problem (Riley et al, 1983; kintsch & Greeno, 1985); (2) the representation is a dual one including a text base consisting of information given in the problem text, and a problem model in which information from the text base is reconstructed in terms of a mathematical structure acquired by the problem solver through development (kintsch & Greeno, 1985); (3) the mapping between text comprehension and mathematical reasoning occurs when information from the text base is being reconstructed in the problem model (kintsch & Greeno, 1985); (4) the problem model is represented as two mental number lines which are coordinated to produce an answer to the problem (Okamoto, 1992); and (5) the problem solving strategy is determined by the problem model, and the observed strategy can be used as an indicator of what kind of problem model is being constructed (Riley et al, 1983; Carpenter, Hiebert, & Moser, 1981).

To study the effect of problem text, the present study employed three types of arithmetic word problems, namely COMPARE, EQUALIZE and WON'T GET. They shared the same mathematical essence, that is, finding the numerical difference between two disjoint sets, and the difference among them is only in problem text. By observing and analyzing the first-grade students' problem solving processes on these three types of problems, this study attempted to investigate the mapping between the liguistic comprehension and mathematical reasoning, and to study this mapping process in detail, in order to obtain data about the effect of the problem texts on the construction of the problem representations. This in turn would reveal the source of difficulty in solving the problems involving comparison.

This study also attempted to investigate the effect of using concrete materials on the problem representations. This is partly because the presence or absence of concrete materials in previous studies has not been considered as a variable affecting problem representation, but some facilitating effect on performance has been found (Riley et al., 1983). Also, knowing whether any interaction between problem texts and concrete materials exists would help clarify the effect of problem text.

CHAPTER II

LITERATURE REVIEW

Conceptual Understanding in Word Problem Solving

Problem Categorization

Studying word problem solving processes starts from the analyses of the problem themselves. The word problems involved in this study are simple arithmetic problems presented in verbal forms. Simply one-step addition and subtraction are involved in the problems, but the problems vary in terms of problem length, grammatical complexity, and order of problem statements, etc. This has led to the issue of word problem classification. Several approaches concerning these variables have been taken (see Riley et al., 1983, for a review). Those classifications deal more with the syntax than semantics of word problems. In order to study conceptual understanding in word problem solving processes, word problems are categorized according to analyses of their semantic structures.

A word problem identifies some quantities and describes a relationship among them. Based on the semantic relations among quantities, three types of word problem, namely, Change, Combine and Compare (see Table 1), were categorized by Heller and Greeno (1978). In Change problems, the relationship among quantities is the action that causes increases or decreases in some quantity. The initial quantity is referred to as the "start set", the increased or decreased quantity as the "change set", and the resulting quantity as the "result set". Both Combine and Compare problems involve static relationships for which there is no direct or implied action. Combine problems involve the relationship about the union of two distinct quantities. The union is frequently referred to as the "superset", and the other two quantities as the "subsets". In Compare problems, the relationship is the comparison between two distinct, disjoint quantities. Since one quantity is compared to the other, it is possible to label one quantity as the "referent set" and the other as the "compared set". The third quantity is the "difference", or the amount by which the larger quantity exceeds the other.

In the study by Carpenter and Moser (1982), another type, Equalize problems (see Table 1), was included. The Equalize type is a hybrid of Compare and Change problems. There is the same sort of action as found in Change problems, but the action is based on the comparison between the two disjoint quantities.

In addition to the various semantic relations, there are other dimensions for which the word problems in Table 1 differ. One is the identity of the unknown quantity. Within each type of problem, different problems can be formed by varying the given quantity and the unknown quantity. In Change problems, any of the "start set", the "change set" or the "result set" can be unknown if the other two are given, yielding three different cases. Furthermore, the direction of change action can either be an increase or a decrease, so

Table 1

Types of Word Problems

Change

Result Unknown

1. Joe had 3 marbles.
   Then Tom gave him 5 more marbles.
   How many marbles does Joe have now.
2. Joe had 8 marbles.
   Then he gave 5 marbles to Tom.
   How many marbles does Joe have now.

Change Unknown

3. Joe had 3 marbles.
   Then Tom gave him some more marbles.
   Now Joe has 8 marbles.
   How many marbles did Tom give him.
4. Joe had 8 marbles.
   Then he gave some marbles to Tom.
   Now Joe has 3 marbles.
   How many marbles did he give to Tom.

Start Unknown

5. Joe had some marbles.
   Then Tom gave him 5 more marbles.
   Now Joe has 8 marbles.
   How many marbles did Joe have in the beginning?
6. Joe had some marbles.

(Continued next page)

Table 1 (Continued)

Then he gave 5 marbles to Tom.

Now Joe has 3 marbles.

How many marbles did Joe have in the beginning?

Combine

Superset Unknown

1. Joe has 3 marbles.

   Tom has 5 marbles.

   How many marbles do they have altogether?

Subset Unknown

2. Joe and Tom have 8 marbles altogether.

   Joe has 3 marbles.

   How many marbles does Tom have?

Compare

Difference Unknown

1. Joe has 8 marbles.

   Tom has 5 marbles.

   How many more marbles does Joe have than Tom?

2. Joe has 8 marbles.

   Tom has 5 marbles.

   How many less marbles does Tom have than Joe?

Compared Quantity Unknown

3. Joe has 3 marbles.

   Tom has 5 more marbles than Joe.

   How many marbles does Tom have?

(Continued next page)

Table 1 (Continued)

---

4. Joe has 8 marbles.

   Tom has 5 fewer marbles than Joe.

   How many marbles does Tom have?

<u>Referent Unknown</u>

5. Joe has 8 marbles.

   He has 5 more marbles than Tom.

   How many marbles does Tom have?

6. Joe has 3 marbles.

   He has 5 fewer marbles than Tom.

   How many marbles does Tom have?

<u>Equalize</u>

<u>Increase</u>

1. Joe has 3 marbles.

   Tom has 8 marbles.

   What could Joe do to have as many marbles as Tom?

   (How many more marbles does Joe have to get to have as many as Tom?)

<u>Decrease</u>

2. Joe has 8 marbles.

   Tom has 3 marbles.

   What could Joe do to have as many marbles as Tom?

   (How many marbles does Joe have to lose to have as many marbles as Tom?)

---

that there are six kinds of Change problems in total. In Combine problems, either the "superset" or one "subset" can be unknown, yielding two sorts of problems. In Compare problems, the unknown quantity may be any of the "referent set", the "compared set", or the "difference", with the direction of difference may be "more" or "less", totally yielding six variations. Six different cases (two shown in Table 1) can also be produced in Equalize problems by varying the unknown among the three quantities analogous to those in Compare problems, and by indicating the different directions of action as in Change problems. Classifying the word problems based on both semantic relations and identity of unknowns results in the problem types in Table 1.

Relative Difficulty

The word problems in Table 1 involve either addition or subtraction as operations for solution. However, problems which require the same operation are not equally difficult. There is a strong influence from semantic structures by which the problems are described. This has been evident in many empirical studies (Carpenter, Hiebert, & Moser, 1981; Riley, 1981; Tamburino, 1980). They found, separately, that Compare-3 and Compare-6 problems are more difficult than either Change-1 or Combine-1 problems, although all four problem solutions involve a simple addition. Similarly, problems involving subtraction can also vary in difficulty across semantic structures. Combine-2 problems and virtually all Compare problems are, in general, more difficult than

Change-2 and Change-4 problems. These findings are consistent with other studies. It has been found that Compare-1 problems are more difficult than Change-2 problems for first-graders (Gibb, 1956; Schell & Burns, 1962; Shores & Underhill, 1976). It has been also found that Combine-2 problems are, in general, more difficult than Change-2 for kindergartners and first-graders (Gibb, 1956; Ibarra & Lindvall, 1979; LeBlanc, 1968; Nesher & Katriel, 1978; Vergnaud, 1982), but are slightly easier than Compare-1 problems (Schell & Burns, 1962).

Attempting to partly explain the influence of a problem's semantic structure on children's solutions to word problems, Riley et al. (1983) suggested that the various semantic structures may correspond to some specific concepts, such as the concepts of quantitative change, equalization, combination, and comparison. They also speculated that these concepts emerge at different times in cognitive development. For example, at a certain age, a specific child might have the concepts of change and combination, but not the concept of comparison.

In addition, problems having the same semantic structure also vary in difficulty. This is the effect of identity of the unknown quantity. Some studies (Carpenter, Hiebert, & Moser, 1981; Riley, 1981; Tamburino, 1980) found children have no difficulty solving Change problems when the "start set" and the "change set" are given and the "result set" is unknown. Even preschool children can solve these problems (Buckingham & MacLatchy, 1930; Hebbeler, 1977). However, many kindergartners and first-graders have difficulty if the "start set" and "result set" are given and they are asked to find "change set". As for Change-5 and Change-6 problems, when "result set" and "change set" are given with "start set" unknown, they become more difficult for even second- and third-graders (Riley, 1981; Hiebert, 1981; Lindvall & Ibarra, 1980a; Vergnaud, 1982), and even more difficult than Combine-2 and Compare-1 problems. Within Combine and Compare problems, difficulty also varies depending on which quantity in a problem is unknown. Combine-2 problems in which one of the subset is unknown are significantly more difficult than Combine-1 problems in which the two subsets are given with "superset" unknown. Compare-5 and Compare-6 problems in which "referent set" is unknown are more difficult than any of the other Compare problems (Riley et al., 1983).

Evidently, word problems differ in the semantic structures as well as in the identity of unknown quantity. The resulting problem types (Table 1) have been employed in studies to reveal the fairly systematic differences in children's performance. Furthermore, the analyses of problem semantic structures serve as the basis for studying the word problem solving processes and children's knowledge required to solve the problems.

RGH Model

Riley, Greeno, and Heller (1983) developed a computer-implemented model, to be called the RGH model (De Corte & Verschaffel, 1988), constituting an account of the internal processes and cognitive structures underlying children's performance on word problems. In this model, three main kinds of knowledge during word problem solving are proposed: (a) problem schemata for understanding various semantic relations described in problem texts; (b) action schemata for representing the model's knowledge about actions involved in problem solutions; and (c) strategic knowledge for planning solutions to problems. When the model is given a word problem to solve, it uses its knowledge of problem schemata to represent the particular problem situation being described. The model's planning procedures then use action schemata to generate a solution to the problem.

Among the three main components of knowledge needed for successful performance, it is proposed that the main source of children's difficulty is not their lack of knowledge about the actions required the solve certain problems. Instead, the main locus of children's improvement in problem solving skill is in the acquisition of schemata for understanding a problem in a way that relates it to already available action schemata. This hypothesis appears to be supported in some studies involving children's performance on slightly reworded Combine-2 problems compared with the performance on Combine 2's stereotype. Consider this as an example: "Joe and Tom have 8 marbles altogether. Joe has 5 marbles. How many marbles does Tom have?". Although the solution procedure for this problem involves three simple actions, namely "make-set", "take-out", and "count-all", which most children have available since they can use it to solve Change-2 problems, most first-graders still find it difficult. However, when the problem wording is slightly changed, their performance is changed. Carpenter et al. (1981) reported that thirty-three out of forty-three first-graders solved it correctly when the problem was changed to: "There are 6 children on the playground. 4 are boys and <u>the rest</u> are girls. How many girls are on the playground?". Another variation, "Together, Tom and Joe have 8 apples. Three of these apples belong to Tom. How many <u>of them</u> belong to Joe?", was found significantly easier for kindergartners than the stereotype by Lindvall and Ibarra (1980b). What is done by the above rewording is that the variations make the relationship among quantities more explicit. Therefore, Riley et al. (1983) concluded that successful solution relies on children's ability to understand problems, that is, the ability to represent the relationships among quantities described in problem situations.

The RGH model emphasizes that the locus of improvement in word problem solving skill lies in the acquisition and development of problem schemata. Based on the analyses of problem semantic structures, three main types of problem schemata are proposed for understanding Change, Combine, and Compare problems. Understanding a problem is defined as building a schematized problem representation. The representation has the form of semantic network structures consisting of elements and relations between the elements. Within each type of schema, three levels of conceptual development are hypothesized. The main differences between the levels relate to the ways in which information is represented and the ways in which quantitative information is manipulated. Those with more detailed representational schemata and more sophisticated action schemata represent the more advanced levels of problem solving skill. Consider Change problems as an example: Level-1 understands quantitative relations by means of a simple schema that limits the representations of Change problems to the external display of objects. This knowledge is sufficient to solve Change-1, Change-2, and Change-4 problems. Level-2 has a Change schema for maintaining an internal representation of increases and decreases in the set of objects it manipulate; the process of building this representation is still relatively "bottom-up" in the sense that it still depends upon the external display of objects. Because of the richer understanding of relationships between quantities and a richer set of action schemata, Level-2 can solve Change-3 problems. Level-3 also has a Change schema for representing relations internally, but it can use its Change schema in a more "top-down" fashion. It has an understanding of part-whole relations, as well as a richer set of action schemata. By transforming the more complicated Change relations into part-whole relations, Level-3 can solve Change-5 and Change-6 problems.

These three levels of development are paralleled in solving Combine and Compare problems. That is, at the lowest level, the child's representations of problems are limited to the external displays of objects; at an intermediate level there are schemata for representing, internally, additional information about relationships among quantities; and at the most advanced level, schemata are available that direct problem representations and solutions in a more top-down manner.

Factors in Conceptual Understanding

The RGH model (1983) was the first attempt to systematically model children's word problem solving processes and their development in terms of conceptual understanding. This model has become influential in studies of word problem solving. First, its emphasis on conceptual understanding is consistent with that in other analyses of mathematical problem solving. Mayer (1985, 1986) distinguished representation and solution as two components of problem solving, and concluded that the construction of an appropriate conceptual problem representation is the crucial component. Second, the problem categorization on which its analyses are based has been widely accepted in researches (De Corte & Verschaffel, 1985, 1987; De Corte, Verschaffel, & De Win, 1985; Carpenter, 1985; Kintsch & Greeno, 1985; Riley & Greeno, 1988; Cummins, Kintsch, Reusser, & Weimer, 1988; Cummins, 1991; Okamoto, 1992). Third, its use of semantic structure analyses as a basis of modelling problem solving processes has been verified in other studies. De Corte & Verschaffel (1985, 1987) have found a strong influence of semantic structures on problem representations and on solution strategies. Carpenter (1985) concluded that we have a reasonably clear picture of how semantic structure affects children's solution processes (p. 26). Finally, relating problem relative difficulty to the development of children's conceptual understanding helps identify cognitive structures or knowledge underlying problem solving skills. This approach is also used in other analyses (Briars & Larkin, 1984; Okamoto, 1992).

