1

PM-1 3 1/2"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file / Votre référence :

Our file / Notre référence :

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Mental Arithmetic from a Psychophysical Perspective:

Studies in the Viability of an Alternative Paradigm

by

Catherine A. Mirsky, B.A.

A thesis submitted to

the Faculty of Graduate Studies and Research

in partial fulfilment

of the requirements for the degree of

Master of Arts.

Department of Psychology

Carleton University

Ottawa, Ontario

September 6, 1995

©copyright

1995, Catherine Mirsky

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-08911-8

Canada

Name ANTHONY J. JONES

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.

PHILOSOPHY - GENERAL

SUBJECT TERM

0629 U.M.I.

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
Polymology 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

The undersigned recommend to the Faculty of Graduate Studies

and Research acceptance of the thesis

Mental Arithmetic from a Psychophysical Perspective:

Studies in the Viability of an Alternative Paradigm

submitted by Catherine A. Mirsky, B.A.

In partial fulfilment of the requirements for

the degree of Master of Arts.

Thesis Supervisor

Chair, Department of Psychology

Carleton University

September 22, 1995

Abstract

Prevailing cognitive models of mental arithmetic are dependent on results using production and verification tasks. In consideration of apparent weaknesses associated with these paradigms, an alternative paradigm was tested bringing together the mental arithmetic and number comparison literatures. Two experiments were conducted in order to examine the processes involved in the activation and use of arithmetic information for addition (e.g., $5+6$) in the context of a number comparison task. Participants were required to make two types of numerical comparisons: a) between numbers only (e.g., 11 and 10) and b) between simple addition problems and numbers (e.g., $5+6$ and 10). Finding tie effects and numerical comparison difficulty effects provide support for the validity of the current paradigm. Further, problem difficulty effects could be inferred. Unlike subtraction (Mirsky, 1993), semantic congruity effects were not found in the add and compare task. Importantly, the processes involved in mental arithmetic and the processes involved in making comparative judgments of numerical magnitude were found to be non-additive in nature. The impact of these findings on psychophysical theories of quantitative numerical representation are discussed.

Acknowledgements

I would like to thank my advisor William M. Petrusic for introducing me to the field of psychophysics and his faith in my abilities to master this formidable area of research. Through his support, guidance and infinite patience I have been able to accomplish a task that I would never have thought possible. I would also like to thank Peter Aubin for all the times I needed an interpretation of what my advisor had explained to me and for giving me the encouragement and strength to continue when things appeared their bleakest. I acknowledge and express my deepest gratitude to Edward Zolpis, who over the years, has taught me to trust and believe in myself. To all the members of the lab, Craig Leth-Steensen, Kate Oakley-McKeen and Derck Harrison, I would also like to extend my thanks for their friendship and support. It would be remiss of me not to express gratitude to my husband, and children for putting up with my long hours at work or my parents who picked up the phone more often than I. Finally, I would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) who provided the funding necessary to complete this project.

Table of Contents

Abstract	iii
Acknowledgements	iv
List of Figures	vi
List of Appendices	viii
Introduction	9
The Bases of Number Processing Models	10
Lesser Known Effects In Mental Arithmetic	13
Analogue Models of Mental Arithmetic and a Preverbal system of Counting and Arithmetic Reasoning	15
Empirical Regularities Upon Which Models are Built	33
Response Time Methods in Mental Arithmetic	34
Comparative Judgment, Addition and Additive Factors	36
Memory psychophysics: Facts and Theories	38
The Psychophysics of Number Comparisons	41
Fundamental Issues	43
Experiment 1	46
Experiment 2	70
General Discussion	85
Appendix A	99

LIST OF FIGURES

FIGURE	DESCRIPTION	PAGE
1	Comparison Type by Comparison Pair Interaction for Correct and Error Response Times	51
2	Comparison Type by Comparison Pair Response Time Interaction Matching Compound Comparisons for Sum	53
3	Comparison Type by Comparison Pair Interaction for Error Rates.	54
4	Instruction by Comparison Type Interaction for Response Time and Error Rate Measures	58
5	Decade by Comparison Type Interaction for Response Time and Error Rate Measures	59
6	Congruity Indices for Subtraction (Mirsky, 1993) and Experiment 1	63
7	Funnel-like Congruity Effects in the Second Decade for Addition and Comparison for Compound Comparisons Involving Unequal Addends	64
8	Addend Order Effects for Addend Split of Two	67
9	Non-Additivity in the Addition and Comparison Paradigm	74
10	Comparison Type by Comparison Pair Response Time Interaction for Correct and Error Responses Collapsed and Error Responses Only	76
11	Comparison Type by Comparison Pair Interaction for Error Rates	77
12	Comparison Type by Decade Interaction for Response Times and Error Rates	79
13	Funnel-like Congruity Effects in the Second Decade for Compound Comparisons	82

LIST OF APPENDICES

Appendix	Description	Page
A.1	Stimulus Set for Experiment 1	74
A.2	Stimulus Set for Experiment 2 - Simple and Compound Comparisons	75
A.3	Adding by Ones Stimulus Set for Experiment 2	76

Introduction

An area of cognitive research which has attracted increasing attention over the last decade is that of numerical cognition. Interest in the structure and processes involved in numerical processing has arisen out of pedagogical concerns and the realization that understanding the way systematically acquired information (i.e. mathematics) is laid down in memory could lead to a more generalized understanding of how material less systematically acquired is organized in memory (Ashcraft, 1992). Suppes and Groen (1967) suggested that a scientific and theoretical explication of the principles of the learning of mathematics was possible. With such a model, it would not only be possible to predict points of difficulty in learning, but also to "predict systematic variations in performance data involving mathematical concepts and skills already taught." (p.36). That is, it would be possible to predict individual differences in performance based on knowledge of the mathematical concepts and skills an individual has learned.

The models of numerical processing that have evolved over the last generation stem largely from response time and error rate data obtained from production and verification tasks (defined on pages 35-36). While these two paradigms often agree in terms of their findings, there are discrepancies that have yet to be satisfactorily resolved. Furthermore, researchers have repeatedly acknowledged that the limitations associated with these paradigms include accuracy of the response times and implicit assumptions regarding the unfolding of cognitive processes (Pachella, 1974; Restle, 1970).

The purpose of the present experiments is twofold. First, to develop an alternative paradigm that at once permits attainment of accurate and reliable response times, and second to examine more explicitly the validity and viability of arithmetical processes. Replicating robust effects that have been found using verification and production tasks would work towards this end. Finding additional significant effects using the present paradigm would call on explanations from available models of numerical cognition, thus providing a further means by which to discriminate amongst them.

The Bases of Number Processing Models

Problem Difficulty Effect.

The most profound and robust effect that has been found with respect to mental arithmetic is the problem difficulty effect (Ashcraft, 1992; Campbell & Graham, 1985; Dehaene, 1992; Groen & Parkman, 1972; Parkman, 1972; Parkman & Groen, 1971). The problem difficulty effect, also commonly referred to as the problem size effect, is defined by the positive relationship between response times and error rates and the size of the operands in arithmetic problems. Specifically, as the magnitude of problem operands increases, response times and error rates increase. It has been demonstrated with all four arithmetic operations: addition (Ashcraft, 1982; Ashcraft & Battaglia, 1978; Parkman & Groen, 1971), subtraction (Mirsky, 1993; Siegler & Shrager, 1984; Woods, Resnick, & Groen, 1975), multiplication (Geary, Widaman & Little, 1986; Miller, Perlmutter & Keating, 1984; Stazyk, Ashcraft & Hamann, 1982) and division (Campbell as cited in Ashcraft, 1992). The problem difficulty effect is observed in both children (Hamann & Ashcraft, 1985; Svenson, 1975) and

adults (Campbell & Graham, 1985). Further and importantly, it is also observed in both verification and production paradigms (Miller, Perlmutter & Keating, 1984; Zbrodoff & Logan, 1990).

One consistent violation of the problem difficulty effect is found in addition problems in which the addends are equal (e.g., $2+2$) (Ashcraft, 1992; Groen & Parkman, 1972; Svenson, 1975; Widaman, Geary, Cormier & Little, 1989). It is characteristic for ties, as they are called, to have uniformly short latencies, with no linear trend reflecting increases in the magnitudes of the addends (Groen & Parkman, 1972). In multiplication, a ties effect has been demonstrated but its form is different from that for addition; e.g., the observed increases are considerably slower than the increases associated with unequal multipliers (e.g., 4×7) (Parkman, 1972; Stazyk et al., 1982). Miller et al., (1984), however, found that response times for ties in *both* addition and multiplication increased more slowly with the size of the problems than for non-ties using a production task.

Split Effects.

As the absolute difference between two numbers increases, the length of time to make a comparative judgment between them decreases. This "split" (distance) effect is a general finding in a variety of task and stimulus situations and there is consensus that it occurs at the decision stage. In verification tasks, it has been shown that response times monotonically decrease as a function of increases between correct and incorrect answers to arithmetic problems (Ashcraft & Battaglia, 1978; Stazyk et al., 1982). In the comparative judgment literature, the numerical split effect is more

commonly referred to as a distance effect, paralleling the ubiquitous distance effect in psychophysical comparisons (e.g., Cattell, 1902; Henmon, 1906; Johnson 1939; Kellog, 1931; Münsterberg, 1894). More recently, in studies where the comparative judgment requires information maintained in long-term memory (e.g., Moyer, 1973), the effect is especially pronounced. Interestingly, split effects with comparisons of numerical magnitude much more nearly resemble those found with perceptual stimuli than those occurring with symbolic remembered stimuli (see e.g., Banks et al., 1976; Dehaene, 1989; Dehaene et al., 1990; Link, 1990; Moyer & Landauer, 1967, 1976; Poltrock, 1989; Restle, 1970).

Split effects found in numerical verification tasks are tentatively explained by associative network models of mental arithmetic as due to a "near neighbour effect" (Ashcraft, 1992, Campbell & Clark, 1989). That is, the difference in activation levels between two nodes that are in close proximity or highly associated will be more difficult to distinguish than nodes that are more distant or less highly associated. Campbell and Graham (1985) view this effect to be the result of interference from other members of an activated set of "candidate answers". The generality of the split effect lends strong supporting evidence for analog models. The more different two magnitude representations are, the less time it takes to distinguish between them on the basis of similarity or relative magnitude. All of the available models assume that arithmetic fact retrieval and comparison processes occur in an analog context (Gallistel & Gelman, 1992; Moyer & Landauer, 1967; Restle, 1970;).

Lesser Known Effects In Mental Arithmetic

Addend Order Effects

The effect of order of operands on response times has received rather negligible attention. The reason for this oversight is that only infrequently has such an effect been found to be statistically reliable. Parkman and Groen (1971), using a verification paradigm to look at addition and comparison of numbers, reported a significant effect of the order of addends on response times. They found that of 44 pairs of problems, 29 resulted in longer response times when the smaller of the two addends was presented on the left of the plus sign in a horizontally presented problem. For example, the response time for $4+2=6$ would be shorter than for $2+4=6$, although semantically the two problems are the same. The mean size of this addend order effect was 12 ms. Assuming left to right encoding of the stimulus pattern, Parkman and Groen conjectured that it might be easier for participants to process problems in which the larger addend was encountered first. This finding, they said was consistent with one of the counting models of arithmetic being scrutinized at the time (Parkman & Groen, 1971). Parkman (1972) and Stazyk and Ashcraft (1982) found no such operand order effect for multiplication or addition. Campbell and Graham (1985), however, reported an operand order effect for multiplication in adults although they did not elaborate or discuss possible reasons for this effect.

Effects of Instruction.

In comparison tasks, response times with the instruction "larger" are characteristically shorter than with the instruction "smaller" (Parkman, 1971). On the

other hand in mental arithmetic experiments in which a verification task is commonly used, individuals have only been asked to indicate the larger value or whether the stated answer to the problem is true or false or is the same as that obtained by the participant (Ashcraft, Fierman and Bartolotta, 1984; Ashcraft and Stazyk, 1981; Miller et al., 1984; Sekuler, Rubin & Armstrong, 1971; Zbrodoff and Logan, 1986, 1990) . As a result, the possible effects of instruction in the context of mental arithmetic are unknown.

Semantic congruity effects.

In comparative judgments with numerical magnitudes, response times are faster with small numbers with the instruction "smaller" than with the instruction "larger". Conversely, with relatively large numbers, response times are faster with the instruction "larger". This effect, referred to as the *semantic congruity effect*, was first explored with numerical magnitudes by Banks, Fujii, & Kayra-Stuart (1976). Dehaene (1989), working with double digit comparisons, also found reliable semantic congruity effects and showed that the reference-point model developed in Jamieson and Petrusic (1975) provided an excellent quantitative account of the data, contrary to the semantic coding model developed by Banks et al., (1976). As with the split effect, a very large literature has developed examining semantic congruity effects in both perception and in memory. Petrusic (1992) summarizes the main properties of these effects with perceptual comparisons and Petrusic, Baranski and Aubin (1995) extend these analyses to remembered comparisons. The present series of experiments was also designed to pursue the question of semantic congruity in context of the adding and

comparing numbers.

**Analogue Models of Mental Arithmetic and a Preverbal system of
Counting and Arithmetic Reasoning**

The earliest models of mental arithmetic were analogue models. Moyer and Landauer (1967) observed that in numerical comparison tasks, response times were inversely related to the distance between the digits being compared. That is, as the distance between the digits increased, response times decreased. Their data most closely fit Welford's (1960) general model for discrimination response times for differences between perceptual stimuli. The ratio of the numerical stimuli was more predictive of the response times than the absolute difference between them. As a result, in their proposed model numbers are internally represented and operated on as analog magnitudes and numerical comparisons are carried out in the same way as comparisons between perceptual stimuli. Restle (1970) obtained data that led him to propose that an analog process occurred for addition through the use of an internal number line and the concatenation of line lengths equal to the magnitudes represented by the digits. Notably, the paradigm used in his experiment is the same as that used in the present study, except that Restle asked individuals to indicate only the larger magnitude. That is, they had to determine which was larger the sum of the addition problem or the number presented.

Gallistel and Gelman (1992) more recently proposed a preverbal system of counting and arithmetic reasoning. Their model postulates that numerosities are internally represented as preverbal analogue magnitudes. It is assumed that operations

can be performed on these magnitudes in a manner isomorphic to arithmetic operations (for a comprehensive review of supporting evidence from animal cognition, developmental and neuropsychological research, the reader is referred to Gallistel & Gelman, 1992). Performance of simple arithmetic fact retrieval is explained via the mapping of verbal and written numbers directly onto preverbal magnitudes. The form of these magnitudes is hypothesized to be analogue in nature and Gallistel and Gelman liken their ordering to that of the various bars found in a histogram. The height of a bar represents the magnitude of the digit. Addition is achieved by concatenation of bars, whereas multiplication requires a more complex manipulation. Number comparison tasks are merely a matter of determining the taller bar. Gallistel and Gelman's histogram arithmetic is reminiscent of Restle's (1970) model.

Early Counting Models

Counting models of mental arithmetic, which have also been referred to as digital models (Miller, Perlmutter & Keating, 1984) arose in opposition to the idea that calculations were carried out using analog representations of numbers (Restle, 1970). As Suppes and Groen (1967) and Groen (1967, as cited in Parkman & Groen, 1971) studied addition algorithms used by children, they were led by their data to propose the notion of an internal counter initially set at the value of the maximum addend and incremented by single units until the answer to a particular problem was obtained eliciting a response. The latency pattern observed by these researchers motivated the design of the model they proposed. Specifically, increases in the magnitude of problem variables are positively correlated with response times. The idea that an

internal counter could be generalized to explain response time data obtained in number comparison tasks (Moyer & Landauer, 1967; Parkman, 1971) lent validity to the model.

Using a verification task, Parkman and Groen (1971) examined the addition and comparison of numbers. They replicated earlier findings of Restle (1970) with regards to the inverse relationship between response times and the addend split. However, Parkman and Groen also examined the potential relevance of other structural variables in addition problems on response times. Following Suppes and Groen (1967), the fit of observed response times to a linear regression line of the form $\text{response time} = a + bx$ was used. Five structural variables (i.e., x) were examined: the minimum addend, maximum addend, sum ($m+n$), addend split, and the $\log(m+n)$, where m and n represent problem addends. These particular variables were investigated because they had each been postulated by earlier researchers as providing the best account of response time data for addition. Out of those five structural variables entered separately into simple linear regression analysis, Parkman and Groen (1971) found that the sum and minimum addend provided the best fitting models to the data. However, the data in Parkman and Groen's (1971) study was subjected to simple linear regression analyses rather than analysis using multiple linear regression. The separate R^2 values obtained for the sum and minimum addend were .712 and .730 respectively. The correlation however, between the min and sum was .867, suggesting considerable overlap in the proportion of variance in response times that each variable explains. Although Parkman and Groen recognized the

dilemma, no satisfactory decision was reached concerning the unique proportion of variance that each variable was able to explain.

Parkman and Groen (1971) also observed the ties effect on response times. Response times for problems with equal addends did not differ as a function of the size of their addends. Their hypothesis regarding this effect was the first in the history of cognitive theories of mental arithmetic to suggest arithmetic fact retrieval through rapid direct memory access. Groen and Parkman (1972) in comparing data obtained from children and adults, proposed a two-process model for addition in which adult performance relied predominantly on the use of fast access memory. They estimated that 5% of the time adults failed to retrieve the answer to a problem from memory and resorted to the same counting method hypothesized to account for children's addition performance. The problem difficulty effect observed in adults was explained per this hypothesis. Multiplication data that Parkman (1972) obtained from adults compared favorably to addition data available at the time. This he interpreted as evidence suggesting common processes underlying both multiplication and addition.

The demonstration of the similarity between addition and multiplication together with the implausibility of a counting model explaining multiplication performance turned theoretical thinking towards exploring the feasibility of a memory retrieval model. Groen and Parkman (1972) had suggested that response times could alternatively be interpreted as reflecting the time required to calculate an index (memory location) to be used in retrieving the appropriate number fact. This set the direction for the research to follow.

Associative Network Models

Developmental research has revealed that during the acquisition of basic number concepts and arithmetic facts, children initially use procedural knowledge (i.e. learned rules) and counting methods which eventually give way to faster, direct memory retrieval from an organized network representation of arithmetic facts. (Ashcraft 1982, 1992; Campbell & Graham, 1985; Siegler & Shrager, 1984; Svenson, 1975).

The neural network theories which have been proposed so far, make three assumptions. First, arithmetic facts are stored in memory as an interrelated network of associated nodes. Second, retrieval is accomplished by means of spreading activation from the separate nodes representing each of the operands (parent nodes) along associative pathways to the node representing the answer. Third, if the level of nodal activation achieves a preset criterion threshold, the answer is retrieved and a response is made. Response time and error probability are assumed to be functions of associative strengths between the operands and the answer. Stronger associative connections are associated with faster response times and lower probabilities of error than weaker associations of lesser strength (Ashcraft, 1982, 1992; Campbell & Graham, 1985).

A variety of associative network models have been proposed over the years in order to better account for various response time and error rate patterns obtained within production and verification task paradigms. Generally, they have evolved from the more rigid table look-up models (Ashcraft & Battaglia, 1978) to the more recent and more flexible distribution of associations model of Siegler (1988; Siegler &

Shrager, 1984) and the network interference model proposed by Campbell (1987a, 1987b, 1990; Campbell & Oliphant, 1992). Each of these will be described briefly.

At the time of Groen and Parkman's (1972) work, the area of numerical cognition had not yet been extensively explored. Experimental work had been done concerning the psychophysical representation of number (Dehaene, 1989; Dehaene, Bossini, & Giraux, 1993; Dehaene, Dupoux, & Mehler, 1990; Rule, 1969; Shepard, Kilpatrick, & Cunningham, 1975) and the processes involved in making judgments of numerical magnitude (Banks, Fujii, & Kayra-Stuart, 1976) but proposals had not yet been made as to the structural representation in memory of information for addition, or the processes involved in the retrieval of such information. The first explicit model addressing these issues was formulated by Ashcraft and Battaglia (1978). They objected to Groen and Parkman's (1972) min model and direct access retrieval explanations for adult performance, first on the basis that the idea of adults still counting after years of practice was counterintuitive, and second, the prediction that retrieval via direct access required a constant amount of time (see also Ashcraft, 1985). Alternatively, Ashcraft and Battaglia examined the viability of the spreading activation theory of semantic processing (Collins & Loftus, 1975) which had achieved considerable popularity within the cognitive community at the time, since the representation of addition information in memory could "parsimoniously be considered analogous with semantic representations" (p.529).

Indeed, Ashcraft and Battaglia (1978) found the correct sum squared to be the best predictor of response times for addition and interpreted it as an index of search time

for retrieval of stored arithmetic facts from memory and evidence against a strict reconstructive counting process hypothesis. Linear effects of problem size on response times were *not* observed for ties, problems with equal addends (e.g., $2+2$), indicating the noninvolvement of counting processes in performance. Further and importantly, significant priming or facilitation effects comparable to those obtained in semantic memory retrieval studies (for a review see Collins & Loftus, 1975) were observed. Response times were significantly faster on those trials which were immediate repetitions to the previous trial, again suggesting a network representation of arithmetic knowledge in memory.

The original memory network retrieval model for addition proposed by Ashcraft and Battaglia (1978) is functionally organized like a printed table with a (0,0) origin, entry nodes (parent nodes) representing addends 0-9 are located on two adjacent sides. Nodes representing sums of various single digit combinations lie at the intersections of rows and columns headed by the parent nodes. The mechanism permitting access and retrieval of information from the table is a process of spreading activation similar to that which Collins and Loftus (1975) proposed for retrieval of semantic information from memory. Ashcraft and Battaglia (1978) interpreted response times to reflect the time required for the intersection of separate searches each beginning at the appropriate entry nodes down the rows and across the columns. For example, for the problem $4+2$ the search would begin with the activation of the entry nodes 4 and 2, spreading down the column and across the row intersecting at the node representing the sum (6) and retrieval would occur.

Contrary to Parkman's (1972) finding of the sum as a better predictor of response times than the minimum addend which implied retrieval from a square look-up table of arithmetic facts, Ashcraft and Battaglia's (1978) results indicated the correct sum squared as a better predictor of response times and that the relationship between response times and the minimum addend was an exponentially rather than linearly increasing function. Consequently, they reasoned that rather than being square, the network must either be stretched in the area of larger sums or that distances between entry nodes increase with the magnitude of the numbers they represent. Although there was some lively debate over the metaphorical shape of the retrieval network, its importance has waned relative to issues such as its development and determinants (Ashcraft, 1985, 1987; Baroody, 1984, 1985; Campbell & Oliphant, 1992; Campbell & Graham, 1985; Geary, Widaman & Little, 1986; Hamann & Ashcraft, 1985).