The RGH model pioneers the systematic analysis of conceptual understanding in word problem solving. However, the model itself has some limitations in its characterization of the "conceptual understanding". The first involves the role of problem schemata. When the RGH model is given a word problem, it uses a problem schema, e.g. Change, Combine, or Compare, to represent the particular situation being described; with the schematic representation, appropriate action schemata are associated by procedure attachments. Also, the development of conceptual understanding is proposed to be within each of these problem categories. This model does not include Carpenter and Moser's (1981) Equalize problems, but these might be put in another category and thus another problem schema must be proposed. The more problem types that are involved, the more problem schemata must be proposed. Therefore, understanding and its development are problem type dependent, and the theory loses its generality across problem types. Another model of word problem solving, Briars and Larkin's (1984) model CHIPS (Concrete Human-like Inferential Problem Solver), has no distinct schemata for representing the different categories of problems and works well. Eventually, in Riley and Greeno's (1988) revision of the RGH model, the problem schemata are no longer essential to the problem solving processes.

The second limitation is related to the nature of hypothesized cognitive structures which underlie conceptual understanding and its development. In the RGH model, the improvement in representing and solving word problems relies on the increasingly sophisticated structures representing set relations. At Level 1, a single set can be represented but without relations to other sets. At Level 2, a certain sort of set relation can be represented internally. Finally at Level 3, all other kinds of relations are assimilated into part-whole relations. Despite the constraints by problem types, all the proposed representing structures are mathematical in nature. According to this model, children's difficulty with certain types of problems can be attributed to the lack of such mathematical structures. This is the hypothesis which has been widely opposed.

De Corte and Verschaffel (1985) suggested that, besides mathematical schemata, word problem solving skill development depends on a more general "word-problem schema" that indicates the structure, role, and intent of word problems in general. The main function of such a word-problem schema is to encode implicit rules, suppositions, and agreements concerning typical word problems that will enable a problem solver to interpret ambiguities correctly and to compensate for insufficiencies in the problem text. In a word problem solving process model developed by Reusser (1990), and

Reusser, Kämpfer, Sprenger, Staub, Stebler, and Stüssi (1990), a concrete, intermediary, non-mathematical "situation model" is proposed to be established before the building of mathematical representation. Mathematical structures alone cannot be sufficient to constitute understanding of word problems. More importantly, since a word problem is a mathematical problem presented in verbal form, language comprehension also plays a crucial role in understanding. With regard to language understanding, the RGH model obviously shows some limitation and has been criticized by many other researchers (Kintsch & Greeno, 1985; Cummins et al., 1988; Commins, 1991). As related to the present study, when the RGH model encounters Compare problems, the limitation becomes more obvious.

The Issue With Compare Problems

Hudson's Study

Compare problems are usually identified as the most difficult type of problems among those in Table 1. In Carpenter et al.'s (1981) study, 81% of first-graders correctly solved Compare-1 problems. This proportion is lower than the proportions for Change-1, Change-2, Equalize-1, Equalize-2, and Combine-1 problems. Furthermore, there was only 28% of first-graders who correctly solved Compare-3 problems. Riley (1981) reported that 17% of kindergartners, 28% of first-graders, and 85% of second-graders could solve

Compare 1 problems correctly. Only third-graders could solve the problems at 100%. Other Compare problems were more difficult for children at each grade level. Again, Compare problems in general were more difficult than Change and Combine problems at any grade level.

However, Hudson's (1980, 1983) study on the word problems involving comparison showed an interesting result. Hudson presented drawings with, for example, five birds and four worms to the subjects, and asked two different questions. One is the question identical to a Compare-1 question: "How many more birds are there than worms?". Another is called a "Won't Get" question: "Suppose the birds all race over and each one tries to get a worm. Will every bird get a worm? How many birds won't get a worm?". He devised a set of eight questions for each type. Children who correctly solve six or more were scored as giving correct response. The result was that only 17% of nursery school children, 25% of kindergartners, and 64% of first-graders gave correct responses to the "how many more ... than ... ?" question. But to the "how many won't get" question, 83% of nursery school children, 96% of kindergartners, and 100% of first-graders gave correct responses.

Hudson's intention was to study two alternative explanations about children's performance in determining numerical difference between disjoint sets. Children's poor performance on the "how many more ... than ... ?" question could be explained as hypothesized by Piaget (1965), namely that the children may be unable to establish suitable one-to-one correspondences between the given sets. Having observed children's successful performance on the "Won't Get" question, Hudson's conclusion was that the children did establish correspondence and were able to determine the numerical difference, but they did not do so because they misinterpreted the "how many more ... than ... ?" question. In addition, Hudson (1983) found that children's performance was also poor when the term "more" in the question was replaced by other comparative terms such as "taller", "longer", and "older". Thus, a linguistic factor, that is, children's limited comprehension of the comparative construction "how many ... [comparative term] ... than ...?", can account for young children's failure in finding the numerical difference between disjoint sets. This interpretation is consistent with much evidence indicating that the range of cognitive abilities elicited by cognitive-assessment tasks can be significantly affected by the language employed by those tasks (Donaldson, 1979; Gelman & Gallistel, 1978; Siegel, 1978). Although Hudson's original purpose was not to study word problem solving, researchers in this field have taken great interest in his findings.

Interpretations About Difficulty With Compare Problems

Cummins (1991) took Hudson's finding as an evidence against the RGH model which she labelled as a "logico-mathematical development view". The RGH model interprets many word problem solution failures as a lack of knowledge concerning set relations, particularly part-whole relations.

Cummins argued that the evidence suggests children often know more about logical set relations than the RGH model supposes. She cited Hudson's (1983) finding about the solution strategies children employed as an evidence of a tacit understanding of logical set relations. The most common strategy the children used involved counting the number of worms, counting out a subset of birds equal to the cardinality of worms, and returning the cardinality of the remaining subset of birds as the answer. Cummins believed this strategy implies a tacit understanding of one-to-one correspondence and subset equivalence (see Briars & Larkin, 1984) of sets with identical cardinalities, as well as the part-whole structure of the sets in question. Therefore, Cummins (1991) proposed a "linguistic development view" which suggests that a major source of difficulty children encounter when solving word problems is properly interpreting certain words and phrases in terms of sets and logical set relations.

Cummins' view is supported by the fact that children often transform comparative terms into simple possession terms when retelling word problems (Cummins et al., 1988), skip over comparative terms when reading phrases containing them (De Corte & Verschaffel, 1986), and perform better when problems containing comparative terms are reworded to exclude them (De Corte et al., 1985; Hudson, 1983). Besides comparative terms, Cummins' view also applies to other ambiguous terms such as "altogether", "each", and "some", etc., in all kinds of word problems.

Extending the work by Cummins et al. (1988), Cummins

(1991) asked children to solve Compare-4, Compare-6, Combine-5, and Change-6 problems and then to select pictures that represented the problems' structures. Her hypothesis in one of the experiments was that solution accuracy could be predicted from children's selection of pictures with correct problem representations, which shows successful understanding of the problematic terms. The results generally support this hypothesis. However, the variance in solution accuracy accounted for by the picture selection task was 43%, 16%, and 22% for Compare-4, Compare-6, and Combine-5 problems, respectively (Change 6 did not correlate with the picture selection task). Although the linguistic view appears to hold some explanatory power, more than half of the variance in each of these problems is yet to be explained (Okamoto, 1992).

Hudson's findings indeed show the RGH model's limitation, especially when it is applied to Compare problems. The model is quite successful in Change and Combine problems. It roughly hypothesizes that the development of conceptual understanding in solving Compare problems parallels that in solving Change and Combine problems. Riley et al. (1983) noticed that Compare-1 problems are usually quite difficult for kindergarten and first-grade children. The explanation the model gives is that the failure is associated with the lack of a schema for understanding the problem situation in a way that makes contact with the model's available action schemata, in this case the match action schema. This explanation attributes the failure to the lack of a specific problem schema -- Compare schema. Linguistic factors are excluded in the model. Hudson's (1980) finding is only taken as an example supporting their general hypothesis that the lack of conceptual understanding instead of solution actions is the major source of children's difficulty in word problem solving.

However, in a revision of the RGH model, Riley and Greeno (1988) admitted the role of a linguistic factor. They agreed that the failure of nearly all the kindergarten and first-grade children to reach Level-1 performance on Compare problems probably indicates that they lack the linguistic and conceptual knowledge to understand the language involving in "how many more ... than ...?". Young children are able to solve problems involving comparisons of sets when there is sufficient linguistic support for their understanding, but if the phrases "more ... than ..." and "less ... than ..." lack quantitative meanings for children, then their ability to infer the differences will not be used.

In the revision (Riley & Greeno, 1988), the model still cannot satisfactorily predict performance on Compare problems. They take the linguistic view as one possibility, indicating that children may need knowledge for understanding specific patterns of information involved in quantitative comparisons. However, they raise an alternative possibility which goes back to their mathematical interpretation, namely, understanding of comparisons depends on acquiring knowledge about differences as a relation between sets. In a part-whole schema, the union of two sets is still a set, but the difference between two sets is not a set. This means that the numbers involved in set unions can be understood as the cardinalities of individual sets, but the numbers involved in set differences must be understood as a relation between sets or between the cardinalities of two sets. What makes this difficult for children is that the "difference" describes a relation between quantities which is not a direct quantification of objects in the real world the way all sets are in Change and Combine problems (Resnick, 1989). When children encounter a phrase like "5 more than ...", they must understand this number as the value of an operator instead of the cardinality of a set. This understanding requires a more advanced understanding of numbers.

Another word problem solving model, Briars and Larkin's CHIPS (1984) includes Hudson's Won't Get problem type and deals with its difference from Compare problems, but takes a different approach. CHIPS manipulates physical or mental "chips" as the model's basic action. It also draws key terms and phrases as cues for its actions. According to CHIPS, the difficulty with Compare problems is that, in their usual wording, these problems describe no actions that the model or a child can imitate. Hudson's Won't Get wording changes a static comparison problem into an action-cued problem so that CHIPS can solve it easily. CHIPS first builds sets of chips representing the birds and worms separately. It then interprets the phrase "how many birds won't get ..." as a cue to match the birds one-by-one to the worms and to count the leftover birds. This is done by a production that makes a "match" schema and a count schema. The match schema holds the knowledge of what sets are to be matched. In this analysis, Carpenter and Moser's (1981) Equalize problems are also categorized as action-cued comparison problems.

When CHIPS encounters a Compare-1 problem, additional language capability is required. Based on the ability to make match schemata in response to action-cued comparison problems, a new production is needed to recognize the phrase "how many more ... than ..." as a cue to build appropriate match and count schemata. It is essential to know that the phrase means to match the two given sets and count what is left over.

CHIPS constructs a display of counters directly derived from certain terms in problem texts, rather than requiring intermediate representations of sets and set relations of the kind that are constructed in the RGH model. This is consistent with Longford's (1986) proposal about "thinking on the table" versus "thinking in the head", assuming that instead of forming a structural mental representation of problem information the child, at least in the easiest problem types, simply takes each piece of information as it comes in and represents it on the table with blocks. However, solving the "action-cued" comparison problems involves two action schemata "match" and "count" which are cued by term "get" and "how many ... won't", separately. What remains unexplained by the above proposal is how these two schemata become associated in solving the problems if there is nothing coordinating the two.

There is an agreement among researchers that Compare problems are difficult, and that certain rewordings can make them easier. However, the explanation of why they are difficult and what the effect of rewording is remains controversial. Each interpretation holds some explanatory power, but all of them have some aspects left unexplained.

Problems Involving Difference Finding

The interest of the present study is drawn from the issue with Compare problems. The present study includes three types of word problems: Compare-1 problems (Riley et al., 1983), Equalize-1 problems (Carpenter & Moser, 1981), and Won't Get problems (Hudson, 1980, 1983). They all involve finding the difference between two disjoint sets. The formal solution to all of them is "larger set, subtract smaller set, equals difference". In other words, they share the same mathematical content and structure. What makes them different to form three types is only the problem texts. The three types of problems, named COMPARE, EQUALIZE, and WON'T GET, respectively in this study, are listed in Table 2. After presenting the two given sets, (1) COMPARE problems ask how many more objects are in the larger set than in the smaller set; (2) EQUALIZE problems ask what to do to make the smaller set have as many objects as in the larger set; and (3) WON'T GET problems ask how many objects from the larger set won't get the objects of the smaller set.

Are the three types of problems as difficult as one

Table 2

Three Types of Problems Involving Difference Finding

COMPARE
Joe has 8 marbles.
Tom has 5 marbles.
How many more marbles does Joe have than Tom?
(Riley, Greeno, & Heller, 1983)
EQUALIZE
Joan picked *a* flowers.
Bill picked *c* flowers ($a < c$).
What could Joan do so she would have as many flowers as Bill? How many more would she need to pick?
(Carpenter, Hiebert, & Moser, 1981)
WON'T GET
There are 5 birds.
There are 3 worms.
Suppose the birds race over and each one tries to get a worm. How many birds won't get a worm?
(Hudson, 1983)

another to children at a certain grade level? If not, what makes one more difficult or easier than another? If it is problem text that makes the difference, what does it affect, and how does it have the effect? These questions are similar to those which were at issue in the studies mentioned above.

None of the previous studies could answer the questions satisfactorily. Mere mathematical reasoning ability or language understanding capability alone cannot account for the relatively complicated word problem solving processes. More detailed analysis of the relationship between language comprehension and mathematical reasoning is needed. Also, it is plausible to propose an internal representation to mediate problem text understanding and mathematical reasoning followed by problem solution in this process. The question remaining is how the representation is constructed, and how it mediates the components in the process.