The mechanism of retrieval from Ashcraft and Battaglia's (1978) network representation was by means of spreading activation along the interconnections amongst the nodes. Lengths of the interconnections were functions of the strengths of relatedness or association between nodes. Reference to "distances" between nodes was a metaphorical expression for associative strengths amongst nodes. Response times and error probabilities were hypothesized to reflect accessibility and location of relevant information in the network. For example, the observation that response times for small problems (e.g., $1+2$) are characteristically faster than response times for large problems ($8+9$) was explained as due to shorter distances and stronger associative strengths between entry and sum nodes for small problems than for large

problems. Such differential problem strengths in long term memory were hypothesized to be functions of early learning, frequency of occurrence, and practice (Ashcraft, 1982, 1992). Considerable support for this hypothesis was obtained in Ashcraft's later developmental research (Ashcraft, 1982, 1990; Ashcraft, Fierman & Bartolotta, 1984; Hamann & Ashcraft, 1985) as well as Siegler (1988: Siegler & Shrager, 1984) and Campbell and Graham (1985).

One further result from Ashcraft and Battaglia's (1978) study provided unanticipated support for a network retrieval explanation for adult mental arithmetic performance. It happened that some of the false addition problems coincidentally correct under subtraction (e.g., $1+8=7$) were associated with abnormally long response times. Ashcraft and Battaglia proposed that on such trials information held in memory for subtraction had interfered with Ss being able to reject a problem more quickly and "the presence of associations among problems or, stated more generally interconnections among memory nodes is a hallmark assumption of network models" (Ashcraft, 1987). Winkleman and Schmidt (1974) found confirming evidence for associations between operation-specific information in memory. Using addition and multiplication problems in a verification task, Winkleman and Schmidt found their participants' response times to be inordinately slow for false problems that would have been correct for the other operation. For example, the false problem $4 \times 5 = 9$ would have been correct had the operation been addition, whereas the false problem $4 + 5 = 20$ would have been correct as a multiplication problem. Thinking that perhaps using both operations had led to individuals misreading the problem's operator on

various trials during the experiment, Stazyk, Ashcraft and Hamann (1985) used only multiplication. Confusion problems were set up such that some of the stated answers to false problems presented in their verification task contained multiples of one of the multipliers (e.g., $4 \times 8 = 24$). The answers to the remaining false problems were not multiplicatively related. Response times were found to be significantly slower for the confusion problems.

In tandem with the idea of spreading activation is the notion of decay. This idea allows predictions concerning priming as a function of lag between the prime and the target. Indeed, positive and negative priming effects were evident in the Stazyk and Ashcraft (1981) data, supporting the notion of spreading activation an inherent aspect of semantic network retrieval models. This notion was later pursued by Campbell and contributed significantly to his "network interference model" describing mental arithmetic performance.

Surprisingly, response times for extreme splits were faster than response times for true problems. Stazyk and Ashcraft (1981) suggested that perhaps a global evaluation stage co-occurs with retrieval. This finding was later used by Zbrodoff & Logan (1990) in their argument that verification tasks and production tasks are not comparable.

Distribution of Associations Model

In Siegler's model, differential strengths of associations develop over time between operands and the correct and incorrect answers given by an individual (Siegler, 1988). Siegler hypothesized that when a problem is presented, all answers that have ever been associated with that problem are activated differentially according to associative strength. The strengths of associative links is postulated to be the result of early learning and the frequency with which different problems are incurred (Siegler & Shrager, 1984). The flatter the distribution of associative strengths the more difficult it is for the individual to determine the most highly activated node representing the answer and the less likely that any one answer will reach the response criterion. The reason for this difficulty is due to the fact that there is no one candidate answer that "ticks out". Therefore, response times and error rates are assumed to reflect the peakedness of associative distributions. (Siegler & Shrager, 1984). Recent findings infer that interrelatedness exists not only amongst arithmetic facts but also between addition and multiplication operations (LeFevre, Bisanz & Mrkonjic, 1988, Zbrodoff & Logan, 1986). Siegler's model includes the use of processing strategies and the idea that over the course of development children make a transition from counting-based strategies not unlike those proposed by Parkman and Groen (1971; Groen and Parkman, 1972) to a faster and easier retrieval strategy. On the basis of their findings, Siegler and Shrager (1984) postulated that in the case of problems with peaked distributions, retrieval is the strategy of choice. Converging evidence from Siegler (1988) and Koshmider and Ashcraft (cited in Ashcraft, 1990)

indicating that estimated associative strength values for multiplication of third and fourth grade children were equally predictive of response time performance across the developmental span into adulthood "suggests that the associative strengths of the basic facts, once established in early grade school, maintain their relative standings to one another across later education and experience" (p. 193, Ashcraft, 1990).

Network Interference Model

Another model concerning the structure and process of mental arithmetic is the network interference model proposed by Campbell (1992). The development of this model over the years has been based almost entirely upon error rate patterns and response times associated with those errors in the context of multiplication. An important difference between Campbell's approach and that of other associative network theorists is his contention that the relationship between structural variables and response times (e.g., min, max, sum, sum^2 , correct product) is artifactual. It is well documented in the literature that response time performance for tie problems in addition or multiplication do not conform to the general pattern observed for other problems. response times for ties are significantly shorter than for other problems and there is generally a very small, if any, relationship between response times and the magnitude of the operands for ties (Groen & Parkman, 1972; Miller et al., 1984; Parkman & Groen, 1971; Stazyk et al., 1982).

Campbell and Graham (1985) demonstrated other violations of the mean response time patterns predicted by structural variables through the analysis of data from Miller et al., (1984). Response time performance for the five times table was found to be 26

ms faster than for 3 times table, 14 ms faster than the 4 times table and about 1 ms faster than the 2 times table. In short, only the 4,8, and 9 times table performance conformed to the response time pattern predicted by structural variables. Although Miller et al., had suggested that the reason for these discrepancies might be due to differences in accessibility to problem answers in the associative network, it was Campbell and Graham who extended this idea to include the notion that such differences might be due to interference or inhibition from competing response candidates. They hypothesized that the degree of interference or inhibition was determined by differential associative strengths amongst related problems.

In examining the nature and acquisition of multiplication facts in children, Campbell and Graham (1985) found a developmental trend to error patterns. By grade 5, error patterns mirrored adult performance. Table related errors, incorrect responses that are multiplicatively related to at least one of the problems operands (e.g., $6 \times 4 = 30$), dominated other types of errors. These results were interpreted as demonstrating the development of an associative network in which retrieval of correct answers is subject to interference from false associations. Campbell (1990) reported that 91% of errors made by participants in a production task were answers to other multiplication tables and 88% of those same errors were multiplicatively related to one of the problem's operands.

Campbell (1987) demonstrated error priming effects which lent further support to a network interference model. Error priming effects are defined as the increase in probability that an incorrect response to a problem will be made due to the correct

retrieval of that same response on an earlier trial. In a series of three production experiments Campbell manipulated the presence or absence of an interference problem set whose answers typically occur as errors for other problems. The frequency of the occurrence of such errors was taken as evidence of false associations between particular problems and answers. He found that when the interference set was embedded in the target problem set, correct response times were longer and error rates were greater than in its absence. Campbell interpreted these error priming effects as reflecting increased associative interference due to residual activation from previously retrieved answers combined with activation due to false associations of problems with those answers.

Having established the existence of error priming, Campbell and Clark (1989) investigated the time course of error priming in number fact retrieval by observing frequencies with which errors matched previously generated answers and noting the number of intervening trials. By comparing error frequencies with those expected by chance, Campbell and Clark found that the probability of an error matching the answer from the immediately preceding answer was below chance. This effect was called negative error priming. Error matching probabilities were greater than those expected by chance however after one intervening trial. Maximum positive error priming effects as they were called occurred approximately 10 s after the generation of the answer causing interference and remain above baseline levels for at least 80 s. These findings allowed further refinement of Campbell's associative interference model. He postulated that about 1.5 s after an answer has been retrieved, inhibitory

influences are stronger than excitatory influences reducing its typical status as a candidate response for the subsequent trial below baseline. The inhibitory influence decays quickly however allowing positive error priming effects to manifest themselves. Since positive error priming effects emerge subsequent to negative error priming the two mechanisms were implicated in arithmetic fact retrieval.

Committed to the idea of inhibition and the idea that specific errors reflect associative strengths in the network, Campbell focused much of his research on error rate patterns and the conditions giving rise to particular types of errors and indeed found much corroborating evidence for his model. For example, Campbell and Graham (1985) observed in their study that the most common multiplication response errors of both children and adults were correct answers to other multiplication problems involving one of the incorrect problem's operands (i.e., a related times table problem).

The concept of interference or inhibitory mechanisms and Campbell's conviction of their mediating role in arithmetic performance drove Campbell's research and were the cornerstones upon which he developed his network interference model. Campbell's latest version of the network interference model is the result of several revisions made in order to accommodate the most recent evidence. The model describes structure and process of memory for arithmetic facts in addition and multiplication.

Problems are represented within the network as nodes, each comprised of a physical code and a spatial magnitude code. The physical code for each node is based

on a modality specific representation of the problem it represents. That is, the nature of the representation is related to phonological or visual features of the problem.

Campbell (1992) describes physical codes as "unitized structures of associated elements" (p. 2) meaning that the operands, operators and answers form an ordered sets of features or units. For example, the physical code for the addition problem $2+4=6$ would be described as [2 4; +; 6]. It is upon these units that we operate. In describing the spatial magnitude code associated with each problem node, Campbell draws on the commonly used analogy of an internal number line. Apparently, each problem in the network has associated with it an analogue representation of its numerical magnitude.

When a problem is presented, parallel activation of both addition and multiplication facts is supposed to occur by way of a feature matching process. Differential activation of physical codes amongst problem nodes is based on the degree of their featural and positional similarity to components of the presented problem. For each featural match, a weight is assigned based on the degree of the positional match. For instance, the presented problem $4 \times 8 = 32$ initially provides excitatory input to both addition and multiplication facts. Nodes having physical codes that include at least one of the operands 4 or 8, the answer 32 or one of its component numbers 3 or 2 will theoretically receive more activation than other nodes (e.g., [8 3; x; 24], [4 6; x; 24] as compared to [1 5; x; 5], [5 10; x; 50]). Further, common features in the same relative position such as the 4 in [4 6; x; 24] would be weighted more in terms of similarity than the 8 in [8 3; x; 24] because of its

positional match. Both of these physical codes would receive a similarity weighting for the match of 4 in the answer 24, but the match that would cause [8 3; x; 24] to be relatively more activated would be the match between the 3 in presented problem's answer $4 \times 8 = 32$ and the 3 in the latter physical code. In other words, the similarity between a physical code and the presented problem is the sum of the weights for featural and positional matches. Additionally, an incremental weight is added for operator matches acting to determine whether addition or multiplication facts are more activated.

Although Campbell has built spatial magnitude codes into his network interference model, their relevance is unclear. On the one hand, he suggests that magnitude similarity between the magnitude code and the magnitude of the presented problem plays a role in the differential activation of nodes. On the other hand, he also states that a given operand pair produces equal activation of magnitude codes associated with their sum and their product, resulting in equivalent similarity values regardless of the operation involved in the presented problem. On the basis of the latter statement, Campbell concludes that differential retrieval of arithmetic facts is due "entirely to differences associated with physical-code similarity" (p.3).