Problem Representation and Competence

Dual Representation

Integrating both text comprehension and problem solving aspects of word problem solving, Kintsch and Greeno (1985) developed a model of understanding and solving word problems. This model includes a more thorough analysis of processes of text comprehension than the RGH model. The general theory of text comprehension used in this model is developed by Kintsch and van Dijk (1978), and van Dijk and Kintsch (1983). The problem solving theory used in this model includes Riley et al.'s (1983) assumptions about the semantic knowledge required for representing the problems and the processes of operating on the numbers to find the answers. The model is implemented in two computer programs, namely WORDPRO (Fletcher, 1985) and ARITHPRO (Dellarosa, 1986).

According to van Dijk and Kintsch (1983), memory representations of texts have two components, a propositional structure of information that is in the text in a specific sense, and a "situation model" that is derived from the text, wholly or in part. The propositional structure, or "text base," is obtained by constructing a coherent conceptual representation of the text itself, called a microstructure, and then deriving from the microstructure a hierarchical macrostructure that corresponds to the essential ideas expressed in the text. If the text is studied in its own right, the text base is adequate for comprehension. However, if the text is merely the medium by which information is transmitted, in other words, the reader's purpose is learning from reading the text, another component is needed for comprehension (Kintsch, 1986). This is called the situation model, because it includes inferences made by using knowledge about the domain of the text information. It is a representation of the content of a text, independent of how the text is formulated, but integrated with relevant experiences in this domain. Its structure is adapted to the demands of whatever tasks the reader expects to perform. When children read a word problem and try to solve it, they construct a representation including both components (Kintsch & Greeno, 1985).

Kintsch and Greeno's model includes a set of knowledge structures and a set of strategies for using these knowledge structures in building a representation and in solving the problem. The representation is a dual one: on one side, the text base represents the textual input, and on the other side, an abstract problem representation, called the "problem model" (instead of "situation model" in van Dijk and Kintsch's term), contains the problem-relevant information from the text base in a form suitable for calculational strategies to yield the problem solution.

Problem representations are built in several steps. The verbal input is transformed into a conceptual representation of its meaning, a list of propositions. The propositions are organized into a task-specific macrostucture that highlights the general concepts and relations mentioned in the text. This organized set of propositions is referred to as the text base. Coordinated with the representation of propositions is the problem model, which reflects knowledge of the information needed to solve the problem. In constructing the problem model, the reader infers information that is needed for solving the problem but is not included in the text base, and excludes information in the text base that is not required for solution of the problem.

The propositions in the text base hold four "slots": object, quantity, specification, and role. The first three can be filled directly by textual input, e.g. Tom owns 3 marbles. The role slot cannot be filled in text base because it can only inferred from mathematical knowledge. Then a set of strategies organizes the propositions into a coherent, task-specific text base, which takes the form of set and set relation. However, because the role slot is not filled yet, the set relation is still in question. The set and the set relation in question, then, constitute the content of the problem model. Now, mathematical knowledge or semantic knowledge referred to as "higher order schemata" are employed to fill the role slot in problem model. These schemata deal with set relations which are critical for deciding how to solve the problem. Five such schemata are proposed in the model. A TRANSFER schema assigns a "start set", "transfer set", or "result set" role to an appropriate set. Two variations of this schema are TRANSFER-IN and TRANSFER-OUT. A PART-WHOLE schema assigns "subset" role and "superset" role to sets. There are also MORE-THAN and LESS-THAN schemata to assign "largeset", "smallset", or "difference" to the empty role slots in a problem model. Once the role slots are filled by inferring the higher order schemata, the goal of the problem solution is determined. Then a set of calculational strategies or actions such as COUNT-ALL, ADD-ON, or SEPARATE-FROM are triggered to provide the answers to the problem.

This model proposes that the problem representation constructed by a child during a solution attempt is a joint product of his or her language comprehension and mathematical reasoning. The interaction between these two components is simulated by the processes of constructing the dual representation: from text base to problem model. This model is the first attempt to include both a linguistic factor and a mathematical factor in the integrated process of word problem solving. The dual nature of this model not only indicates the necessity of both components in the process but, more importantly, reveals the interdependency between the two components. On the one hand, the problem model uses information in text base as its content. Thus, the prerequisite for building a problem model is that children understand words such as "have", "give", "all", "more", and "less" in a general way, and also in a special, task-specific way. This means they have to represent propositions involving "having", "giving", and so on with arguments that refer to sets of objects. This also includes extensions of the ordinary use of terms such as "all" and "more" to more complicated constructions involving sets, denoted by "altogether" and "more than". Furthermore, the inference to mathematical schemata is cued by certain key propositions such as HAVE-ALTOGETHER in text base. The construction of the problem model is initiated by the text base.

On the other hand, textual cues do not directly lead to the solution operations. In the problem-solving process, information from text base is reconstructed in problem model. This is evident in a problem recall study reported by Kintsch (1986). Dellarosa, Weimer, and Kintsch (1985) presented word problems in both easy versions and hard versions on the basis of relative difficulty of the problems, and asked 30 second-graders to recall the problems in two conditions. One was to recall the problems without solving them, the other was to recall after solving the problems. When the subjects only had to recall the problems but not solve them, there was no significant difference between the recall of easy and hard problems, confirming that the text base of these problems were quite comparable. However, after solving the problems, the subjects recalled significantly more easy problems than hard problems. In other words, although the hard problems were not inherently less recallable than the easy problems, they were not recalled as well after a solution attempt had been made. The complexity of the problem model required for solving hard problems confused the subjects. There was a tendency to misrecall the problems as if they were easier ones. The subjects who misrecalled did not recall the text directly, but rather they recalled the problem model they had formed from the text base. After the mistakes in recall were analyzed (Kintsch, 1986), it was clear that the text bases of the problems were not used as the basis for the recall; rather, the texts were reconstructed from the problem model.

The distortion of the textual information in recall of word problems supports the existence of the "problem model" and the role of mathematical knowledge in the representation construction. This means that the textual information must be reconstructed in the problem model so that the information can be used for problem solution. The problem model is represented in a form of a mathematical structure which consists of relations among sets. Any linguistic input, if needed for problem solution, must be understood in its specific mathematical sense.

Kintsch and Greeno's (1985) model is apparently more sophisticated than previous word problem solving models. It in turn has become the basis for further studies in this area (Riley & Greeno, 1988; Cummins et al., 1988; Okamoto, 1992). Researchers from either the mathematical perspective or the liguistic perspective use the same model to simulate the problem solving processes. Both of them have found support for their own view of the competence underlying the dual representation.

Competence for Understanding Word Problems

In view of the dual representation assumption, any emphasis on one factor of understanding no longer means excluding the role of another factor. The linguistic point of view and the mathematical point of view both agree that the competence for understanding word problems consists of both linguistic and mathematical aspects. The argument now concerns along what line the competence is developed.

Cummins' (1991) linguistic development view suggested that the major source of children's difficulty with certain types of word problems is their misinterpretation of certain words and phrases in problem texts. This misinterpretation does not mean they do not understand the expressions at all, but there has been a failure in mapping the expressions onto appropriate mathematical structures. In other words, children who fail do not understand the words and phrases in terms of mathematical set relations. This in turn, leads to a failure to access the mathematical knowledge which has been available.

Cummins suggested that development of the competence depends on the acquisition of new meanings of the terms and phrases which are already understood in a general sense. Through instruction or experience, children can learn the mathematical meanings of the language and familiarize themselves with the mapping onto mathematical relations. This line of development starts from language with general meaning towards language understood in mathematical sense.

On the contrary, there is another developmental line proposed as starting from a more "pure" mathematical competence to linguistically related competence. Gelman and Greeno (1989) hypothesize three levels of understanding of numerals. At the simplest level, the meanings of numerals include reference only to individual objects and the results or arguments of counting operations. Implicitly, numerals are associated with sets in the process of counting, but this does not imply that the representations of meaning include explicit references to sets. At a second level, numerals denote the cardinalities of sets, and reference to sets is included in the meaning of propositions that have numerals and other quantifiers, such as "some". For example, when a child hears or reads a phrase "three marbles," he or she understands that there is a set of marbles and that "three" denotes the cardinality of the set. The competence for understanding propositions including reference to sets is referred to as a "linguistic cardinality principle". At a third level, numerals also denote the numerical differences between sets. At this level, the meaning of a sentence such as "Tom has two more marbles than Joe" includes reference to the set of Tom's marbles, the set of Joe's marbles, and a third entity, the numerical difference between the two sets. In this usage, numbers are properties of a relation between sets. This concept of number therefore is more complex than that in which numbers are only cardinalities of individual sets. This competence for understanding propositions that include reference to set differences is called by Gelman and Greeno the principle of "linguistic numerical difference".

According to Gelman and Greeno (1989), the development of competence for understanding numbers starts from a nonlinguistic principle of cardinality and proceeds to the two linguistic meanings. Young children have some counting-specific competence at the beginning, then the learning of linguistically related principles can be based on this competence. The process of learning such new principles happens as part of their instruction in arithmetic, where they learn to add and subtract. Some data reported in Riley at al.(1983) is cited to support this point, specifically that Compare-1 and -2 problems were solved correctly by 80% of the children who were near the end of second grade, but by only 25% of the children who were near the end of first grade.

The linguistic development view proposes a developmental sequence which gradually increases mathematical components in language understanding (Cummins, 1991). Gelman & Greeno (1989) proposed a developmental sequence towards a "linguistic numerical difference" principle which includes linguistic components in a mathematical structure. It is not hard to find that these two lines of competence development start from contrasting originss but come to the same end, only expressed differently, either understanding language in terms of mathematical relations or understanding mathematical relations in linguistic forms. In other words, they both point out the mapping between the two components. This, indeed, is the key to the understanding of word problems.

Coming to this point does not mean they have answered all the questions to be answered, nor does it mean the issue with Compare problems or with difference-finding problems ceases to exist. The first remaining question involves the development of the competence. They both attribute it to instructions and experience, supposing children can learn as long as they are exposed to the problems. But, is there any initiation device or constraint on such learning? Why have second-graders learned but first-graders not learned the "numerical difference" concept? Is this only a matter of time, exposure, or familiarization?

Another problem with both sides is that they each have pointed out the mapping, but neither of them has described the process of mapping. This process is implied as naturally happening because children's competence has achieved the understanding of mathematical meaning in a language form. But if all the competence is term-specific or form-specific, how many different kinds of competence must be hypothesized if children repeatedly encounter different types of difference-finding problems?

Their failure in describing the mapping process, especially in solving Compare problems, also reflects the researchers' uncertainty as to what mathematical structures the linguistic input can map onto. Cummins (1991) proposed only part-whole relations to be mapped onto. She interpreted the matching strategy used by Hudson's (1983) subjects as a tacit understanding of part-whole relations. Cummins' proposal is based on Riley et al.'s (1983) hypotheses. However, Riley and Greeno (1988) have abandoned using part-whole schema as the mathematical set relation inferred to construct problem models for Compare problems, because the difference between sets cannot be treated as a set as in part-whole relations. Gelman and Greeno's (1989) hypotheses about the "linguistic numerical difference" is an effort to re-explain the process of solving Compare problems. In addition, mapping from Compare problem texts to the part-whole schema requires a much harder transformation from the problem situation to the structure, and the processing demand would be heavily increased if mapping to part-whole schema were not automatized. However, they have not proposed any other problem model construction to replace part-whole structure.

Further explanation about the underlying competence and its development is needed, as is a basis for simulating the mapping processes.

Central Numerical Structure in Word Problem Solving

The word problem solving models mentioned above cannot go beyond their own limitations when they attempt to explain young children's difficulty with Compare problems and the relative difficulty of problems involving difference finding. Their domain-specific perspectives prevent them from looking at children's more generalized potential and limitations in the development of the competencies underlying word problem understanding. In turn, they fail to find a cognitive structure base on which they can simulate the process of mapping between linguistic input and mathematical reasoning. This difficulty is overcome by a new word problem solving model developed by Okamoto (1992) who integrates a generalized developmental theory (Case, 1992) into the study of word problem solving. The new model suggests that a set of "central numerical structures" proposed by Case (1992) underlies the understanding of word problems. This conception provides a plausible base for simulating the mapping between problem text and mathematical reasoning. Okamoto's model is also to be seen as an application of Case's general theory of cognitive development in a new domain. Although the issues regarding specificity and generality in cognitive development are not the concern of the present study, this developmental theory does produce a framework for the present study that is not possible from the previous theories.

Case's General Theory of Cognitive Development

In Case's (1985) theory of cognitive development, the basic constructs are control structures, stages, and substage. The control structures are what children construct to cope with problems in their daily life, and contain three components: a representation of the essential features of some particular class of problem, a representation of the goals that this problem class most frequently occasions, and a representation of a sequence of operations that will bridge the gap between the problem's initial and terminal states. According to this theory, much of children's development stems from a change in their control structures. This change is hypothesized to be constrained and potentiated by a set of changes that are system-wide and that have a strong biological component. These changes influence the highest level of intellectual operation that children can execute successfully under optimal environmental conditions, as well as their working memory for the products of such operations. As these upper limits change, the control structures are believed to progress qualitatively through a universal sequence of four stages that are labelled as the sensorimotor stage, interrelational stage, dimensional stage, and vectorial stage. Within each of the four stages, a recursive sequence of structural changes is proposed: at the first substage, children assemble a new class of operations, by coordinating two well-established executive structures that are already in their repertoire. As their working memory increases (as a function of maturation and practice), they enter a second substage in which they are capable of executing two such operations in sequence. Finally, with further growth in working memory, they enter a third substage in which they become capable of executing two or more operations of the new sort in parallel, and integrating the products of these operations into a coherent system. Once consolidated, these integrated systems then function as the basic units from which the structures of the next stage are assembled, and the process of integration repeats.

Central Conceptual Structures

In Case's (1985) earlier work, a control structure is defined as a tripartite entity consisting of three representations regarding a particular class of problem. Although the assembly of any given control structure and the construction of any new control structure are constrained and potentiated by some system-wide components, and the development of control structures occurs through a universal sequence of four stages, the control structures, by definition, are task-specific. The formation of a control structure is a function of the particular question which has been posed to the child. Thus, the control structures within any one phase of development are constructed in isolation of each other, in an independent, unrelated fashion.

More recently, a series of studies by Case and his associates (Case, 1992) resulted in an extension of Case's general theory of cognitive development. It was found that "it is a mistake to see children as assembling executive control structures for each separate task in complete isolation from those for each other task, subject only to an upper bound on their processing capacity. Rather, it seemed more appropriate to view children as assembling a central conceptual structure that is applicable to a broad range of tasks, then utilizing this central structure, more or less successfully, as a guide for assembling the particular executive control structures that each new task may require" (Case, 1992, p. 355).

By postulating a "central conceptual structure", the hypothesized locus of generality in children's performance is shifted from the emphasis on the size of their working memory to including conceptual structure assembled under the constraints of the working memory. This central structure exists prior to a problem being posed, thus, the control structures can no longer be seen in an independent, unrelated fashion. Rather, they must been seen as being part of a network of related control structures, which are tied together by a common conceptual core.

A central conceptual structure (Case & Griffin, 1990; Case & Sandieson, 1992) is an internal network of concepts and conceptual relations that plays a central role in permitting children to think about a wide range of situations at a new epistemic level and to develop a new set of control structures for dealing with them. By a "structure," the notion means an internal mental entity that consists of a number of nodes and the relations among them. By "conceptual" it means that the nodes and relations are semantic rather than syntactic. And by "central," it means structures that (a) form the core of a wide range of more specific concepts, and (b) play a pivotal role in enabling the child to make the transition to a new type of thought, where these concepts are of central importance.

Central conceptual structures are found in several domains of child development, such as logico-mathematical thought, social and emotional thought, and spatial thought, and in motor development. It is also found that these central conceptual structures bear a certain resemblance to each other, both in their form and in the timing of their emergence. These commonalities suggest that all the structures may be subject to a common set of constraints, in speed of processing or in working memory (Case, 1992).

Central Numerical Structures

In its application to the domain of mathematics, the general developmental theory identifies a set of central conceptual structures regarding quantitative variables. Specifically, children's development of number concepts from approximately four to ten years of age is explained in terms of the acquisition of central numerical structures as shown in Figure 1 (Case & Griffin, 1989). Based on the results of Resnick's (1983) and Fuson's (1982) studies, Case and Sowder (1990) proposed the "mental number line" by which the central numerical structure is represented.

Children at the age of four are capable of counting

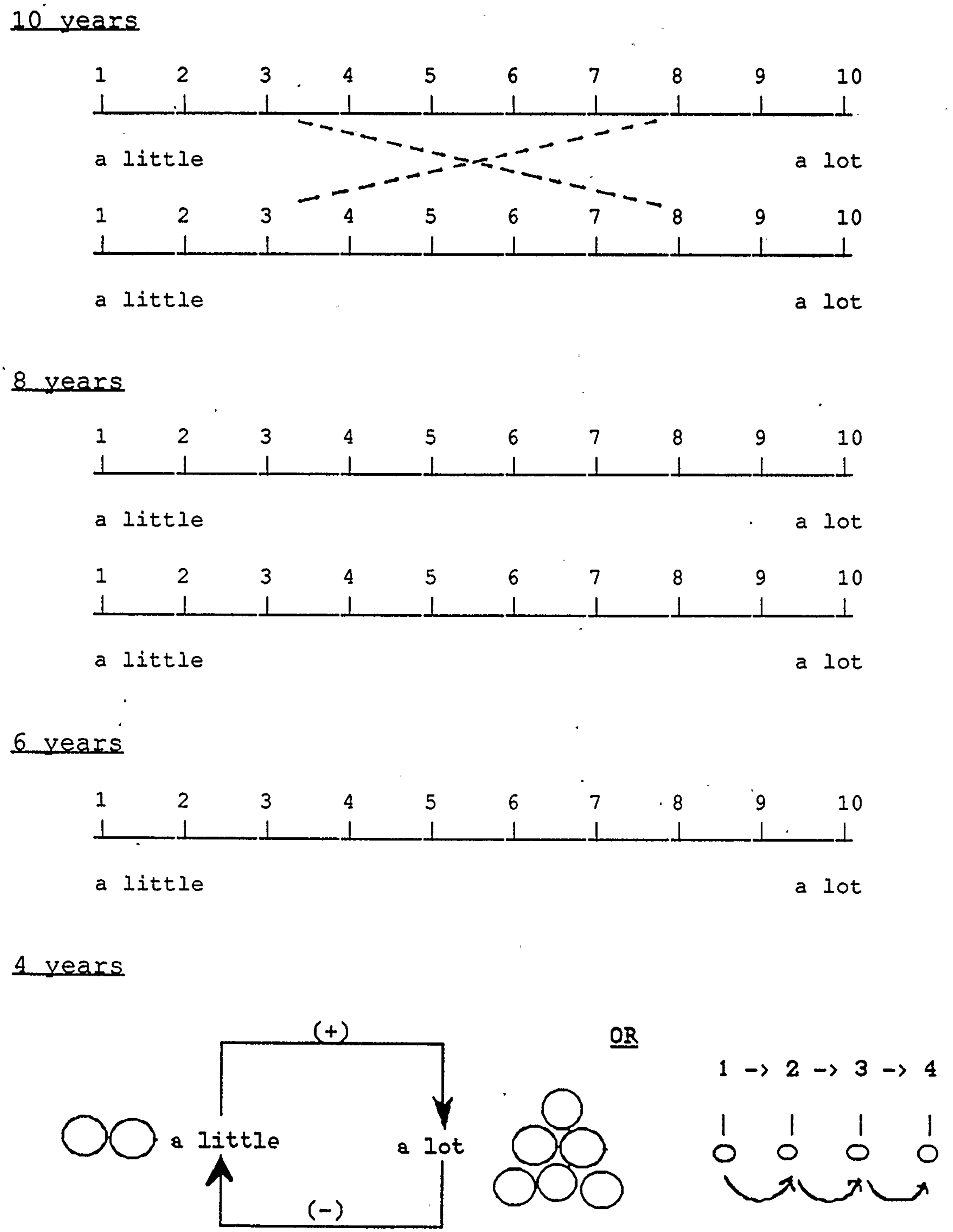

Figure 1. Development of Central Numerical Structures (4 to 10 years of age) (From Case & Griffin, 1989)

(structure for enumeration) or making judgment of relative quantity (structure for quantity evaluation), but are not able to integrate these two structures. At the age of six, there is a qualitative shift in thought as children enter the first substage of the dimensional stage. They understand the relation between the two structures and are able to integrate them and, thus, think in terms of a single number line. At the second substage, 8-year-olds are able to coordinate two unit of operations, that is, to think in terms of two number lines, at least in a sequential fashion. A full capacity to coordinate the two number lines develops at the last substage where 10-year-olds are able to think in terms of two mental number lines in an integrated or on-line fashion, and appreciate the relationship between them. Finally, the integrated structure will function as the initial single unit to be reorganized into the structures of the next stage.

These central numerical structures represented as number lines have been identified in children's development of, for example, computational estimation (Case & Sowder, 1990), scientific reasoning (Marini, 1992), everyday mathematical knowledge including time-telling and money-handling (Griffin, Case, & Sandieson, 1992), and sight-reading of musical notation (Capodilupo, 1992). These cross-task studies have shown the existence of such central numerical structures and their roles in the development in the domain of quantity, which have verified Case's general description of cognitive development, as well as proved the applicability of the construct to a broad range of tasks, also showing a potential to apply to the domain of word problem solving.

Okamoto's Model

Okamoto (1992) applied the notion of central numerical structures to the development of children's word problem solving ability. By integrating the models of word problem solving regarding numerical competence by Riley and Greeno (1988) and by Gelman and Greeno (1989), models regarding text processing by van Dijk and Kintsch (1983) and by Kintsch and Greeno (1985), and the developmental theory by Case (1985, 1992) as it has been applied to the domain of number (Case & Griffin, 1989; Case & Sandieson, 1988; Case & Sowder, 1990), Okamoto assumed that the word problem solving process involves (a) the construction of semantic networks, as a result of comprehending problem texts, and (b) the construction of problem models using the various number lines that are available at different levels of development.

Okamoto (1992) developed a set of computational models to simulate children's solutions of the set of word problems categorized by Heller and Greeno (1978) (Table 1) at each of three developmental levels of knowledge. Each computational model performs three functions. First, it constructs semantic networks representing propositions extracted from the problem text, in Kintsch and Greeno's (1985) terms, forming the text bases. Second, it constructs problem models which are simulated in a form of arrays which increase in complexity from a mental object line to two coordinated mental number lines as a function of development. Prototypes of the problem models at different levels of development are shown in Figure 2. A choice for one form over the others depends upon linguistic information represented in the text base and the levels of central numerical structures assumed to be available. Third, it generates answers. The specific characteristics of the problem model are checked against a set of production rules. When appropriate conditions are met, manipulations of number on the mental arrays take place, simulating children's counting behavior. Based on the three functions, three levels of processes are simulated by applying the three different cognitive models in Figure 2 to each category of word problems.

Level-1 processes are limited to the lowest level of dimensional thought (unidimensional thought). At this level, objects described in a problem can be represented and manipulated mentally only on a single dimension. Specifically, Level-1 processes can line up objects internally on a single mental object line, add to or take away mental objects, count those objects that are present on a mental object line, and cite as an answer the last mental objects counted. Therefore, Change-1, Change-2, and Combine-1 problems can be solved at this level. When solving a Combine-1 problem ("Joe has 2 marbles. Tom has 6 marbles. How many marbles do they have altogether?"), for example, Level-1 processes construct a problem model representing 2 marbles possessed by Joe and 6 marbles owned by Tom internally on one mental object line with the owner-specification "Joe and Tom". Then the answer is obtained by

<u>Level 3</u>: Two mental number lines, well coordinated

Owner 1 1 <-> 2 <-> 3 <-> 4 <-> 5 <-> 6 <-> 7 <-> 8 <-> 9 <-> 10

Owner 2 1 <-> 2 <-> 3 <-> 4 <-> 5 <-> 6 <-> 7 <-> 8 <-> 9 <-> 10

<u>Level 2</u>: Two mental number lines

Owner 1 1 <-> 2 <-> 3 <-> 4 <-> 5 <-> 6 <-> 7 <-> 8 <-> 9 <-> 10

Owner 2 1 <-> 2 <-> 3 <-> 4 <-> 5 <-> 6 <-> 7 <-> 8 <-> 9 <-> 10

<u>OR</u>

Owner 1 <-> 2 <-> 3 <-> 4 <-> 5 <-> 6 <-> 7 <-> 8 <-> 9 <-> 10

Counter 0 <-> 1 <-> 2 <-> 3 <-> 4

<u>Level 1</u>: Single mental object line

1 2 3 4 5 6 7 8 9 10

o <-> o <-> o <-> o <-> o <-> o <-> o <-> o <-> o <-> o

a little a lot

<u>Notes</u>. <-> indicates a "next-to" relation.

⟷ can point to any number.

<u>Figure 2</u>. Prototypical Problem Models at Three Levels of Development (From Okamoto, 1992)

counting all eight objects on this mental object line.

Level-2 processes include a simple coordination of two mental number lines (bidimensional thought). That is, the unknown quantity is represented by coordinating two known quantities (each of which is represented on a mental number line). The relational information concerning the known and unknown quantities can be used to specify how one mental number line is to be manipulated with regard to the second. One of the simplest of such operations is "reflection", in which the quantity derived on one line is marked off or "reflected" on the other. Thus, Level-2 processes are able to solve Combine-2, Change-2 to Change-6, Compare-1 and Compare-2 problems. For example, a Combine-2 problem ("Joe has 2 marbles. Tom has some marbles. They have 6 marbles altogether. How many marbles does Tom have?") is solved in the following manner. A problem model is constructed consisting of two mental number lines, one representing Joe's 2 marbles, another representing the 6 marbles with the owner-specification "Joe and Tom". To answer "how many marbles does Tom have?", these two mental number lines are coordinated, that is, Joe's number line is reflected on the second line (whole) to indicated the part owned by Joe, and therefore the rest is inferred as the part owned by Tom. Finally, a counting procedure is carried out mentally to produce the answer.

Level-3 processes can fully coordinate two mental number lines (integrated bidimensional thought), that is, reverse the reflective operations executed at Level-2. Level-3 processes understand that (1) a numerical difference can represent the amount above (more) or below (less) a criterion, and (2) a positive difference on one line can be compensated for by a negative difference on the other line. Compare-5 and Compare-6 problems can be successfully solved at this level. A Compare-5 problem ("Joe has 6 marbles. Tom has some marbles. Joe has 2 more marbles than Tom. How many marbles does Tom have?") can be used as an illustration. The problem model is constructed containing one mental number line representing Joe's 6 marbles, another line representing "some" marbles owned by Tom. The way to coordinate the two lines is that "2 more" on Joe's line is reflected in reverse as "2 more to reach the number of marbles Joe has" on Tom's line, by which Tom's line has "2 fewer than Joe's 6". Then the answer is produced by counting mentally.

Okamoto's model was examined through a series of empirical studies involving word problem solving by school children at ages from six to ten. The results showed that there was a reasonable fit of the data to the computational model. That is, the difficulty levels of the problems identified by the children's actual performance were quite similar to those predicted by the model regarding the central numerical structures children possess at different development levels. This model consists of same semantic networks for building text base as in Riley and Greeno's (1988) and Kintsch and Greeno's (1985) models. What makes Okamoto's model different from the previous models is its assumption of problem model construction by applying Case's central numerical structures to this domain, so that her model can accurately simulate the entire solution process, especially the process of problem model construction. Furthermore, Okamoto's model successfully predicted the difficulty levels for Compare problems which was not explained by Riley and Greeno's (1988) model.

Number Line Coordination in Solving Compare Problem

The most important contribution of Okamoto's (1992) model to the analyses of problem solving processes for Compare problems is that it describes the problem model as represented by two number lines, and in solving a problem, the number lines are coordinated by a reflective operation. This is the most plausible assumption because of the nature of the problem itself. A Compare problem involves the difference between two disjoint sets. To find the difference, the most plausible way is to act on one set while constraining the action by referring to the criterion represented by the other set. It is less plausible to use a part-whole schema as the mathematical set relation inferred to construct problem models for Compare problems, because the difference between sets cannot be treated as a set as in part-whole relations. In addition, mapping between Compare problem texts to a part-whole schema requires a much harder transformation from the problem situation to the structure, and the processing demand would be heavily increased if mapping to part-whole schema were not automatized. Okamoto's model accurately describes the construction of the problem model as represented by number lines and the operations leading to the solution to the problem.