While Campbell suggests that the spatial magnitude code is independent from the operation itself, the generation of the magnitude that the presented problem represents means that the operation must necessarily be taken into account. Retrieval of an answer from the network occurs after a criterion threshold level of activation is reached. This involves a series of processing cycles in which differential activation

amongst the network's nodes is calculated by the system according to the amount of excitation or inhibition exerted on individual nodes. The degree to which a specific node is excited is determined by its similarity to the presented problem, whereas the amount of inhibition it receives is relative to the total amount of activation amongst other activated nodes (Campbell 1992, Campbell & Oliphant, 1992). Campbell postulates that when the system reaches a state of equilibrium the processing cycles cease and the node that is most highly activated is selected as the response.

Dehaene's Triple Code Model

After an extensive review of the current literature, Dehaene (1992) proposed a model designed to reconcile the ongoing debate regarding psychophysical representation of number. In other words, what is the nature of the numerical representations that exist when doing calculations, comparing numerical magnitudes or making approximations of number? This model assumes three types of numerical abilities, each having a specific format in which numerical information is processed. Processing involved in addition and multiplication retrieval as well as counting sequences is hypothesized to be done on analogues of word sequences. The label "auditory-verbal" is applied to the learned lexicon of verbal associations or notation format (e.g. six/hundred) in which these types of processing are conducted.

Numerical comparison and approximation abilities require an analogue magnitude code which is thought to be part of a primitive preverbal system of arithmetic reasoning. Evidence of this system comes from animal cognition and developmental research. The ability to approximate and compare numerosity has been found in

animals and infants (Gallistel & Gelman, 1992). It is the last ability left in severely aphasic and acalculic individuals (Dehaene, 1992). In numerical comparison tasks, normal individuals' response times decrease as the absolute difference between the numbers being compared increases (Banks, Fujii & Kayra-Stuart, 1976; Dehaene, 1989; Link, 1990; Poltrock, 1989).

Finally, Dehaene proposed that an Arabic notation is used by digital manipulation abilities which require a "dedicated positional notation system" (p.34), such as in multi-digit calculation and parity judgment. Dehaene notes that this model must assume that it is possible to mentally translate from one notation to another depending on the demands of the task. A significant prediction of this modular theory is that translation from one notation to another takes a certain amount of time which can be experimentally measured.

Empirical Regularities Upon Which Models are Built

As mentioned previously, models of mental arithmetic have arisen out of theorists' interpretations of repeatedly demonstrated response time effects and error rate patterns in experiments examining performance in simple addition, multiplication or numerical comparison tasks. Although there are numerous effects, especially in multiplication data, the three most commonly encountered effects are problem difficulty, split and congruity effects. It therefore seems relevant to review the extent to which aforementioned models are able to explain and predict these effects.

It is generally viewed by associative network theories that the problem difficulty effect reflects the duration and degree of retrieval difficulty (Ashcraft, 1992;

Dehaene, 1992). Ashcraft's (1992) model points to metaphorical distance within a network of associated nodes, with distance representing associative strength resulting from frequency of retrieval and practice. In addition to frequency and practice, Campbell's (1992; Campbell & Oliphant, 1992) model also incorporates original acquisition to explain the problem difficulty effect as representative differential interference from false associations. Analog models and Dehaene's (1992) triple code model account for problem difficulty effects by assuming Fechnerian-like compression of linear magnitudes. As well, differential strengths between associated nodes might also be assumed to be the explanation in the triple code model. Gallistel and Gelman (1992) however, are able to explain and predict the problem difficulty effect in terms of speed-accuracy trade-offs and the variability in mapping of verbal and written digits to preverbal magnitudes. According to their model, variability found in the mapping from the verbal domain to preverbal magnitudes is Weberian in nature. That is, variability in preverbal magnitudes increases proportionately with increasing numerosities. Therefore, problem difficulty effects are the result of the increased amount of time necessary to "achieve a mapping of acceptable reliability" (p.61). Although generally indistinguishable from the analogue view, their model makes the important prediction that the problem difficulty effect should be enhanced under speeded responding or early in development.

Response Time Methods in Mental Arithmetic

The most common technique used to infer the nature of the processes and the architecture of arithmetical knowledge has been through the analysis of response time

and error rate data obtained from production or verification tasks. Drawbacks have been associated with both of these techniques. One method often used in obtaining production times, for example, how long it takes to solve $5+3$, is voice keying, whereby the participant activates a voice key with a verbal response (e.g. Miller et al., 1984). Inherent difficulties in securing accurate response times using this method have been noted (Pachella, 1974; Restle, 1970). Groen and Parkman (1972) pointed out that extraneous effects are produced by recording mechanisms in production tasks, possibly masking "small, but systematic latency differences". The method proposed in the present series of experiments is an extension of Donder's subtractive-method which assumes independent stages of processing. It further assumes that it is possible to manipulate one stage without affecting the nature of processing at other stages (Donders, 1868/1969; Sternberg, 1969). If this method is valid, its use in the study of mental arithmetic might provide additional empirical insight regarding the psychophysical representation of numerical magnitudes.

Until recently, verification tasks have been assumed to be composed of two independent stages of processing; a production stage plus a comparison stage. Zbrodoff & Logan (1990) produced evidence indicating error in this assumption and suggest that the nature of a task may determine differences in processing. Further, verification tasks generally require a binary decision (i.e. same-different, right-wrong) from the participant. Actual computation of the arithmetic problem can be avoided because there are a number of strategies that can successfully be used. It is possible for individuals to respond correctly by comparing structural characteristics of

internally generated digits with those of digits proposed as answers, or by comparing whole equations against memory for "resonance" (Zbrodoff et al., 1990). There is no way of determining whether an individual has indeed solved or compared numerical quantities.

Another drawback to the binary same-different, yes-no, true-false responses of verification tasks is that the comparison aspect of these tasks is made implicit, precluding examination of properties of the comparison process. Campbell (1987) demonstrated that presented answers in verification tasks act as priming stimuli whereby true answers have facilitating effects and false answers having associative links to the presented problems interfere with both retrieval and accuracy. The result is a distortion or masking of effects related to retrieval processes, rendering the validity of interpretations based on such measures questionable. At the very least, Campbell points out that processing models based on evidence from verification procedures may not be generalizable to the processes of retrieval involved in arithmetic production performance. In consideration of the noted weakness in a verification task, the present study is designed to increase the likelihood that in those trials where addition problems occur answers will actually be retrieved.

Comparative Judgment, Addition and Additive Factors

A commonly used method for determining temporal characteristics of stages of processing that was adopted in the present study is Sternberg's (1969) additive-factors method which he developed as an extension of Donder's (1868/1969) subtractive method. It is hypothesized that if addition and comparison of numerical magnitudes

constitute two separate stages of processing, then it will be possible to estimate the length of time to access information for addition. As suggested initially by Münsterberg (1894), this could be done by subtracting response times for trials in which Ss made simple comparisons of magnitude between two single digits (e.g. 8 vs 5 or 4 vs 9) from response times for trials in which the identical magnitudes are compared but one magnitude is represented in the form of a simple subtraction problem (e.g. 8 vs 7-2 or 5-1 vs 9).

Münsterberg's method of investigation provides an opportunity to test the validity of using Sternberg's (1969) additive factors idea as a means to determine various characteristics of arithmetical knowledge. Sternberg's additive factors method is heavily relied upon in verification paradigms used in mental arithmetic research to date. As previously mentioned, the comparisons in these paradigms have been implicit and therefore examination of the characteristics of this "stage" especially in terms of its independence from the retrieval stage are especially relevant. Unlike verification tasks, the number comparison aspect of the paradigm that will be used in the present series of experiments is explicit. Participants will be required to do both mental arithmetic and quantitative comparisons between numbers. Importantly, by making the comparison explicit it is possible to examine the comparison process itself.

One of the basic underlying assumptions of Sternberg's additive factors idea is that it is possible to manipulate one stage of processing without altering or influencing the nature of processing at other stages. Various factors have main effects on single stages which influence overall length of processing in an additive fashion. This

assumption was tested by manipulating three independent variables well known for their effects in addition, multiplication and number comparison literature: (a) problem difficulty, (b) comparison "split" (the absolute distance between the numerical quantities to be compared), and (c) instruction. If production and comparison stages are independent then problem difficulty and comparison split would not interact in their influence on response times. Alternatively, an examination of the available literature in which comparisons with both remembered and perceptual stimuli are required reveals a clear violation of separate comparison and memory retrieval stages (e.g., Banks, 1977; Banks et al., 1982; Moyer & Bayer, 1976; Petrusic, Baranski & Aubin 1995). The main features of these experiments are now briefly reviewed with a view toward developing their theoretical implications for understanding both the memory and comparison processes in the context of mental arithmetic, specifically, addition.

Memory psychophysics: Facts and Theories

Typically, in experiments examining the differences and similarities in comparisons with perceived and with remembered stimuli (e.g., Banks, 1977; Banks et al., 1982; Moyer & Bayer, 1976; Petrusic, Baranski & Aubin, 1995) participants first learn to associate labels or names (CVCs) with perceptual magnitudes. In the second stage, comparisons are made strictly from memory (i.e., "which label corresponds to the larger circle?") for some participants (e.g., Moyer & Bayer, 1976) and other participants simply compare the perceptual stimuli. In other experiments (e.g., Banks et al., 1982; Petrusic, Baranski & Aubin, 1995), the two types of trials

are intermixed. Invariably, performance in the memory and perceptual conditions is errorless or nearly so. Two important and robust findings characterize these experiments: both types of comparisons exhibit distance and semantic congruity effects but these effects are greatly amplified with the very much slower comparisons with remembered stimuli. For example, congruity effects are approximately 20-30 ms with perceptual comparisons but they are 200-300 ms with the memory comparisons.

Taken, together, this configuration of findings has been taken as providing powerful support for the view that decision processing in perception and in memory is fundamentally different and that the comparisons are also based on different representations (e.g., Banks, 1977; Banks et al., 1982). Memory processing is based on propositionally-based semantic codes and perceptual comparisons are based on information given directly from sensory analyzer systems. However, in support of the analogue position developed originally in Moyer & Bayer (1976), Petrusic (1992) and Petrusic, Baranski & Aubin (1995) have argued, that to the contrary, heightened congruity and distance effects follow from the view that memory maintains analogue-like information and that this information is processed by a common evidence accrual based decisional system.

Evidence accrual theories of response time liken the overt decision process to that involved in the sequential testing of statistical hypotheses (e.g., Wald, 1945). First, all theories assume that the nervous system generates noisy internal representations of the magnitudes of each of the stimuli to be compared. Second, it is assumed that during each pre-decisional moment, the decisional system computes the difference

between the two representations momentarily generated by the nervous system. The task for the decisional system is to determine if the evidence over a series of such samples and their resultant differences provides greater evidence for the hypothesis that the differences sampled arise from a distribution with a positive mean (e.g., corresponding to the response alternative the larger stimulus magnitude is on the right of the screen) or for the hypothesis that the distribution of differences has a negative mean. Sample differences are accumulated until the threshold of sufficient evidence for one hypothesis over the other is reached and the response corresponding to that hypothesis is then emitted.

Thus, in most models (see Luce, 1986 for the underlying theoretical developments) the overt response time arises as a sum of independent and identically distributed random variables, each representing the duration of the generation of the neural representations of the magnitudes, the computation of their differences, and examination of the accumulated competing evidence strength relative to the criterion and an additional motor-output stage. Although alternative quantitative theories differ in detail, in terms of specific assumptions for the termination of the evidence accrual process (i.e., in terms of stopping rules), they all predict: (a) more accruals are required when the difference to be discriminated is smaller; (b) when the internal representations are noisier, as they are in memory, more accruals will also be required.