As an illustration, consider a Compare-1 problem: "Joe has 2 marbles. Tom has 6 marbles. How many more marbles does Tom have than Joe?". This problem can be solved by Level 2 processes, which means the solver must be around eight years old, at the level of bidimensional thought in development. According to Okamoto's model (Figure 3), initially, one number line represents Joe's 2 marbles; the second number line represents Tom's 6 marbles. Then, Joe's 2 marbles are reflected onto Tom's number line to indicate the difference between Joe's 2 marbles and Tom's 6 marbles. The difference is inferred as the amount that Tom has in excess of Joe's. Finally, one number line acts as a counter to count the difference marked off on the other number line to produce an answer.

This model is based on children's capability of representing two mental number lines. Coordination of the two number lines by a reflective operation is the key to solving the problem. Although, at Okamoto's Level-2 processes, the coordination is more or less sequential, rather than performed in an on-line, integrated fashion, it is adequate for solving this particular type of problem. This combination of plausibility and accuracy have never been achieved by any of the previous models. Furthermore, Okamoto's empirical studies (1992) showed that children's performance on this problem is well predicted by her model.

Problem Model for Compare-1

```
                          P1

Array (Joe)   1 <-> 2 <---------------------->

Array (Tom)   1 <-> 2 <-> 3 <-> 4 <-> 5 <-> 6

                                            P2
```

Compare-1

Joe has 2 marbles.

Tom has 6 marbles.

How many more marbles does Tom have than Joe?

Notes. P1: Proposition 1 [Owner-specification = Joe, quantity = 2, objects = marbles]. P2: Proposition 2 [Owner-specification = Tom, quantity = 6, objects = marbles]

Figure 3. Problem Model for a Compare-1 Problem (from Okamoto, 1992)

In the RGH model (1983) and Riley and Greeno's (1988) model, Compare-1 problems are supposed to be solved at Level-1, where children are able to make individual sets by manipulating external display of objects. No description of the solution process is explicitly given. It seems that the model suggests children make the two given sets externally, then if they have a "compare schema" available, children will be able to understand "how many more ... than ...?" question and relate it the an "action schema" called "match". However, Riley and Greeno (1988) find that the data pattern in their study "casts considerable suspicion on the models' characterization of knowledge for Compare problems. Although the kinds of knowledge assumed in Level-1 and Level-2 for Compare problems are similar to the kinds of knowledge for those levels for Combine and Change problems, many more kindergarten and first-grade children were at Level-1 or Level-2 for Combine and Change problems than for Compare problems. The scalability analyses and ordinal analysis were less successful for Compare problems than they were for Combine and Change problems. In the proportions of children matching model performance, statistical agreement was good for kindergartners and first-graders, but only because most children solved none of the problems. ... These different levels of knowledge distinguish among kindergarten and first-grade children's performance on the Combine and Change problems, but few of those children had response patterns on Compare problems that were consistent with any of the models except the null model that predicts no success." (p. 84)

In Briars and Larkin's CHIPS (1984) model, no intermediate internal representations are hypothesized for solving Compare 1 problems as well as Won't Get and Equalize problems. Linguistic inputs are directly mapped to two solution actions, "match" and "count". This pattern is probably a result of children's having well established the coordination between two number lines and having been extensively exposed to these kinds of problems. Otherwise it is hard to explain why the two actions can be used together for one of these problems. This can be done only when the problem solver uses one of the two given sets as the criterion for matching and counting the rest of the objects after matching, which requires the coordination proposed by Okamoto's model.

In terms of the present study, Okamoto's model not only successfully describes the entire process for solving Compare 1 problems, but also can be applied to all the problems involving difference finding between disjoint sets. It's assumption about number line coordination as the key to the problem solution also provides a basis for simulating the mapping between problem texts and mathematical operations in the processes.

The Effect of Concrete Materials

In previous studies, it is hard to find any discussion about the effect of concrete materials on the solution processes of Compare problems. It is assumed that at lower levels or young ages children are limited to direct modeling of problem situations and to solutions using concrete objects (Carpenter, 1985; Riley et al., 1983; Briars & Larkin, 1984). As a result, researchers consistently use certain kinds of concrete materials when they present Compare-1 problems to young children. Riley (1981) presented blocks with Compare-1 problems to kindergartners and first-graders and found 17% of the former and 28% of the latter correctly solved the problems. Carpenter, Hiebert, and Moser (1981) made a set of red and white Unifix cubes available to their subjects and told them to use the cubes to help them solve the problem if they needed the cubes or not sure of their answers. Their results showed 29 out of 43 first-grade students correctly solved Compare-1 problems. Hudson (1983) used a set of drawings of the objects described in his Compare-1 and Won't Get problems. The proportions of correct Compare-1 responses were 17% for nursery school children, 25% for kindergartners, and 64% for first-graders. The concrete materials were included in these studies simply because they were believed to be needed by the young children and would have no effect on the problem difficulty. Hudson (1983) mentioned that children responded incorrectly to "how many more ... than ...?" questions even when the given sets were block rows placed side by side so that appropriate one-to-one correspondence were visually understood, suggesting that the difficulty from the problem text could not be overcome just by using blocks.

However, there is some evidence for a facilitation effect of concrete materials on word problem solving other than Compare problems. Riley (1981) found there was a general improvement in kindergartners' performance when they used objects to solve all types of Change problems except Change 5 and 6. Steff and Johnson (1971) obtained a similar result for four types of Change problems and Combine-1 and -2 problems. In addition, Carpenter et al. (1981) observed that first-graders preferred to use blocks if they had the choice.

The concrete materials used for word problem solving can be included in the concept of "task environment" defined as comprising all the elements of a task that are available and perceived by the problem solver (i.e., the "givens" of a problem) (Resnick & Ford, 1981). Generally, task environment provides the raw materials out of which the information-processing system builds a representation of the problem. This in turn determines which solution strategy is selected. Information about how concrete materials facilitate solution correctness does not show their effect on the problem representation. This effect is assumed to be reflected in the strategies the children use under different conditions.

Carpenter et al. (1981) found an effect for the availability of Unifix cubes on some strategies used by first-graders in solving arithmetic addition and subtraction word problems. In addition problems, children were able to use a counting-on strategy, but the availability of cubes influenced children to use a counting-all rather than a counting-on strategy. In subtraction problems, a matching strategy was only feasible when objects were available. These findings support the position that the availability of concrete materials as a part of the task environment influences the construction of problem representations.

In previous studies, considerable attention has been given to determining how problem semantic structures influence representations, how linguistic factors influence representations, even how children use concrete materials to model the problem structures. However, there has seldom been research on how concrete materials influence representations, especially with Compare problems. Thus some questions remain: do children use a matching strategy in solving Compare problems? Do they use it only when concrete materials are available? Is there any influence of concrete materials on their Compare problem representation construction?

CHAPTER III

RATIONALE AND HYPOTHESES

Rationale

The purpose of the present study was to investigate the mapping between text comprehension and mathematical reasoning in the processes of solving difference-finding word problems by first-graders. The theoretical framework regarding word problem solving process adopted by the present study includes these assumptions: (1) successful solution to a problem relies on the conceptual understanding of the problem, that is, building a coherent mental conceptual representation of the problem; (2) the representation is a dual one including a text base consisting of information given in the problem text, and a problem model in which information from the text base is reconstructed in terms of a mathematical structure acquired by the problem solver through development; (3) the mapping between text comprehension and mathematical reasoning occurs when information from the text base is being reconstructed in the problem model; (4) the problem model is represented as two mental number lines which are coordinated to produce an answer to the problem; and (5) the problem solving strategy is determined by the problem model, and the observed strategy can be used as an indicator of what kind of problem model is being constructed. Based on this framework, several inferences on the processes of solving difference-finding problems and on the mapping between text comprehension and mathematical reasoning in the processes can be made.

Mapping Onto the Coordination of Two Number Lines

First, to solve the three types of difference-finding problems, namely COMPARE, EQUALIZE, and WON'T GET (Table 2) requires the same mathematical reasoning competence, that is, the coordination of two number lines. The purpose of all the three types is to find the numerical difference between two disjoint sets. The two sets are given. They are represented in each problem model as two number lines. Then in the case of COMPARE, the question is "how many more objects are in the larger set than in the smaller set". According to Okamoto (1992), two sets of objects cannot be lined up on a single mental object line to capture the comparative nature of the problem. Coordinating two number lines is a prerequisite of comparing two disjoint sets. In the case of EQUALIZE, the question is "what can be done to make the smaller set have as many objects as in the larger set". It does not matter which number line the problem solver works on, either to make the smaller set larger or to make the larger set smaller, but it is necessary to refer to the criterion about how far to go indicated by the other line. In the case of WON'T GET, the question is "how many objects from the larger set won't get the objects of the smaller set". A matching strategy is proposed in almost all previous studies to solve the problem. Matching actually means making correspondence between the two number lines. Furthermore, in order to reach a solution, matching must be followed by a counting-the-rest action. This action is constrained by the criterion about where to start the counting. This starting point is marked by the smaller set line when it is matched onto the larger set line. Therefore, coordination between the two number lines, as in the cases of COMPARE and EQUALIZE, is required in the case of WON'T GET.

Second, successful solution for any of the three types of problems depends upon if the problem texts can be mapped onto the coordination of the two number lines. This is the reason for any difference in difficulty among the three types. In the EQUALIZE problem text, the question sentence contains the phrase "as many ... as ...". It cues the problem solver to take both sets into consideration, and to use one set as criterion to constrain the action on the other set. If the problem solver possesses the competence to coordinate two number lines, the EQUALIZE problem text can be easily mapped onto this structure to build an appropriate problem model, and then the coordinating operation can be applied to produce an answer. In the WON'T GET problem text, the phrase "each one tries to get a ..." cues the problem solver to build the one-to-one correspondence between the two sets, in other words, to project one number line on the other in the problem model. Then the phrase "won't get" cues the counting of the unmatched objects. However, this counting action is not independent of the coordination of the two number lines. It uses the mark made by the end of the smaller set line as the criterion to start counting.

Although the whole process may occur in sequence, that is, taking the result of projection as a "given", and then counting the unmatched, instead of in a "on-line" fashion, the problem text is clearly mapped onto the coordination of the two number lines. In the COMPARE problem text, however, no cues are available. The phrase "how many more ... than ..." does not cue to build any operational relationship between the two number lines. Even if the problem solver possesses the competence of coordination of two number lines, he or she may not be able to apply the structure in this COMPARE problem context. In summary, unlike EQUALIZE and WON'T GET problem texts, the COMPARE problem text does not facilitate the mapping onto the coordination of two number lines.

Third, whether the problem text is mapped onto a mathematical structure to build a problem model can be reflected by the strategy the problem solver uses to solve the problem. For the difference-finding problems, three strategies are proposed, namely PART-WHOLE (Riley et al., 1983; Cummins, 1991), ADD-ON (Carpenter et al., 1981), and MATCH (Hudson, 1983; Carpenter et al., 1981; Briars & Larkin, 1984). The PART-WHOLE strategy treats the larger set as whole, the smaller set as one part, and the difference as the other part. The interrelationship in this structure is "whole = part + part". The strategy then takes away the given part from the whole and counts the objects in part left as the difference. As mentioned in a previous section (see Chapter 2), mapping onto a PART-WHOLE structure from the difference-finding situation requires a harder transformation, so that the competence of simple coordination of two number lines does not easily produces a PART-WHOLE strategy. The ADD-ON strategy adds more objects to the smaller set to make it equal to the larger set, then takes the added objects as the difference. This strategy is clearly based on the coordination between two number lines. It acts on the smaller set line by adding more objects on the line, but constrains the action by referring to the criterion marked by the larger set line, that is, continuing to add until the smaller set line reaches the end of the larger set line. The MATCH strategy matches up the objects of the two given sets and find the unmatched objects as the difference. This strategy also reflects the coordination between the two given sets. The matching action reflects the projection of one number line onto the other, and the counting-the-left action uses the mark made by the end of the smaller set line as the starting point to count. Thus, if either the ADD-ON or MATCH strategy is observed, it can be inferred that the problem model is being represented as the coordination between the two number lines. As for the PART-WHOLE strategy, it does not reflect a simple coordination between two sets, it shows that a more demanding transformation is being used and a more sophisticated representation is being built.

In summary, all the three types of difference-finding problems require the ability to coordinat two number lines; certain types, like EQUALIZE and WON'T GET facilitate the mapping from their problem texts onto the coordination of two number lines, but certain types, like COMPARE, do not. If the mapping is successful, the problem model represented as the coordination of two number lines can be constructed, and, the successful problem model construction can be inferred by the presence of certain problem solving strategies, like ADD-ON and MATCH.

Developmental Level

The COMPARE problems are reported to be difficult for young children up to first-grade students (Riley, 1981; Hudson, 1980, 1983). This is well explained by Okamoto's (1992) model which is based on Case's (1992) theory of the development of central numerical structures (see Chapter 2). Across a wide range of tasks, children show a similar pattern of development. At the age of four, prior to the dimensional reasoning, children count and make judgments of relative quantity, but are not able to integrate the two structures. At about six years of age, children can understand the relation between enumeration and quantity evaluation. This relation allow them to think in terms of a single mental object line or along one dimension. This structure still does not allow children to deal with the word problems involving comparison between two disjoint sets. It is at about age eight that children enter the bidimensional substage and are able to construct two number lines to coordinate them. Not until this level do children possess the competence to solve the problems involving comparison.

Children at age of seven would be the best candidates to reveal the differences of the three types of difference-finding word problems in mapping onto the coordination of two number lines. Generally speaking, seven-year-old children's processing capacity has allowed them to begin to build two number lines and to coordinate the two lines, but this newly developed structure has not yet consolidated. Their ability to construct such a structure in a problem solving situation is task-dependent, that is, to represent the problem in terms of such a structure heavily depends on the factors such as problem formation, familiarity, and task environment. In the context of solving difference-finding word problems, this task-dependency would be mainly reflected by the fact that the problem model construction in terms of the coordination of two number lines relies on the problem text. It is reasonable to assume that if the text cues the coordination, the newly developed structure can be applied to the problem; if not, the structure is not accessible.

Using Blocks

Based on Resnick and Ford's (1981) argument that the task environment influences problem representations, and Carpenter et al.'s (1981) findings that the availability of concrete materials influences strategy use by first-graders in solving word problems, it can be inferred that the presence of certain concrete materials affects the construction of representations of difference-finding problems. Specifically, if the COMPARE problem texts do not cue the mapping between the two number lines representing the two given sets, the availability of certain objects like blocks may provide a perceptual cue for the coordination. That is, when children have two piles of blocks to represent the two given sets, and they arrange the two sets of blocks in two rows side by side, they may see that the row of blocks representing the larger set exceeds the row of blocks representing the smaller set. Then when they hear the question "how many more ... than ...?", the removability of the blocks may cue them to try to manipulate the two block lines to build a one-to-one correspondence between them, and then they see the unmatched blocks, and count them to produce the difference. Thus, it can be assumed that the perceptual cues may compensate for the lack of textual cues in building a appropriate problem model.

According to this assumption, the availability of blocks would reduce the difficulty of COMPARE problems. However, an alternative is also plausible. The above manipulation and arrangement with the blocks may be merely an external representation of the problem solver's mental number lines. This reduces the problem solver's processing load and this then increases correct solutions. This means the availability of blocks does not influence the construction of the problem model, only re-represents the problem model by means of the concrete materials.

The hypotheses of the present study were derived from the main theoretical framework provided by the previous studies, and from the inferences made on the mapping of texts onto the coordination of two number lines, the development level, and the effect of using blocks.

Hypotheses

The general hypothesis of the present study was that, among the three types of simple arithmetic word problems involving finding the numerical difference between two disjoint sets, the EQUALIZE and WON'T GET problem texts facilitate their mapping onto the problem models represented as the coordination of two mental number lines, whereas the COMPARE problem texts do not facilitate the mapping. This general hypothesis yields the following specific predictions.

1. First-grade students will perform better in terms of correct solutions in solving EQUALIZE and WON'T GET problems than in solving COMPARE problems.

2. First-grade students are more likely to use ADD-ON and MATCH strategies to solve EQUALIZE and WON'T GET problems and they are less likely to use the two strategies to solve COMPARE problems.

3. First-grade students will perform better in terms of correct solutions in solving all the three types of problems when the problems are presented with blocks than without blocks.

4. When the COMPARE problems are presented with blocks, first-grade students are more likely to use ADD-ON and/or MATCH strategies to solve the problems.

CHAPTER IV

METHOD

Sample

Twenty-nine first-grade students were recruited with parental consent from two Catholic schools in northwest Calgary, Canada. Both schools were located in middle-class neighborhoods. One student withdrew during the experiment, so that the final data were collected from twenty-eight subjects, including 16 boys and 12 girls, with a mean age of 7.01 (standard deviation =.40). None of the subjects had received any previous formal instructions on the problem types involved in the experiment.

Materials

Problems

Thirty simple arithmetic word problems were created based on previous studies (Riley et al., 1983; Carpenter et al., 1981; Hudson, 1983; De Corte, Verschaffel, & De Win, 1985). The thirty problems shared the same mathematical structure, specifically finding the numerical difference between two disjoint sets. The text of these problems distinguished three types of problems: COMPARE, EQUALIZE, and WON'T GET (see Table 2). Each type consisted of ten problems.

The COMPARE problems involved a static comparison between two disjoint sets. The two sets were presented and the question was "how many more objects of the larger set are there than those of the smaller set". For example: "John has 9 apples. Ann has 4 apples. How many more apples does John have than Ann?"

The EQUALIZE problems involved actions reqired to make two disjoint sets numerically equal. The two sets were presented and the question was posed as "what could be done to the smaller set to make it equal to the larger set". For example: "Fred has 9 buckets. Betty has 5 buckets. How many more buckets does Betty have to get to have as many buckets as Fred?"

The WON'T GET problems involved finding element correspondence between two disjoint sets. The two sets were presented, then the question was "how many objects of the larger set won't get the objects of the smaller set". For example: "8 children went to a store to buy hats. There were only 5 hats in the store. How many children would not get a hat?"

Among all the problems, the objects of the sets and the owners of the objects varied. The numbers used to present the larger set were not greater than 9. The number triples involved in the problems had ten variations: 9-5-4, 9-4-5, 9-6-3, 9-3-6, 8-5-3, 8-3-5, 7-4-3, 7-3-4, 7-5-2, and 6-4-2. (For all the problems, see Appendix A)

Blocks

A set of 20x20x5mm wooden square blocks (Activity

Resources Company Inc., 1973) were used. The blocks were presented with half of the problems of each type. The presence and absence of the blocks produced two problem solving conditions, namely "without blocks" and "with blocks".

Tasks

The subjects were required to solve the three types of problems without blocks and then with blocks. Five problems of each type would be solved without blocks. Another five problems of each type would be solved with blocks. The tasks included (1) giving an answer to the question of each problem, then (2) identifying the strategy he/she tried to use to solve the problem. Based on previous studies of solving problems involving difference finding between disjoint sets, three kinds of strategies were expected, namely PART-WHOLE, MATCH, and ADD-ON.

The PART-WHOLE strategy treats the larger set as whole, the smaller set as one part, and the difference as the other part.

The MATCH strategy matches up the objects of the two sets and finds the objects left in the larger set as the difference.

The ADD-ON strategy adds more objects to the smaller set to make it equal to the larger set, then takes the added objects as the difference.

Procedures

The testing was conducted in May of 1992, when none of the subjects had yet received any formal lessons on the types of problems involved in the experiment. Each subject was tested in an individual interview situation in his/her school. Each subject was tested in three sessions (20 minutes each), with a one week interval between sessions.

The subjects from one school started solving the 15 problems (5 from each type) presented without blocks, and then solved the other 15 problems (5 from each type) presented with blocks. The subjects from the other school started with the problems with blocks, and then the problems without blocks. Every subject solved exactly the same problems. However, the sequence of presentation of the problems was randomly arranged for each subject.

At the time of testing, each problem was read to the subject. A problem might be read twice, if required by the subject. First the subject gave an answer for each problem. Under the without-blocks condition, using fingers was prohibited. Under the with-blocks condition, the researcher would make two piles of blocks (without any particular pattern of arrangement) on a table to represent the two given sets as reading the problem. The subject's numerical solution to the problem was recorded.

After solving a problem, the subject was asked to identify the strategy he/she used to get the solution, no matter whether he/she got a correct or incorrect numerical solution. For this purpose, the researcher would ask two questions "how did you get the answer?" and "why?", and only these two questions. The questions were asked in a neutral manner, no guidance or directions were provided. Under the without-blocks condition, the subject was asked to verbally explain the strategy. Under the with-blocks condition, the subject was allowed to rearrange the blocks to show how he/she got the answer.

The numerical solution to each problem and the strategy presented by the subject were recorded on a sheet on which the question itself, "right" and "wrong" marks, and the categories of the strategies had been printed.

Scoring

Correct Solutions

A score of "1" was awarded to each correct numerical solution to the problem, and a "0" to each incorrect solution. Because there were five problems for each type of problem under each condition, the score for each subject ranged from 0 to 5 for each type of problem under each condition.

Strategies

The strategy a subject used to solve a problem would fall into one of the following four categories: (1) PART-WHOLE, when the subject presented "larger set - smaller set = difference" or "smaller set + difference = larger set"; (2)

MATCH, when the subject matched the objects of the smaller set to those of larger set, then counted the objects left in the larger set as difference; or took away the objects from larger set until the two sets matched up, then counted the objects taken away as the difference; (3) ADD-ON, when the subject added more objects to the smaller set until it equalled to the larger set, then counted the added objects as the difference; (4) OTHER, when the subject stated "I have no idea", "I can not remember", "I don't know", and "I guessed", etc., or when the strategy presented by the subject did not fit any of the above known strategies.

A score of "1" was awarded to one of the four categories according to the subject's presentation for each problem, no matter he/she got a correct numerical answer to the problem or not. Because each subject solved five problems of each type under each condition, his/her possible score for each category would range from 0 to 5.

CHAPTER V
RESULTS

A set of repeated-measures analyses of variance were performed to analyze (1) the correct solutions of the problems, and (2) the strategies for solving the problems. The analyses were conducted by SPSS/PC$^+$ through SPSS MANOVA (Norusis/SPSS Inc., 1990). The F values and their significance were calculated by the averaged test of significance, which is equivalent to a mixed-model univariate analysis of variance (Winer, 1971). The Mauchly Sphericity Test was used to check the violation of the assumption that the variance of all the transformed variables for an effect be equal and that their covariance be zero. When the assumption was violated, the degrees of freedom for the F test were adjusted by the Huynh-Feldt Epsilon (Huynh & Feldt, 1967). Only the final, adjusted results will be reported in the following section. Focused analyses for significant main effects were performed by pairwise contrast analyses (Rosenthal & Rosnow, 1985), and the corresponding parameter estimates will be reported.

Correct Solutions

The means and standard deviations of the correct solutions of the three types of problems under two conditions of availability of blocks are shown in Table 3. A 3 x 2 (Type x Condition) repeated-measures analysis of variance was conducted to reveal the difference in correct solutions among the three problem types under different conditions. Significant differences were found in the main effect of Type, $F(1,37) = 10.45$, $p < .001$, and Condition, $F(1,27) = 19.68$, $p < .001$. There was no significant interaction between Type and Condition, $F(2,54) < 1$.

For the Type effect, three pairwise contrast analyses were performed (Table 4), and showed that the mean solutions of COMPARE problems was significantly lower than those of both EQUALIZE and WON'T GET problems, $t = 3.04$, $p < .01$, and $t = 3.70$, $p < .01$, respectively, while there was no difference between EQUALIZE and WON'T GET problems, $t = 1.20$, $p > .05$.

The results showed that (1) no matter whether the problems were presented with or without blocks, COMPARE problems were the most difficult problem type compared to EQUALIZE and WON'T GET problems, and (2) across the three types, the problems presented with blocks were easier than those presented without blocks.

Strategies

Without Blocks

Table 5 shows the means and standard deviations of the strategies for solving each type of problems presented without blocks. A significant interaction between Type and Strategy was found through a 4 x 3 (Strategy x Type) repeated-measures analysis of variance, $F(3,94) = 53.63$, $p < .0001$. This analysis indicated that the frequencies of using different strategies depend upon the types of problems. Therefore, a set of one-way repeated-measures analyses of variance on the strategies for each Type were performed to reveal the differences of usage frequency among the strategies. Alpha was set at .017 (.05/3) to control the Type I error rate.

For COMPARE problems, no significant difference was found among the strategies, $F(3,81) = 1.79$, $p > .05$. This indicates that the first-graders were not clear how to solve COMPARE problems, because they did not use any strategy more often than any other, and especially because there was no difference between OTHER, the unclassifiable strategies, and the strategies suitable for difference finding problems.

When solving EQUALIZE and WON'T GET problems, the tendency to use one strategy over the others was significant. For EQUALIZE problems, the Strategy main effect was significant, $F(2,54) = 31.09$, $p < .0001$. The contrast analyses further revealed the use of the ADD-ON strategy overwhelmingly more than any other strategies (Table 6), indicating that the first-graders mainly tried to use the ADD-ON strategy to solve EQUALIZE problems.

For WON'T GET problems, the Strategy main effect was also significant, $F(2,56) = 38.81$, $p < .0001$. The pairwise contrast analyses showed that among the strategies, MATCH was most frequently used (Table 7), which suggested that WON'T GET problems were usually solved by using MATCH strategy.

To summarize, when the three types of problems were presented without blocks, using a particular strategy to solve each of the problems depended on the problem type which determined by the problem text. The kind of strategy that was used for COMPARE problem was not clear. However, it was clear that the first-graders mostly used the ADD-ON strategy to solve the EQUALIZE problems, and the MATCH strategy to solve the WON'T GET problems.

With Blocks

When solving the problems presented with blocks, the first-graders' strategy use showed a similar pattern to that for solving the problems presented without blocks. The means and standard deviations of the strategy scores for each type of problem presented with blocks are shown in Table 8. A significant interaction between Type and Strategy was found through another 4 x 3 (Strategy x Type) repeated-measures analysis of variance, $F(4,104) = 31.55$, $p < .0001$. Furthermore, a set of one-way repeated-measures analyses of variance on the strategies for each Type were performed to reveal the differences of usage frequency among the strategies. Alpha was set at .017 (.05/3) to control the Type I error rate.

For EQUALIZE problems, a significant Strategy main effect was found, $F(2,57) = 17.00$, $p < .0001$. Through the pairwise contrast analyses (Table 9), the ADD-ON strategy was found as the most frequently used strategy.

For WON'T GET problems, a significant Strategy main effect was found, $F(1, 39) = 154.36$, $p < .0001$. A set of pairwise contrast analyses (Table 10) showed that the MATCH strategy was used more than other strategies.

The presence or absence of blocks did not seem to influence the use of strategies to solve the EQUALIZE or WON'T GET problems. No matter whether the problems were presented with blocks or not, the first-graders tended to use the ADD-ON strategy to solve the EQUALIZE problems, and the MATCH strategy to solve the WON'T GET problems.

However, presenting the problems with blocks did make a minor difference in strategy use for solving COMPARE problems. Under this condition the Strategy main effect for COMPARE problems was significant, $F(3,81) = 3.75$, $p = <.017$. The only two significant results by the contrast analyses on strategies for COMPARE problems (Table 9) were between the MATCH and PART-WHOLE strategies, and between MATCH and ADD-ON strategies. These showed that when solving COMPARE problems presented with blocks, the MATCH strategy was used more than the PART-WHOLE and the ADD-ON strategies. The comparison between MATCH and OTHER, however, was not significant.