Thus, the specific effects observed with remembered comparisons are predicted on this analogue evidence accrual view because more accruals are required with the less

discriminable pairs (i.e., those adjacent on the magnitude continuum) and because more accruals are required with the noisier memory representations. As well, it may also be the case that each accrual moment is slower in memory than in perception (because more neural operations may be required to activate the memory representation than to process the sensory representation. Petrusic (1992) demonstrated that the semantic congruity effect arises on each accrual moment. Consequently, just as above, heightened congruity effects will be evident with remembered comparisons. With errorless perceptual comparisons, very few accruals are required and the congruity effect should be minimally evident, which is, in fact the case.

In the present context of the compound, add and compare task, studied in this thesis, the accrual notion makes a clear prediction. On the view that during each accrual moment, the memory operations involved in addition must be reactivated, the difference between the time to compare and add and compare only will increase with increasing decision difficulty and congruity will be enhanced in the add and compare condition. Finally, because the memory processes in addition must be reactivated on each accrual, the properties of the addition process will be most evident when decision difficulty is held constant and at the smallest difference (i.e., a split of 1).

The Psychophysics of Number Comparisons

Despite the extensive literature on the internal representation of number, there is as yet no consensus as to the nature of its internal representation (see Link, 1990, for a review). There is evidence in the number comparison literature, however, that

numbers display the same properties as perceptual and remembered stimuli in the context of prothetic (quantitative) continua (Petrusic et al., 1995). Moyer and Landauer (1967) demonstrated that response time and error rates decrease as a function of the distance between the numbers to be compared. Since this distance effect mirrors the effect found with comparisons with physical stimuli, Moyer and Landauer concluded that numbers are transformed into analogue magnitudes for comparison. This effect has since been found by others (Buckley & Gillman, 1974; Hinrichs, Yurko & Hu, 1981; Holyoak, 1978; Sekuler, Rubin & Armstrong, 1971).

Another well documented effect is the magnitude effect, where the time required to choose between two relatively small numbers (e.g., 1 vs. 2) is shorter than the time required to decide between two relatively large numbers (e.g., 8 vs. 9). Banks, Fujii and Kayra-Stuart (1976) suggested that this effect was due to the compressive nature of the subjective number scale. That is, numbers get subjectively closer together as one ascends the subjective number scale (Schneider, Parker, Ostrosky, Stein and Kanow, 1975). This interpretation has been further corroborated in the context of similarity experiments (Shepard, Kilpatrick & Cunningham, 1975).

Banks, Fujii and Kayra-Stuart (1976) also demonstrated the semantic congruity effect for numbers. In their second experiment, participants were required to choose the larger or smaller of two numbers. Results indicated that response times for choosing the larger of two relatively large digits (e.g., 8 vs. 9) were shorter than those for choosing the smaller digit of the same pair. Conversely, participants chose the smaller of two relatively small digits more quickly than the larger of the same

pair. Collectively, these effects, which suggest analogue-like properties of the internal representation of number impose constraints on models of mental arithmetic (Dehaene, Dupoux & Mehler, 1990).

Fundamental Issues

It is clear that the results of mental arithmetic and number comparison research have given rise to quite different interpretations as to the manner in which numerical magnitudes are manipulated. On the one hand, a strong argument for models of mental arithmetic proposing retrieval of arithmetic facts from a highly interconnected network of associated nodes through a process of spreading activation has emerged (Ashcraft 1992, Campbell 1992). However, these models are silent or unclear as to the internal representation of the facts. Since these facts have meaning in that they are representative of a magnitude quantity, it would be important to resolve the nature of that representation. On the other hand, number comparison research clearly addresses the issue of representation suggesting that comparing the magnitudes of numbers is done similarly to perceptual comparisons such as line lengths. By combining an arithmetic task with an explicit number comparison task, there is much that can be learned concerning the independence or interdependence of arithmetical and number comparison processes.

In a preliminary investigation, Mirsky (1993) applied Sternberg's (1969) additive factors method to Munsterberg's subtract and compare paradigm and found strong evidence suggesting non-independent stages of processing for subtraction and comparison of numbers. Participants were required to either compare two single digits

or to compare a single digit with a single digit subtraction problem presented without its answer. A clear interaction was found between the type of comparison and the distance between the magnitudes to be compared (split). Additivity requires that response times and error rates for comparisons involving subtraction, parallel those involving straight number comparisons at each comparison split. Further, the finding of a strong instruction by pair interaction (semantic congruity effects) indicates that in the context of arithmetic, the nature of the task changes the way in which numbers are processed. These findings challenge theories supporting the notion of modularity in numerical processing which suggests that the way in which numbers are processed is task specific. That this early study revealed an interaction between arithmetic fact retrieval and number comparison suggests that the story is much more complicated than has previously been assumed. Supporting this view was a serendipitous and perhaps revealing finding that subtracting by one was reliably faster than subtracting by any other single digit. This result suggests subtracting by one is qualitatively different than subtracting other single digits. We can no longer look at various aspects of number processing in isolation but must look at understanding the nature of the relationships between them. What does one aspect tell us about another, and how can this help us further discriminate between available models of numerical processing?

The current experiments were designed to provide a more rigorous test of the viability of the operate and compare paradigm as an alternative to production and verification paradigms. An open question was whether the psychophysical paradigm would carry the signature of a production task, a verification task, a number

comparison task, all or none? In the least, for the viability of the operate and compare paradigm to be supported, the replication of robust effects found using the traditional paradigms was a requirement (e.g., problem difficulty effects, magnitude effects, split effects, semantic congruity effects). Further, it was hoped that the increased precision in response time measurement through the use of explicit comparisons would reveal reliable results demonstrating effects only sporadically reported in the literature. Specifically, response times and error rates were examined for evidence of addend or presentation order effects. That is, was the position of the larger addend in the addition problem important or whether the number to be compared was on the left or on the right on operate and compare trials? The effect of instruction was also examined since, as explained earlier, individuals have traditionally been asked to determine the larger when comparing perceptual or remembered magnitudes.

Experiment 1 was designed primarily to replicate the major findings obtained for addition using production and verification tasks, such as problem difficulty and tie effects. Since semantic congruity and magnitude effects were found for subtraction (Mirsky, 1993) using the operate and compare paradigm, the presence of these effects were also examined for addition.

Experiment 2 was designed as an extension to Experiment 1 in order to examine the effects of comparison split, addend order at operand splits between one and three, as well as to investigate possible effects of adding by ones comparable to those found for subtraction.

Establishing the viability of the currently proposed paradigm would extend the methodological repertoire available to those individuals working in the domain of numerical cognition, perhaps providing a vehicle for the convergence of two large, but very separate literatures.

Experiment 1

Method

Subjects. Twenty-three undergraduate students with normal or corrected-to-normal vision enrolled in a first year introductory psychology course at Carleton University (Ontario, Canada) participated for course credit. Relevant demographic characteristics such as age, sex and handedness were recorded.

Apparatus. For this experiment, a 486DX IBM compatible computer was used. The computer was programmed to randomize trial presentation for instructions, stimulus pairs and positional order, record trial order, the instruction and order of stimuli on each trial, responses, error rates and response times for each individual. The accuracy of the computer's timing device was ± 1 ms. A Zenith colour video monitor was used to display instructions and stimuli in white on a black background. The monitor was located upon the central processing unit (CPU) and positioned on a table, such that the stimuli were viewed at eye level when the participant was seated at approximately 50 cm from the monitor. Each block of experimental trials was initiated by the pressing of any key on a keyboard positioned on the table directly in front of the CPU. Responses were indicated by the participant depressing one of two specified keys on a custom designed response panel located on the table next to the

keyboard on the side of the preferred hand.

Stimuli. Comparisons containing endpoints 0 or 20 were excluded to minimize response bias, since 0 would always be the smaller in any comparison and conversely 20 would always be the larger. Comparison split was held constant at a split of 1, the most difficult level of comparison. Further, all comparisons were presented in both orders and with both instructions.

The total stimulus set was made up of two combined subsets (see Appendix A-1). One stimulus subset included 36 simple number comparisons (i.e., digit vs. digit) each presented twice. After balancing for order and instruction the total number of simple comparison trials equalled 288. The second subset of stimuli included compound comparisons (i.e., digit vs. addition problem). Only single digit addition problems with an addend split (i.e., the difference between the addends) less than or equal to 2 (e.g. $8+8$, $8+7$, $8+6$) were used, keeping sums below 19. Non-tie addition problems were presented in both addend orders (e.g. $2+3$, $3+2$) for a sub-total of 60 trials. The total after balancing for order and instruction was 240. Each tie addition problem was presented twice to match the number of times comparable non-tie problems were presented. That is, the 18 tie problems were doubled for addend order and balanced for presentation order and instruction for a total of 148 trials. There were 368 compound comparisons (i.e., ties and non-tie addition) and 288 simple comparisons for a total of 656 experimental trials. Together with 16 practice trials, each participant was presented a total of 672 trials.

For each participant, the complete stimulus set was randomly generated and

divided into 6 blocks of 112 trials. The 16 practice trials were randomly selected with replacement for each participant.

Procedure. Each individual was tested individually in a quiet room dimly lit by a lamp with a 25 watt bulb behind the computer. The experimenter was not present during the experimental sessions which lasted approximately 1 hour and 20 minutes. After having read and signed an informed consent, participants were instructed as to the types of trials that would be encountered.

The experimental session began with the practice block of trials followed by the experimental trials. At the beginning of the session, a message appeared on the monitor screen informing participants that they were about to begin the practice phase. When the practice phase was completed, another message appeared advising the participant that the experimental phase was beginning.

Practice and experimental trials were of two types. On some trials two numbers in Arabic format appeared, and in others a single Arabic number appeared with a single digit addition problem (as described above). All trials began with the appearance of either the instruction "LARGER" or "SMALLER" at approximately the middle of the monitor screen. The numeric stimuli appeared 2 seconds later, 2 cm below the instruction and spaced horizontally 2 cm apart. In accordance with the particular instruction, it was the participant's task to determine the location of either the larger or smaller numeric quantity and respond by pressing the response panel key corresponding to the position of their choice. For example, if the instruction "SMALLER" was followed by the comparison "11 12", the correct response would

be have been to press the left response key. Alternatively, if the comparison contained a single number/addition combination such as "8+3 12", the correct response would also have been the left response panel key. The inter-trial interval (ITI) between trials was 2 s.

Participants were told that speed and accuracy were equally important. In order to minimize fatigue effects, participants were allowed rest periods between each of the 8 blocks. The length of these respites was controlled by the participants, who when ready to continue, began the next block of trials by pressing any key on the keyboard. At the end of the experiment, participants were given their experimental credits and debriefed.

RESULTS

Twenty-three participants (9 females and 14 males) between the ages of 17 and 25 years provided a total of 15,456 trials in Experiment 1. Trials with response times below 200 ms or exceeding 10,000 ms were treated as outliers and discarded prior to analysis. According to these criteria it was only necessary to remove 4 trials from the data, less than 0.001%, leaving 16,508 trials for further analysis. Of the remaining trials, there were 738 errors resulting in an error rate of 4.47%.

Repeated measures analyses of variance were used to examine all reported effects. The levels of significance reported for both Experiments 1 and 2 are reported based on the Greenhouse-Geisser criterion. The degrees of freedom reported are defined by the design of the respective experiment. All results of error rate analyses reported are based on analyses of arcsine transformed error means .

Effects of Comparison Type As expected and as shown by Figure 1, simple comparisons between two digits had the fastest response times followed by comparisons with tie addition problems and finally comparisons with addition problems with unequal addends. The respective mean response times were 993 ms, 1335 ms, and 1656 ms. In terms of differences, simple comparisons were 502 ms faster than compound comparison with ties and unequal addend addition problems combined. Further, compound comparisons involving equal addends were on average 320 ms faster than those involving unequal addends. The same response time patterns were observed for both correct and incorrect responses.

Differences in response times between the three comparison types were established as significant by a 3-way analysis of variance (ANOVA) with response time as the dependent variable and comparison type (simple, compound with tie problems and compound with addition problems with unequal addends), instruction (larger, smaller) and comparison pair (2 vs 3 through 17 vs 18) as independent variables. The main effect of comparison type was reliable [$F(2,44)=131.13$, $p<0.0001$]. Contrast analyses revealed not only a reliable difference between comparisons involving 2 digits (simple) and comparisons involving ties and unequal addition problems (compound) [$F(1,22)=246.56$, $p<0.0001$], but also a reliable difference between comparisons involving ties and comparisons involving unequal addends [$F(1,22)=51.79$, $p<0.0001$].

Replication of the ties effect found in experiments using verification and production paradigms (Ashcraft, 1992; Dehaene, 1992; Zbrodoff & Logan, 1990) is

Figure 1 Shows times to make simple comparisons (filled circles), compound comparisons involving ties (open triangles) and compound comparisons involving unequal addend addition (filled triangles) as a function of comparison pair. Tick labels for comparison pair refer to the minimum magnitude of each comparison pair (e.g., 2 is the minimum comparison magnitude in the pair 2 vs 3). The characteristic effect of tie problems being faster than addition problems at larger addend splits can be readily observed as can magnitude effects in the increases in response times with increases in the magnitudes of comparison pair.

evidence in favour of the viability of the operated and compare paradigm since it successfully reveals that the same arithmetic processes are at work when operating in conjunction with processes involving the comparison of numerical magnitudes.

To ensure that the difference between the two types of compound comparisons was not due to collapsing over the different unequal addend addition problems associated with each of the numbers in the comparison (e.g., $2+3$ vs 6 and 5 vs $4+2$), addition problems whose sums were not equal to the sums of tie problems (i.e. problems with addend split of 1) were removed from the data and the analysis was repeated. Figure 2 shows response times plotted as a function of comparison pair for each comparison type. The familiar pattern is evident and it is not surprising that reliable differences were found between three comparison types [$F(2, 44)=118.54$, $p<0.0001$]. The results of contrast analyses replicated the original analyses finding significant differences between simple and compound comparisons [$F(2,44)=237.29$, $p=<0.0001$] as well as between the two types of compound comparisons matched for common sum [$F(1,22)=44.53$, $p=0.0001$]. The nature of the differences suggests that ties are processed differently from addition problems with unequal addends. At the very least ties are merely processed more quickly irrespective of the task at hand.

Effect of Pair. It is evident from Figures 1 and 2 that response times for all comparison types increased as a function of the magnitude of the numbers being compared resulted in increased response times for both correct and error responses. Indeed, a reliable main effect of pair was found [$F(15,330)=28.80$, $p<0.0001$].

Before continuing it seems prudent to admit that one drawback in using the

operate and compare paradigm for addition, is that it is not possible to determine whether the magnitude effect is due to the magnitude of the operands involved in the addition or the magnitude of the numbers being compared. Although the response time - comparison pair function for simple comparisons appears to be relatively flat in Figures 1 and 2, a significant linear effect found for the simple comparison condition [$F(1,22)=13.51$, $p=0.0013$] suggests that there are in fact magnitude effects occurring. The increasing steepness of slopes for compound comparisons relative to that for simple comparisons suggests that problem difficulty acts in concert with comparison magnitude effects. The apparent interaction between comparison type and comparison pair was indeed statistically reliable [$F(30,660)=12.19$, $p<0.0001$].

Support for the presence of problem difficulty effects was provided by Mirsky (1993). Using the same paradigm in the context of subtraction, Mirsky (1993) was able to disentangle the problem difficulty effects from magnitude effects arising from comparison processes. By holding constant the comparison split and the split between the minuend and subtrahend, response times were shown to increase monotonically with the magnitude of the operands. For example, response times of 9-7 compared to 3 were longer than response times of 6-4 compared to 3. It would therefore seem reasonable to hypothesize that magnitude effects occurring in the present context of addition arise from two sources. One source being the robust problem difficulty effect noted in the cognitive arithmetic literature (Ashcraft, 1992; Dehaene, 1992) and the second source rising from the comparison of two numerical magnitudes (e.g., Link, 1990).

Figure 2 shows response times as a function of comparison pair for simple comparisons (filled circles), compound comparison involving ties (open triangles) and compound comparisons involving unequal addend additions (filled triangles). Unequal addition problems include only those problems whose sums are equal to the sums of tie addition problems (e.g., $3+3$ in the comparison pairs 5 vs. 6 and 6 vs. 7 is matched with $4+2$). Tick labels for comparison pair refer to the minimum magnitude of each comparison pair (e.g., 2 is the minimum comparison magnitude in the pair 2 vs 3).

Figure 3 shows how error rates vary as a function of comparison pair for each comparison type. Note their similarity to response time patterns as shown in Figures 1 and 2. Tick labels for comparison pair refer to the minimum magnitude of each comparison pair (e.g., 2 is the minimum comparison magnitude in the pair 2 vs 3).

In order to further characterize the nature of the interaction between comparison type and comparison pair, trend analyses were conducted for each comparison type with pair entered as the quantitative variable. Clear and significant linear trends were observed for simple comparisons [$F(1,22)=13.51$, $p=0.0013$], compound comparisons with tie problems [$F(1,22)=39.91$, $p<0.0001$], as well as for compound comparisons involving unequal addends [$F(1,22)=39.91$, $p<0.0001$]. While comparisons involving ties exhibited a significant quadratic trend [$F(1,22)=31.57$, $p<0.0001$] compound comparisons involving unequal addends exhibited a significant cubic trend [$F(1,22)=4.82$, $p=0.0391$].

An analysis of errors rates using an arcsine transformation also provided support for the findings related to the effects of comparison type and comparison pair on response times. Main effects were found for comparison type [$F(2,44)=8.51$, $p=0.0031$] and pair [$F(15,330)=3.30$, $p=0.0030$] in addition to a significant comparison type by comparison pair interaction [$F(30,660)=1.90$, $p=0.0478$]. Figure 3 shows, although noisy, the similarity of error rate patterns to the response patterns seen in Figure 1, that is faster response times are generally associated with lower error rates.

Effects of Instruction. One result that is particularly relevant to research in the area of mathematical cognition is the effect of instruction. Researchers in this area using a verification paradigm to examine the characteristics of the processes involved in mental arithmetic have either asked whether the answer proposed by the experimenter is the same or different to the answer obtained by the participant or

which answer is larger. The results of Experiment 1 indicate that in the latter case, the instruction "Which is smaller?" should also have been asked. In Figure 4 response times (left panel) and error rates (right panel) are plotted as a function of comparison type for each instruction. Overall, response times for the instruction "larger" were only 39 ms faster than response times for the instruction "smaller". Although a small difference, an analysis of variance with comparison type, instruction and comparison pair as the independent variables showed that it was significant [$F(1,22)=10.67$, $p<0.0035$]. Looking more closely at Figure 4, differences in response times between the two instructions vary according to comparison type. Indeed, the 2-way interaction between instruction and comparison type was significant [$F(30, 660)=12.19$, $p<0.0001$]. Simple effects analyses showed that the effect of instruction was significant for compound comparisons involving ties [$F(1,22)=5.80$, $p=0.0248$] and with compound comparisons involving unequal addends [$F(1,22)=12.57$, $p=0.0018$] but not for simple comparisons; in each case, with faster responding with the instruction "larger".

An analysis of error rates using an arcsine transformation of proportion correct responses found a reliable instruction by comparison type interaction [$F(2,44)=3.47$, $p<0.05$]. As can be seen in the right panel of Figure 4 and determined by an analysis of simple effects, the effect of instruction was significant only for comparisons involving unequal addends [$F(1,22)=4.77$, $p<0.05$].

Effect of Decade. In order to investigate the effects of decade on comparison

Figure 4. Response times (left panel) and probability of error (right panel) plotted as a function of comparison type for each instruction. Tick labels for comparison pair refer to the minimum magnitude of each comparison pair (e.g., 2 is the minimum comparison magnitude in the pair 2 vs 3).

Figure 5. Response times and error rates plotted as functions of unit steps within a decade. Note the additivity of response time patterns for compound comparisons involving unequal addends between decades.

type, comparison pair and instruction the comparison pairs for their respective positions were matched within each decade. For example, the comparison pair 2 vs 3 was matched with comparison pair 12 vs 13 and as they are the second unit step within each of their respective decades if it is assumed that 0 and 10 mark the beginning of the first and second decades, respectively. Endpoints 1 and 20 were not included in the design of the experiment in order to minimize the potential for endpoint strategies. Consequently, there was no match for pairs 10 vs 11 or 11 vs 12. Also since the largest addition problem involving unequal addends was $9+8$, no matches were available for pairs 8 vs 9 and 9 vs 10. In total, six comparison pairs were obtained representing the unit steps 2-3 to 7-8.

Mean response time and error rate patterns are plotted in Figure 5 as a function of unit step within a decade separately for the first decade (1-10) and second decade (10-20). Generally, it appears from the plots that response times increase with each unit step within each decade and that response times are longer for simple and compound comparisons in the second decade compared to the first decade. Although noisier, error rate patterns (lower panels) seem to be similar to response time patterns. An analysis of variance was conducted with response time as the dependent variable and comparison type, instruction, decade and unit steps within a decade as the independent variables. As Figure 5 suggests, significant main effects were found for comparison type [$F(2,44)=132.71$, $p<0.0001$], decade [$F(1,22)=53.89$, $p<0.0001$], and unit steps within a decade [$F(5, 110)=14.45$, $p<0.0001$]. Reliable interactions were found between comparison type and decade [$F(2,44)=27.06$,

$p < 0.0001$], as well as decade and unit steps within a decade [$F(5, 10) = 3.96$, $p = 0.0055$]. The two interactions involving comparison type, decade and unit steps with a decade were qualified by a significant 3-way interaction between these three factors [$F(10, 220) = 3.96$, $p = .0026$].

Figure 5 suggests that the 3-way interaction can be accounted for by change in the interaction between decade and pair across comparison types.

It can be seen in that for simple comparisons (far upper left and lower left panels) both response times and error rates behave similarly within a decade irrespective of the magnitudes being compared. Response times increase slightly towards the upper end of each decade and simple comparisons in the second decade are 46 ms slower overall than comparisons in the first decade. Simple effects analyses confirmed the reliability of these observations with simple main effects for decade [$F(1, 22) = 11.21$, $p = .002$] and unit steps within a decade [$F(5, 110) = 9.17$, $p < 0.0001$]. These results were taken as supportive evidence for the presence of magnitude effects during the processing involved in comparing two numerical magnitudes in the operate and compare paradigm.

Viewing the upper middle panel of Figure 5, it is apparent that response times for compound comparisons involving ties in the second decade are longer than comparisons in the first decade. The differences however seem to vary according to the particular unit step at which the comparison is being made. The interaction between decade and unit steps within a decade was reliable [$F(5, 110) = 8.52$, $p < 0.0001$]. Further analyses showed that although significant, the interaction appears

to be due to noise and cannot be interpreted further.