These results from the analyses on the COMPARE problems presented with blocks suggested that, when the blocks were available, the first-graders tended to use the MATCH strategy to solve the COMPARE problems rather than use the PART-WHOLE or the ADD-ON strategies. However, the MATCH strategy was used only more often than the PART-WHOLE and the ADD-ON strategies. It was this difference that produced the significant main effect. Another important piece of information from the analyses was that the MATCH strategy was not used significantly more often than the OTHER strategy. The rather high value for unspecified or random strategies does not give a clear picture of exactly what strategies they could use to solve the COMPARE problems.

The results from the analyses on the strategies clearly showed the dependency of strategy use on the problem type which is determined by the problem text, across the two conditions of block availability. The EQUALIZE problems were mostly solved by using the ADD-ON strategy, and the WON'T GET problems were mostly solved by using the MATCH strategy. Under both conditions, the first-graders tried to use some unclassifiable or random strategies as well as the strtegies suitable for difference-finding problems to solve the COMPARE problems. There was no significant difference between the use of unclassifiable strategies and the use of the suitable strategies, indicating that the first-graders were not clear what they could use to solve the COMPARE problems.

The effect of the blocks on strategy use showed on the COMPARE problems. When the blocks were not available, no one strategy was used significantly more often than any others. When the blocks were available, the MATCH strategy was used more often than the ADD-ON and the PART-WHOLE strategies.

Table 3

<u>Mean Solutions and Standard Deviations (in parentheses) by Type and Condition</u>

Condition	Type of problem		
	COMPARE	EQUALIZE	WON'T GET
Without blocks	3.04	3.82	3.89
	(1.90)	(1.28)	(1.20)
With blocks	3.68	4.68	4.93
	(1.93)	(0.61)	(0.26)

Note. Maximum score = 5.

Table 4

Contrast Analyses for Type Main Effect

Comparison	t
COMPARE vs. EQUALIZE	3.04**
COMPARE vs. WON'T GET	3.70**
EQUALIZE vs. WON'T GET	1.20

**p<.01.

Table 5

Means and Standard Deviations (in parentheses) of Strategies Used for Each Type (Without Blocks)

Type	Strategy			
	PART-WHOLE	MATCH	ADD-ON	OTHER
COMPARE	1.18	0.61	1.64	1.57
	(1.68)	(0.92)	(1.62)	(2.08)
EQUALIZE	0.68	0.25	3.43	0.64
	(1.49)	(0.52)	(1.50)	(1.03)
WON'T GET	0.71	3.68	0.25	0.36
	(1.41)	(1.56)	(0.65)	(0.95)

Note. Maximum score = 5.

Table 6

Contrast Analyses on Strategies for EQUALIZE (Without Blocks)

Comparison	*t*
ADD-ON vs. PART-WHOLE	5.18**
ADD-ON vs. MATCH	10.44**
ADD-ON vs. OTHER	7.09**

**$p<.01$.

Table 7

Contrast Analyses on Strategies for WON'T GET (Without Blocks)

Comparison	*t*
MATCH vs. PART-WHOLE	5.67**
MATCH vs. ADD-ON	9.89**
MATCH vs. OTHER	8.26**

$^{**}p<.01$.

Table 8

Means and Standard Deviations (in parentheses) of Strategies Used for Each Type (With Blocks)

	Strategy			
Type	PART-WHOLE	MATCH	ADD-ON	OTHER
COMPARE	0.68	2.28	0.75	1.29
	(1.52)	(2.14)	(1.30)	(1.94)
EQUALIZE	0.71	1.11	2.96	0.21
	(1.27)	(1.40)	(1.77)	(0.63)
WON'T GET	0.32	4.54	0.04	0.11
	(0.95)	(1.17)	(0.19)	(0.57)

Note. Maximum score = 5.

Table 9

Contrast Analyses on Strategies for EQUALIZE (With Blocks)

Comparison	*t*
ADD-ON vs. PART-WHOLE	4.40**
ADD-ON vs. MATCH	3.37**
ADD-ON vs. OTHER	14.89**

**p<.01.

Table 10

Contrast Analyses on Strategies for WON'T GET (With Blocks)

Comparison	*t*
MATCH vs. PART-WHOLE	10.91**
MATCH vs. ADD-ON	19.81**
MATCH vs. OTHER	14.89**

$^{**}p < .01$.

Table 11

Contrast Analyses on Strategies for COMPARE (With Blocks)

Comparison	t
MATCH vs. PART-WHOLE	2.79*
MATCH vs. ADD-ON	2.85**
MATCH vs. OTHER	1.45

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

CHAPTER VI
DISCUSSION

A New Interpretation of Text Effect

Among the three types of word problems involving finding the numerical difference between two disjoint sets, COMPARE, EQUALIZE, and WON'T GET, the COMPARE problems were found to be the most difficult type in this study. This is consistent with the findings of Hudson (1980, 1983) and the prediction of Briars and Larkin's (1984) word problem solving model. It is also consistent with the first hypothesis of the present study, which was that the COMPARE problem texts do not facilitate the mapping onto the coordination of two number lines, with the result that the first-graders encountered more difficulty than solving the EQUALIZE and WON'T GET problems. This is the present study's interpretation of the effect of problem text.

This interpretation was supported when the strategy data were analyzed. When solving the EQUALIZE problems, first-graders tended to use the ADD-ON strategy. When solving the WON'T GET problems, they tended to use the MATCH strategy. Although they are different in terms of involving different actions, the ADD-ON and MATCH strategies have one thing in common, that is, they both try to coordinate the two number lines and to constrain their actions by using the criterion set by the coordination. The ADD-ON strategy acts on the number line which represents the smaller set, adding more objects to it, but the action of adding is always limited by the length of the other number line which represents the larger set.

According to the descriptions of previous studies (Carpenter et al., 1981; Hudson, 1983), the MATCH strategy first projects the shorter number line onto the longer line, building one-to-one correspondence, and then counts the units unmatched on the longer line. This was found in the present study, too. In addition, the present study observed another variation of matching. The problem solver arranged two object lines, side by side, then took away objects from the longer line, one by one, until the two lines had the equal length, i.e., matched. The taking away action was always constrained by the length of the shorter number line. No matter which kind of MATCH strategy the problem solvers used, it was done by coordinating the two number lines.

Using a particular strategy reflects the kind of representation the problem solver has been constructing. The ADD-ON and MATCH strategies showed that the problem models built by the first-graders were represented by the coordination of the two number lines, when they tried to solve the EQUALIZE and the WON'T GET problems. On the other hand, first-graders tried to use various other strategies to solve the COMPARE problems, including the strategies which were unclassifiable and undefinable. Generally, among those alternative strategies, no one was used more times than any others. This reflected the uncertainty of first-graders on what kind of problem model they could build according to the

COMPARE problem texts.

The only difference among the three type of problems was problem text. Thus the source of the relative difficulty among the problems can only be traced back to the problem texts. Through the analyses of the strategies used for solving the problems, it can be concluded that the problem texts of the EQUALIZE and the WON'T GET problems could be easily mapped onto the coordination of two number lines, so the appropriate problem models could be constructed to produce the correct solutions. Specifically, the phrase "what to do to get as many ... as ..." in the EQUALIZE problem texts, and the phrase "how many won't get ..." in the WON'T GET problem texts, cued the coordination of two number lines. The COMPARE problem text did not cue the coordination of two number lines, which is the prerequisite for solving such problems, so that this type became the most difficult difference-finding problem.

This interpretation produced an answer to the question "what makes the problem involving comparison so difficult," which was central to the main issue of the greater difficulty of the Compare problems. Riley et al. (1983) and Cummins (1991) could not provide the answer because they assumed that the part-whole structure was responsible for the problem representation. When Riley and Greeno (1988) found that the part-whole structure was not appropriate, and when Cummins (1991) proposed that the source of difficulty was the failure of mapping, they still failed to find a new underlying mathematical structure. Okamoto (1992) proposed the coordination between two number lines as the underpinning of the construction of the problem model, which provided the structural base for the study of mapping. However, Okamoto compared the Compare problems with other categories such as the Change problems and the Combine problems. The lack of comparability across the categories reduces her model's explanatory power on the issue of the relative difficulty of comparison problems. The present study adopted Okamoto's proposal about the coordination of two number lines as the problem model, and employed three types of word problems sharing the same mathematical essence. Thus the effect of the problem texts was revealed.

The present study failed to find a differential effect of using blocks on the problem representation. The result that using blocks generally facilitated the problem solving performance for all types was open to various interpretations, for example, the blocks might not influence the representation, but merely reduce processing load.

When the blocks were available, first-graders tended to use the MATCH strategy more often than another coordination-appropriate strategy, ADD-ON, to solve the COMPARE problems. Exactly why one appropriate strategy was cued by the blocks but not the other, cannot be explained within the framework of the present study. The effect of concrete materials on the problem representations and its interaction with the problem texts is a topic for the future studies. The blocks used in the present study were still relatively abstract, because they were used as referents for any kinds of objects.

If some "naturally paired" objects were presented, such as eggs and egg cups, desks and chairs, birds and nests etc., the MATCH strategy would be more easily cued. Although this would seem a fruitful direction for future research, it is beyond the scope of the present study. What was found in the present study was that the MATCH strategy was not used more often than those unclassifiable and undefinable strategies even when the blocks were available. Therefore, what made the COMPARE problem difficult was still the problem text, which apparently does not cue the coordination of two number lines.

Limitations

Although this study extends previous research, it was limited to problems involving finding the numerical difference between two disjoint sets. The interpretation derived from this study may not be able to explain the effect of rewording in other problem categories. The rewording effect on the Combine problems found in Carpenter et al.'s (1981) and Lindvall and Ibarra's (1980b) studies seems to be related to whether the texts can make the part-whole structure more explicit. De Corte, Verschaffel, and De Win (1985) found such an effect of rewording on the Change problems. Because the Change problems do not involve disjoint sets, the problems probably can be solved by acting on one number line.

Also, the present study did not deal with the social context and situational aspects of the problem texts. It was found that the context personalization (Davis-Dorsey, Ross, & Morrison, 1991), the structure, role, and intent of the word problem texts (De Corte & Verschaffel, 1985), and the situation described in the problem texts (Reusser, 1990; Reusser et al., 1990) all affect relative difficulty. However, the present study focused on just the one aspect of the problem texts relating to the more "pure" mathematical structures.

Another possible limitation of this study is the methodology used to get the strategy data. Pressley (1992) questioned the microgenetic method for studying cognitive development, pointing out that the children's choice of strategy may be cued by the experimenters' promoting specific options during the interviews. This was not the case in the present study, in which the experimenter only asked "how did you get the answer" and "why", and provided no guidance and directions. However, this raises the question of whether the children's own report, especially when there was no blocks, could be taken as the evidence of their strategy use. This method is not without its critics (Ginsburg, Kossan, Schwartz, & Swanson, 1983). In fact, one of the main criticisms is that a child may fail to report a considerable part of his/her thinking process. However, the interview method and the children's verbal report are still the major source of any information of problem solving processes. According to Ginsburg et al. (1983), to judge the value of subject's descriptions of some of their cognitive processing depends on how one construes the notion of processing. If the notion is limited to descriptions of neural processing, the introspective reports will be no use; if stages or steps taken in solving a problem, such as in mathematics, are the aspects of processing that are of interest, verbal reports may be a valuable source of information. Certainly, it would be more reliable if this method were combined with some other more "objective" method, such as De Corte and Verschaffel's (1990) collection of eye-movement data while children read and solve word problems. However, it is not clear just how well the physiological and self-report assessments match up, and in any event the strategies found here have high face validity.

Implications

This study examined the very nature of word problems, that is, the fact that the word problems are mathematical problems presented in verbal forms. From this perspective, further research on word problem solving processes should focus on the interaction between linguistic comprehension and mathematical reasoning. This means not only trying to find what kind of linguistic competence is required, or what kind of mathematical structure is underlying the problem representation, but also we should examine how the two components interact. To analyze the mapping processes in detail is one way to understand the interaction.

In future studies, this method should be extended to all other problem types to study the different mapping processes under the current categorization. It may also be used in studying word problems beyond arithmetic, such as algebra word problems. Another related area may be to study the degrees of children's dependency on linguistic comprehension at different cognitive developmental level, or with different amount of exposure and exercise. This could help explain why the COMPARE problems were difficult for first-graders but not for third- or second-graders.

Educational research should integrate the findings of word problem solving studies into the studies of mathematics curriculum and instruction. At the present time the problems involving comparison are introduced to first-grade textbooks without any preparation for the students. The information-processing analysis and detailed descriptions of word problem solving processes, including the mapping processes between linguistic comprehension and mathematical reasoning, would provide the base on which curriculum revision should be developed and instruction should be designed. Also, the findings from the studies on the development of word problem solving ability should be integrated into educational studies, so that curriculum can be sequenced according to the development of children's ability, and educators can know how to get around the difficult mathematical points which originate in the constraints in cognitive development by making use of the influence of problem texts.

During the course of data collection for the present study, a first-grade teacher predicted that the problems used here would be too difficult for the students, because (by the end of the first year) the students "have met some of the COMPARE problems, but have not seen the EQUALIZE and the WON'T GET types at all". The teacher suggested that the researcher do this experiment in grade two instead. This example shows there is a great deal of information not delivered to teachers, including that (1) the three types share the same mathematical structure and require the same problem solving competence; (2) the COMPARE problems are more difficult than the other two types, and this is not the problem of exposure but the basic problem text; (3) it would be better to let students encounter the EQUALIZE and the WON'T GET problems before the COMPARE problems; and (4)maybe a better way to teach solving the COMPARE problems is to let the students first do the mapping in the manner that they do in solving the EQUALIZE and WON'T GET problems, that is, by using matching or adding-on strategies, so that they may avoid the difficult point intrinsic to the COMPARE problem texts.

REFERENCES

Briars, D. J., & Larkin, J. H. (1984). An integrated model of skill in solving elementary word problems. Cognition and Instruction, 1(3), 245-296.

Buckingham, B. R., & MacLatchy, J. (1930). The number abilities of children when they enter grade one. In 29th Yearbook of the National Society for the Study of Education. Bloomington, IN: Public School Publishing.