The most interesting aspect of the comparison type by decade by unit step within a decade interaction is seen in the far right panel of Figure 5. Similar to simple comparisons and compound comparisons involving ties response times increase with increases in unit steps within both decades. Unlike, the other two comparison types however, a strikingly clear additive effect of decade is evident for compound comparisons involving unequal addends. The apparent simple main effect of decade was verified as reliable by the simple effects analysis [$F(1,22)=38.30$, $p<0.0001$] as was the simple main effect of unit step [$F(5,110)=6.66$, $p=0.0005$]. Most importantly, there was not a significant simple interaction between decade and unit steps indicating additivity. While the response time patterns within each decade reflect similar difficulty effects, an additional constant amount of processing time processing was required to perform comparisons in the second decade. While this additive effect was due to the inclusion of values greater than ten, it is uncertain as to whether the source of this additivity was the addition or the comparison process.

Error rate analysis using arcsine transformation showed a similar pattern of results thus supporting the findings of the response time analysis (see Figure 5). Overall, significantly more errors were made when comparisons involved the second decade [$F(1,22)=11.61$, $p=0.0025$] and generally increased with increases in the magnitudes being compared [$F(5,110)=2.88$, $p=0.0383$] again reflecting problem difficulty effects, magnitude effects in the comparison process or both. As with response times, these main effects of decade and unit step were qualified by a 2-way

interaction of decade and pair [$F(5,110)=2.62$, $p=0.0408$] and a 3-way interaction between comparison type, decade and pair [$F(10,220)=2.48$, $p=0.0413$] but these effects were not readily interpretable.

Semantic Congruity Effects. Semantic congruity describes a particular form of an instruction by pair interaction where response times for the selection of smaller of two relatively small numbers are faster than response times for the selection of the larger of the same pair. Conversely, response times for the larger of two relatively large numbers are faster than response times for the smaller of the same pair. Figure 6 contrasts the results of Mirsky (1993) (left panel) with those found in the current experiment (right panel) for simple and compound comparisons. When the arithmetic operation was subtraction semantic congruity effects (expressed as mean response times with the instruction "smaller" subtracted from mean response times for the instruction "larger") were greatly enhanced in the subtract and compare condition relative to the compare only condition (simple). These effects mirrored enhanced congruity effects found in comparisons with remembered stimuli relative to comparisons with strictly perceptual stimuli (Petrusic, Baranski & Aubin, 1995). As can be seen in Figure 6, (right panel), the present experiment involving addition is striking in its complete absence of congruity effects across the two decades. In the analysis examining effects within decades, a decade by instruction interaction approached significance [$F(1,22)=3.40$, $p=0.07$]. Examination of the far right panel of Figure 7 suggests a funnel-like congruity effect in the second decade. The funnel effect typically results from a markedness effect where faster response times for the

Figure 6. Congruity index (mean response times with the instruction "smaller" subtracted from mean response times with the instruction "larger") for the adjacent (one-step) comparisons for simple comparisons (filled triangles) and compound comparisons (filled circles). The left panel shows enhanced congruity effects in the context of a subtract and compare task (Mirsky, 1993), while the right panel reveals the striking absence of congruity effects in the context of addition.

Figure 7. Shows response time plotted as a function of comparison pair for both instructions ("larger", "smaller") for each comparison type. No congruity effects are manifest across the three comparison types although in the case of comparisons involving unequal addend addition there is some suggestion of funnel-like congruity effects in the second decade. Tick labels for comparison pair refer to the minimum magnitude of each comparison pair (e.g., 2 is the minimum comparison magnitude in the pair 2 vs 3).

unmarked form of the comparative (i.e., "larger") reduce and or eliminate congruity for the pairs at the lower end of the number continuum (see Petrusic, 1992). While no congruity effect is evident for the relatively small pairs (e.g., 11 vs 12) response times for the larger pairs (e.g., 14 vs 15, 15 vs 16 and 16 vs 17) show a clear advantage with the instruction "larger". This effect however failed to reach significance.

It seems logical from the contrasting results between the subtraction (Mirsky, 1993) and addition experiments that the degree of semantic congruity depends on the congruence of the arithmetic operation with the unmarked instruction ("larger"). Subtraction requires that a smaller number be obtained, an idea incongruent with the unmarked instruction ("larger"). As a result the markedness effect is countered and semantic congruity enhanced. Conversely, addition requires that a larger number be obtained, an idea congruent with the unmarked instruction. In this case, the markedness effect is enhanced minimizing semantic congruity.

Addend Order Effect. As explained previously, the addend order effect refers to the effect of the relative positions of the larger and smaller addends in an addition problem on response time. A preliminary examination of Figure 8 suggests that although there does not appear to be an addend order effect at an addend split of one (left panel), there seems to be one at the addends split of two (right panel). Response times for the order "larger + smaller" appear to be somewhat faster than for the order "smaller + larger". In fact, this response time pattern was obtained in 29 out of 44 cases incidentally replicating the proportion found by Parkman and Groen

(1971) in their examination of the addend order effect.

An analysis of variance with addend split (1,2), pair (14) and addend order (2) as independent variables and response time as the dependent variable was conducted. Subsequent error rate analyses used the same independent variables but used the arcsine transformation of proportion of correct responses as the dependent variable. Main effects of addend split [$F(1,22)=4.69$, $p=0.014$] and pair [$F(13, 286)= 23.72$, $p<0.0001$] were found. Indeed as predicted a borderline significant addend split by addend order interaction was obtained [$F(1,22)=4.28$, $p=0.0505$] showing that the addend order effect does depend on addend split. Further simple effects analyses indicated that addend order had no effect at an addend split of one but was significant at an addend split of two [$F(1, 22)=4.40$, $p=0.0478$].

In summary, Experiment 1 validated the operate and compare paradigm by replicating fundamental effects observed in verification and production tasks typically used in mental arithmetic research. These included a ties effect, problem difficulty effects, and magnitude effects. A second experiment was conducted to replicate these findings and explore further some unexpected results. First the addend order effect which has only been sporadically reported in the literature (Parkman & Groen, 1971) seems to depend on addend split. Experiment 2 attempted a replication and an extension by including addend splits of 1,2 and 4. Second, the almost complete absence of the semantic congruity effect was surprising given its clear and very substantial magnitude in subtraction. As a result, congruity was also a focus for Experiment 2. As well, Mirsky (1993) found clear violations of additivity in the

Figure 8. Left and right panels (addend splits 1 and 2) show response times as functions of comparison pair for the two addend orders. At an addend split of 2 (e.g. 6+4) response times were significantly longer for order, smaller+larger (filled triangles) than for the order, larger+smaller (open diamonds) suggesting some support for counting models. Tick labels for comparison pair refer to the minimum magnitude of each comparison pair (e.g., 2 is the minimum comparison magnitude in the pair 2 vs. 3).

context of subtraction and comparison of numbers which suggesting non-independence of processing. In order to establish a more complete parallel with the results of the subtraction experiment, Experiment 2 examined the additivity of the arithmetic operation and comparison process by including two comparison splits.

Presentation Order Effects. A clear, albeit small, main effect of the position of the larger magnitude was found [$F(1,22)=5.82$, $p<0.0247$]. When the larger magnitude (addition or digit) was on the right, overall response times were 22 ms faster than when the larger magnitude was on the left. Closer examination of the data revealed that the differences in response times for simple and compound comparisons (involving unequal addends) were 32 and 31 ms respectively. These results are surprising for two reasons. First, presentation of the stimuli on each trial was simultaneous and second, such magnitude order effects are invariably found with sequential presentation of perceptual magnitudes and have been called Space Errors possibly interpretable as Time-Order Errors (TOEs) (Fechner, 1869/1966). Negative time-order errors are typically characterized by longer response times and greater error rates for perceptual stimuli presented in the large-small order compared to the small-large order. Various theories have been proposed as to the basis for this effect including loss of information from the internal representation of the first stimulus (Kohler, 1923), drift of the internal representations of current stimuli towards the geometric mean of all previously presented stimuli (Woodrow, 1933), and more recently, differential sensation weighting (Hellstrom, 1979).

Differential sensation weighting (Hellstrom, 1979) best accounts for TOE

effects including magnitude of the stimuli, interstimulus interval, and stimulus range. Hellstrom (1979) suggested that successive stimuli are differentially weighted with respect to their distance from the geometric mean of all previously presented stimuli. While the stimuli in the current experiment were presented simultaneously, an interstimulus interval can still be said to exist since stimuli may conceivably be scanned from left to right as is the case for reading. Research with physical stimuli suggest that TOEs do in fact occur with perceptual stimuli but not with stimuli in memory. Using the method of paired comparisons with simultaneously presented squares, Petrusic, Baranski and Aubin (1995) found a 100 ms advantage for the small - larger order over the large - small order. Replicating studies employing successively presented stimuli (Baranski & Petrusic, 1992), Petrusic et al. (1995) did not find TOEs with comparisons of memory based symbolic stimuli suggesting that TOEs are unique to strictly perceptual comparisons. More importantly and critically was the finding that TOEs also occurred when one of the squares was held in memory while the other was perceptually available. Taken together with these findings the occurrence of presentation order effects with simple and compound numerical comparisons suggests that the representations of number underlying these comparisons is analogue based. Interestingly, Restle (1967) and more recently Dehaene (1992) have both advanced the notion that numbers have a functional analogue representation.

Experiment 2

Method

Subjects. Participants were twenty-three undergraduate students (5 females and

18 males), 18 to 25 years of age with normal or corrected-to-normal vision and were enrolled in a first year introductory psychology course at Carleton University (Ontario, Canada) for course credit.

Apparatus. The apparatus of Experiment 1 was used.

Stimuli. The total stimulus set was made up of two combined subsets, simple and compound comparisons. The number of comparison splits was increased to include splits 1 and 3. There were a total of 78 simple comparisons ranging between 2 vs.. 3 and 17 vs..18 for comparison splits of 1, and ranging between 1 vs.. 4 and 16 vs.. 19 for comparison splits of 3 (see Appendix A-2 and A-3). After taking into account two orders and two instruction there were a total of 312 simple comparisons.

Unlike Experiment 1, only compound comparisons involving unequal addends were included and addend splits were increased to include problems with a split of 4 in addition to those with a split of 1 and 2 (e.g. $8+7$, $8+6$, $8+4$). All addends were single digits. The addition problems were presented in both addend orders (e.g. $2+3$, $3+2$) for each of the three comparison splits. At an addend split of 1 there were 16 addition problems at a comparison split of 1 and 14 addition problems at a comparison split of 3 for a total of 30. At an addend split of 2 there were 14 addition problems at both comparison splits 1 and 3 totalling 28. At an addend split of 4 there were 10 addition problems at both comparison split of 1 and 3 for a total of 20 problems. Since the addition of 1 to a single digit were of interest, those adding by 1 problems not yet included were added at only a comparison split of one. This added a further 10 compound comparisons (e.g., $5+1$, $6+1$). The total number of trials for

all addend and comparison split combinations plus the addition adding by ones combinations was 126. After multiplying by two addend orders, two instructions and two presentation orders, the total number of number/addition problem comparisons will be 1016.

For each participant, the complete stimulus set was randomly generated and divided into 8 blocks of 127 trials. The 16 practice trials were randomly selected with replacement for each participant.

Procedure. The procedure was exactly the same as for Experiment 1 except that the experimental session was approximately 1 hour and 15 minutes.