Capodilupo, A. M. (1992). A Neo-Piagetian analysis of children's response to instruction in the sight-reading of music notation. In R. Case, The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge (pp. 99-115). Hillsdale, NJ: LEA.

Carpenter, T. P. (1985). Learning to add and subtract: An exercise in problem solving. In Edward A. Silver (Ed.), Teaching and learning mathematical problem solving: Multiple research perspectives (pp. 17-40). Hillsdale, NJ: LEA.

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem-solving skills. In T. P. Carpenter, J. M. Moser and T. Romberg (Eds.), Addition and subtraction: A cognitive perspective (pp. 9-24). Hillsdale, NJ: LEA.

Carpenter, T. P., & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. Journal for Research in Mathematics Education, 15(3), 179-202.

Carpenter, T. P., Hiebert, J., & Moser, J. M. (1981). Problem structure and first-grade children's initial solution processes for simple addition and subtraction problems. Journal for Research in Mathematics Education, 12(1), 27-39.

Case, R. (1985). Intellectual development: Birth to adulthood. New York: Academic Press.

Case, R. (1992). The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge. Hillsdale, NJ: LEA.

Case, R., & Griffin, S. (1990). Child cognitive development: The role of central conceptual structure in the development of scientific and socila thought. In C. A. Haver (Ed.), Advances in Psychology: Developmental Psychology (pp. 193-229). North Holland: Elsevier.

Case, R., & Sandieson, R. (1992). Testing the presence of a central quantitative structure: Use of the transfer paradigm. In R. Case, The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge (pp. 117-132). Hillsdale, NJ: LEA.

Case, R., & Sowder, J. T. (1990). The development of computational estimation: A Neo-Piagetian analysis. Cognition and Instruction, 7(2), 79-104.

Cummins, D. D. (1991). Children's interpretations of arithmetic word problems. Cognition and Instruction, 8(3), 261-289.

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

Davis-Dorsey, J., Ross, S. T., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. Journal of Educational Psychology, 83(1), 61-68.

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

De Corte, E., & Verschaffel, L. (1987). The effect of semantic structure on first graders strategies for solving addition and subtraction word problems. Journal for Research in Mathematics Education, 18(5), 363-381.

De Corte, E., & Verschaffel, L. (1988). Computer simulation as a tool in research on problem solving in subject-matter domains. International Journal of Educational Research, 12(1), 49-69.

De Corte, E., Verschaffel, L., & De Win, L. (1985). Influence of rewording verbal problems on children's problem representations and solutions. Journal of Educational Psychology, 77(4), 460-470.

De Corte, E., Verschaffel, L., & Pauwels, A. (1990). Influence of the semantic structure of word problems on second graders' eye movements. Journal of Educational Psychology, 82(2), 359-365.

Dellarosa, D. (1986). A computer simulation of children's arithmetic word-problem solving. Behavior Research Methods, Instruments, and Computers, 18(2), 147-154.

Dellarosa, D., Weimer, R., & Kintsch, W. (1985). Children's recall of arithmetic word problems (Institute of cognitive Science Tech. Rep. No. 14). Univeristy of Colorado, Boulder.

Donaldson, M. (1979). Children's Minds. New York: Norton.

Fletcher, C. (1985). Understanding and solving arithmetic word problems: A computer simulation. Behavior Research Methods, Instruments, and Computers, 17(5), 565-571.

Fuson, K. C. (1982). An analysis of the counting-on solution procedure in addition. In T. P. Carpenter, J. M. Moser and T. Romberg (Eds.), Addition and subtraction: A cognitive perspective (pp. 67-81). Hillsdale, NJ: LEA.

Gelman, R. & Callistel, C. R. (1978). The Child's understanding of number. Cambridge, Mass.: Harvard University Press.

Gelman, R. & Greeno, J. G. (1989). On the nature of competence: Principles for understanding in a domain. In L. B. Resnick (Ed.), Knowing, learning and instruction. Hillsdale, NJ: LEA.

Gibb, E. G. (1956). Children's thinking in the process of subtraction. Journal of Experimental Education, 25, 71-80.

Ginsburg, H., Kossan, N. E., Schwartz, R., & Swanson, D. (1983). Protocol methods in research on mathematical thinking. In H. P. Ginsburg (Ed.), the development of mathematical thinking (pp. 153-196). New York: Academic Press.

Griffin, S., Case, R., & Sandieson, R. (1992). Synchrony and asynchrony in the acquisition of children's everyday mathematical knowledge. In R. Case, The mind's staircase:

Exploring the conceptual underpinnings of children's thought and knowledge (pp. 75-97). Hillsdale, NJ: LEA.

Hebbeler, K. (1977). Young children's addition. The Journal of Children's Mathematical Behavior, 1, 108-121.

Heller, J. I., & Greeno, J. G. (1978). Semantic processing in arithmetic word problem solving. Paper presented at the Midwestern Psychological Association Convention, Chicago.

Hiebert, J. (1981). Young children's solution processes for verbal addition and subtraction problems: The effect of the position of the unknown set. Paper presented at the 95th annual meeting of the National Council of Teachers of Mathematics, St. Louis.

Hudson, T. (1980). Young children's difficulty with "How many more _____ than _____ are there?" questions. (Doctoral dissertation, Indiana University). Dissertation Abstract International, July 1980, B41, (01).

Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. Child Development, 54, 84-90.

Huynh, H. & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. Journal of Educational Statistics, 1, 69-82.

Ibarra, C. G., & Lindvall, C. M. (1979). An investigation of factors associated with children's comprehension of simple story problems involving addition and subtraction prior to formal instruction on these operations. Paper presented at the annual meeting of the National Council of Teachers of Mathematics, Boston.

Kintsch, W. (1986). Learning from text. Cognition and Instruction, 3(2), 87-108.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. Psychological Review, 92(1), 109-129.

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

LeBlanc, J. (1968). The performance of first-grade children in four levels of conservation of numerousness and three IQ groups when solving arithmetic subtrction problems. Unpublished doctoral dissertation, University of Wisconsin.

Lindvall, C. M. & Ibarra, C. G. (1980a). Incorrect procedures used by primary grade pupils in solving open addition and subtrction sentences. Journal for Research in Mathematics Education, 11(1), 50-62.

Lindvall, C. M. & Ibarra, C. G. (1980b). A cinical investigation of the difficulties evidenced by kindergarten children in developing "models" in the solution of arithmetic story problems. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Langford, P. E. (1986) Arithmetic word problems: Thinking in the head versus thinking on the table. Educational Studies in Mathematics, 17, 193-199.

Marini, Z. (1992). Synchrony and asynchrony in the development of children's scientific reasoning. In R.

Case, The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge (pp. 55-73). Hillsdale, NJ: LEA.

Mayer, R. E. (1985). Mathematical ability. In R. J. Sternberg (Ed.), Human abilities: An information processing approach (pp. 127-150). New York: Freeman.

Mayer, R. E. (1986). Mathematics. In R. F. Dillon & R. J. Sternberg (Eds.), Cognition and Instruction (pp. 127-154). New York: Academic Press.

Nesher, P. (1982). Levels of description in the analysis of addition and subtraction. In T. P. Carpenter, J. M. Moser, & T. Romberg (Eds.), Addition and Subtraction: A cognitive perspective (pp. 25-38). Hillsdale, NJ: Erlbaum.

Nesher, P. & Katriel, T. (1978). Two cognitive modes in arithmetic problem solving. Paper presented at the second annual meeting of the International Group for the Psychology of Mathematics Education, Onabruck, West Germany.

Norusis, M. J./SPSS Inc. (1990). SPSS advanced statistics user's guide. SPSS Inc.

Okamoto, Y. (1992). A developmental analysis of children's knowledge and processes for solving arithmetic word problems. Unpublished doctoral dissertation, Stanford University.

Piaget, J. (1965). The child's conception of number. New York: Norton.

Pressley, M. (1992). How not to study strategy discovery. American Psychologist, October, 1240-1241.

Resnick, L. B. (1983). A developmental theory of number understanding. In Herbert P. Ginsburg (Ed.), The development of mathematical thinking (pp. 109-151). New York: Academic Press.

Resnick, L. B. (1989). Developing mathematical knowledge. American Psychologist, 44, 162-169.

Resnick, L. B. & Ford, W. W. (1981). The psychology of mathematics for instruction. Hillsdale, NJ: LEA.

Reusser, K. (1990). Understanding word arithmetic problems: Linguistic and situational factors. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Reusser, K., Kämpfer, A., Sprenger, M., Staub, F. Stebler, R., & Stüssi, R. (1990). Tutoring mathematical word problems using solution trees. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Riley, M. S. (1981). Conceptual and Procedural knowledge in development. Unpublished Master's thesis, University of Pittsburgh.

Riley, M. S. & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. Cognition and Instruction, 5(1), 49-101.

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In Herbert P. Ginsburg (Ed.), The development of mathematical thinking (pp. 153-196). New York: Academic Press.

Rosenthal, R. & Rosnow, R. L. (1985). Contrast analysis: Focused comparisons in the analysis of variance. Cambridge: Cambridge University Press.

Schell, L. M., & Burns, P. C. (1962). Pupil performance with three types of subtraction situations. School Science and Mathematics, 62(3), 208-214.

Shores, J., & Underhill, R. G. (1976). An analysis of kindergarten and first-grade children's addition and subtraction problem-solving modeling and accuracy. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Siegel, L. S. (1978). The relationship of language and thought in the preoperational child: A reconsideration of nonverbal alternatives to Piagetian tasks. In C. J. Brainerd & L. S. Siegel (Eds.), Alternatives to Piaget: Critical essays on the theory. New York: Academic Press.

Steff, L. P., & Johnson, D. C. (1971). Problem-solving performances of first-grade children. Journal for Research in Mathematics Education, 2, 50-64.

Tamburino, J. L. (1980). An analysis of the modelling processes used by kindergarten children in solving simple addition and subtraction story problems. Unpublished Master's thesis, University of Pittsburgh.

van Dijk, T. A. & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.

Vergnaud, G. (1982). A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. In T. P. Carpenter, J. M. Moser and T. Romberg

(Eds.), Addition and subtraction: A cognitive perspective (pp. 39-59). Hillsdale, NJ: LEA.

Winer, B. J. (1971). Statistical principles in experimental design. New York: McGraw-Hill.

APPENDIX A

Word Problems Used in the Experimentation

COMPARE

1. Joe has 8 marbles.

   Tom has 5 marbles.

   How many more marbles does Joe have than Tom?

2. Peter won 4 prizes at a fair.

   Mary won 7 prizes.

   How many more prizes did Mary win than Peter?

3. John has 9 apples.

   Ann has 4 apples.

   How many more apples does John have than Ann?

4. Tom has 9 toy cars.

   And he has 5 toy trucks.

   How many more toy cars does he have than his toy trucks?

5. You got 8 books.

   I got 3 books.

   How many more books did you have than I?

6. Mark and Sue like cups.

   Mark collected 3 cups.

   Sue collected 7 cups.

   How many more cups did Sue collect than Mark?

7. Susan puts some pencils in her pencil box.

   There are 6 green pencils.

(Continued next page)

(Continued)

And there are 4 red pencils.

How many more green pencils are there than red pencils?

8. There are 9 blue balloons in the sky.

   And there are 6 yellow balloons there.

   How many more blue balloons are there than yellow ones?

9. My brother read 5 books in a week.

   My sister read 7 books.

   How many more books did my sister read than my brother?

10. In a zoo, there are 9 monkeys.

    And there are 3 panda bears.

    How many more monkeys are there than panda bears?

EQUALIZE

1. Joan picked 9 flowers.

   Bill picked 4 flowers.

   Bill wanted to have as many flowers as Joan.

   How many more flowers would he need to pick?

2. Fred has 9 buckets.

   Betty has 5 buckets.

   How many more buckets does she have to get to have as many buckets as Fred?

3. There are 8 desks in a classroom.

   And there are 3 chairs there.

   If we need to have as many chairs as desks,

(Continued next page)

(Continued)

How many more chairs will we need?

4. Some children are playing marble games.

   Steve has 5 marbles.

   Wade has 8 marbles.

   How many marbles does Steve have to win to have as many marbles as Wade?

5. At first, Wendy drew 7 pictures.

   Jill drew 3 pictures.

   Jill drew some more so she had as many pictures as Wendy.

   How many more pictures did Jill draw?

6. One day, children were telling stories.

   A girl told 7 stories.

   A boy told 4 stories.

   The boy told more stories later so he told as many stories as the girl.

   How many more stories did the boy tell?

7. 6 boys and 4 girls came to Bob's birthday party.

   But Bob invited as many girls as boys to his party.

   So we know some girls were late.

   How many girls were late?

8. Tony's mom gave him 9 crackers yesterday.

   And gave him 6 today.

   Tony wanted to have as many crackers today as he got yesterday.

(Continued next page)

(Continued)

So he asked mom for more.

How many more would he ask for?

9. Some children were preparing for a picnic.

They got 7 apples and 5 bananas.

But they need as many bananas as apples.

How many more bananas do they have to get?

10. You have 9 pencils.

I have 3 pencils.

If I want to have as many pencils as you,

How many more pencils do I have to get?

WON'T GET

1. There are 7 riders.

But there are only 5 horses.

How many riders won't get a horse?

2. There are 9 children in a room.

And there are 3 chairs in the room.

How many children won't get a chair?

3. Here are 9 children.

And here are 4 candies.

How many children can not get a candy?

4. 9 children went to a store to buy hats.

There were only 5 hats in the store.

How many children would not get a hat?

(Continued next page)

(Continued)

5. Here are 8 birds.
   And here are 3 worms.
   Suppose every bird wants to get a worm.
   How many birds won't get a worm?
6. There are 7 dogs.
   They are playing with 3 cats.
   Each dog wants to catch a cat.
   How many dogs won't get a cat?
7. 8 people went to a movie.
   But there were only 5 tickets left.
   How many people couldn't get a ticket?
8. There are 7 people getting on a bus.
   There are 4 empty seats.
   How many people won't get a seat?
9. There are 6 drivers.
   There are 4 cars can be driven.
   How many drivers will not get a car to drive?
10. Here are 9 pilots, and
    here are 6 planes at an airport.
    How many pilots won't get a plane?