Results

Of twenty-three participants, one was excluded for not completing the task and another was excluded for contributing an extremely high proportion of outliers. The latter participant contributed 56.25% of the total number of outliers below 200 ms in addition to a high error rate (14%). The 21 remaining participants provided a total of 21,336 trials. Trials with response times below 200 ms or exceeding 15,000 ms were treated as outliers and discarded prior to analysis. According to these criteria it was necessary to remove but 42 trials from the data, less than 0.002%, leaving 21,294 trials for further analysis. Of the remaining trials, there were 808 errors resulting in an error rate of 3.66%. Since a comparison split of three was used only for the purpose of determining additivity of addition and comparison processes, all analyses except for that for additivity were conducted at a comparison split of 1. In addition, compound comparisons involving ties were not used in this experiment. As a result,

the expression "compound comparisons" will refer to compound comparisons involving unequal addends.

Additivity. As is evident in Figure 9 (left panel), differences in response times between comparison types decreased as the comparison split increased for both correct and error responses suggesting non-additivity of the comparison and addition processes. Additivity would require that the response time functions plotted in Figure 9 be parallel to each other which they are not. At comparison splits of 1, the most difficult level of comparison (Dehaene, 1989; Link, 1990; Moyer & Landauer, 1967), the compound comparisons not only exhibited the slowest response times but also the greatest error rates. These differences were attenuated at a comparison split of 3. This major result of Experiment 2 was confirmed by a significant comparison by type comparison split interaction [$F(1,20)=26.62$, $p<0.0001$]. This same effect was evident in separate ANOVAs for correct [$F(1,20)=26.35$, $p=0.0001$] and incorrect [$F(1,20)=6.35$, $p=0.0204$] response times.

Effects of Comparison Type. As is clear from Figure 10 (left panel), overall response times for compound comparisons were considerably slower than those for simple comparisons replicating the results of Experiment 1. A 3-way ANOVA with response time as the dependent variable and comparison type (simple, compound)instruction (larger, smaller) and comparison pair (2 vs.3 through 17 vs.18) as independent variables. The main effect of comparison type was indeed reliable [$F(1,20)=89.45$, $p<0.0001$]. On average, simple comparisons were 703 ms faster

Figure 9. The left panel shows mean response times for correct (filled points) and error (unfilled points) responses for the simple (digit vs digit) comparisons (triangles) and compound (digit vs addition) comparisons (circles) as a function of the absolute difference between sum the two addends and the comparison digit. Probabilities of error responses in the simple and compound comparison conditions are presented in the right panel.

than compound comparison. Response times for errors reflected patterns similar to those obtained for overall responses.

Effect of Pair. Figure 10 shows that response times increase as a function of the magnitude of the numbers being compared for both comparison types. A reliable main effect of pair was found [$F(15,300)=14.62$, $p<0.0001$]. Similarly to Experiment 1, the form of the relation between response time and comparison pair differs according to comparison type. This apparent interaction was confirmed by a significant 2-way interaction [$F(15,300)=13.66$, $p<0.0001$]. The effect of comparison pair on response times is indeed dependent on the comparison type involved.

As was the case in Experiment 1, the increasing steepness of slope for compound comparisons relative to that for simple comparisons suggests that problem difficulty and comparison magnitude effects are jointly at work. Trend analyses for each comparison type showed clear and significant linear trends for simple comparisons [$F(1,20)=5.43$, $p=0.0304$], reflecting pure magnitude effects, and for compound comparisons [$F(1,20)=40.65$, $p<0.0001$]. Compound comparisons were further characterized by a significant quartic trend [$F(1,20)=8.13$, $p=.0099$] suggesting, as for Experiment 1, that the importance of the abrupt increase in response times at the junction of the first and second decades for compound comparisons.

An analysis of errors rates using an arcsine transformation demonstrated support for the findings related to the effects of comparison type and comparison pair on response times. Main effects were found for comparison type [$F(1,20)=23.12$,

Figure 10. The left panel shows response time plotted as function of comparison pair collapsing over correct and error responses. Note that response times for compound comparisons (open triangles) increase more quickly across comparison pairs relative to response time increases for simple comparisons suggesting non-additivity of processing. Neither error response time patterns of simple nor compound comparisons (left panel) were significantly different from the patterns displayed by correct responses, permitting both kinds of responses to be examined together. Tick labels for comparison pair refer to the minimum magnitude of each comparison pair (e.g., 2 is the minimum comparison magnitude in the pair 2 vs 3).

Figure 11. Shows error rates plotted as a function of comparison pair for simple (filled circles) and compound (unfilled triangles) comparisons. Note the similarity of error rate patterns to response time patterns of Figure 10 both within and between comparison types. Tick labels for comparison pair refer to the minimum magnitude of each comparison pair (e.g., 2 is the minimum comparison magnitude in the pair 2 vs 3)

$p=0.0001$]. Figure 11 shows the similarity of error rate patterns to the response patterns seen in Figure 10, that is faster response times are generally associated with lower error rates.

Instruction. Overall, response times for the instruction "larger" were 82 ms faster than response times for the instruction "smaller" consistent with a markedness effect also demonstrated in Experiment 1. Although a significant main effect was found for instruction [$F(1,20)=6.42$, $p=0.0197$] for response times, no significant effect of instruction occurring for errors was evident.

Effect of Decade. As for Experiment 1, comparison pairs were matched for their respective positions within each decade in order to investigate the effects of decade on comparison type, comparison pair and instruction. Paralleling Figure 5 in Experiment 1, Figure 12 shows mean response time and error rate patterns plotted as a function of unit step within a decade separately for the first decade (1-10) and second decade (10-20) for simple (upper and lower left panels) and compound comparisons (upper and lower right panels). Unlike Experiment 1, there was no significant decade effect for simple comparisons (upper left panel) but the similarity of the results with respect to compound comparisons is clear. On average, response times for compound comparisons were 556 ms slower in the second decade relative to the first decade. The strength of difference in response time patterns between comparison types was reflected in a reliable main effect of comparison type [$F(1,20)=78.96$, $p<0.0001$].

Replicating Experiment 1, the main effect of decade was reliable

$$[F(1,20) = 35.02,$$

Figure 12. Response times (upper panels) and error rates (lower panels) plotted as functions of unit steps within the first decade (open circles) and second decade (filled circles). The similar additive nature of response time patterns for compound comparisons between decades again suggests categorisation of numbers in the first decade as relatively small and numbers in the second decade as relatively large and the basis of the funnel-like congruity effects seen in Figure 13.

$p < 0.0001$] but as expected this effect was qualified by a reliable comparison type by decade interaction [$F(1,20)=29.10$, $p < 0.0001$]. Although not as descriptively tidy as in Experiment 1 additive effects of decade on response times are apparent in the upper right panel of Figure 12. This observation was supported by a non-significant decade by unit step interaction. A simple main effect of unit step within a decade [$F(5,100)=3.21$, $p < 0.05$] was interpreted as reflecting magnitude effects. The additive effects of decade suggests equivalent effects of magnitude within each decade. That is, irrespective of whether processing single digits found in the first decade or double digits found in the second decade, magnitude effects are of relatively similar strength.

Error rates displayed in Figure 12 (lower panels) show the same patterns as for response times (upper panels). Specifically, no decades effect for simple comparisons but a marked effect for compound comparisons.

Presentation Order Effects In replication of Experiment 1, response times were 52 ms faster when the larger comparison magnitude was presented on the left than when it was presented on the right. This difference in response times was confirmed as being statistically significant [$F(1,20)=7.85$, $p=0.0110$] and provides further support for the idea that analogue information is being retrieved from the internal representation of numerical magnitudes.

Semantic Congruity Effects and Addend Order Effects. Figure 13 clearly shows an absence of congruity across the range of comparison pairs again replicating Experiment 1. A funnel effect is also observed in Experiment 2 with response times

Figure 13. Shows overall response times for compound comparisons plotted as function of comparison pair for the instruction "larger" (filled circles) and the instruction "smaller" (open triangles). Funnel-like congruity effects can be seen in the second decade which contains large numbers relative to the first decade.

for the instruction "smaller" being slower than those for the instruction "larger" especially at the upper end of the second decade.

In contrast to Experiment 1 no addend order effects were observed for either addend splits of 2 or 4 suggesting that the effect observed in Experiment 1 was anomalous.

Adding by One. In Figure 14 response times are plotted as a function of comparison pair for compound comparisons involving three types of addition problems: 1) additions in which 1 is not an addend, 2) additions in which 1 is an addend and the other addend does not match the comparison number (e.g., $3+1$ vs. 5), and finally, 3) comparisons in which 1 is an addend and the other addend is equal to the comparison number (e.g., $4+1$ vs 4). An overall ANOVA with comparison type and pair as independent variables and response time as the dependent variable found a significant main effect of comparison type. Newman-Keuls performed on the levels of the comparison type factor revealed that while comparisons involving adding by one were each significantly different from other types of compound comparisons [$t(3,252)=6.55$, $p<.01$; $t(2,252)=4.85$, $p<.01$], adding by ones comparisons did not differ from each other [$t(2,252)=1.70$, $p>.05$]. As is evident, comparisons involving addition by 1 are faster than other compound comparisons. Further, the apparent lack of a difference in response times between the two types of comparisons involving addition by 1 was confirmed through a non-significant result. This was surprising considering that when one of the addends is equal to the comparison number, it is not necessary to determine the sum before

Figure 14. Response times plotted as a function of comparison pair for compound comparisons in which one of the addends is not 1 (filled circles), compound comparisons in which one of the addends is 1 (unfilled triangles), and compound comparisons in which one of the addends is 1 and the other addend is equal to the comparison number (filled triangles). Comparisons involving addition by one were consistently and significantly faster than other compound comparisons. Both types of addition by 1 exhibited similar response time patterns and were significantly different from each other.

making the comparative judgment. Any number plus one will be a larger number than itself, the decision is straight forward, no sum is required.

General Discussion

Collectively, the results of the present experiments replicate and extend those obtained with subtraction (Mirsky, 1993) strongly supporting the validity of the operate and compare paradigm. These data replicate not only the ties and problem difficulty effects, commonly found in paradigms associated with the area of mathematical cognition (verification and production tasks), but also split and magnitude effects revealed in the psychophysical comparison literature involving numerical, perceptual and remembered stimuli. The operate and compare paradigm permits the opportunity to bring together what have been two separate literatures to provide a richer characterization of the properties of and the relationships between arithmetical and comparison processes.

The violation of additivity is a clear and important finding since non-independence of arithmetic and comparison processes is not intuitive. As was shown for subtraction by Mirsky (1993), there was a significant interaction between the time to add (problem difficulty) and the degree of difficulty of subsequent numerical comparisons (split) attesting to the non-independence of these intuitively separate stages (Sternberg, 1969). Response time patterns were paralleled by error rate patterns.

Non-additivity in the context of adding and comparing numbers are consistent with findings of Zbrodoff & Logan (1990) who used both production and verification

paradigms to examine the properties of addition. Taken together, Experiments 1 and 2 provide converging evidence suggesting similarity between the processes involved in binary same-different decisions in the context of verification tasks and decision concerning relative magnitudes.

Findings of non-additivity lends strong support for theories that posit analog representations of number as a fundamental element of human numerical competence. Neither associative network (Ashcraft, 1982; Campbell et al., 1985; Siegler et al., 1984) nor Dehaene's (1992) triple code model can account for the interaction found in this experiment. For example, if the sequence of events was, to first retrieve the required arithmetic fact from the associative network of facts in long term memory and to secondly, translate the retrieved answer to an analog representation in order to make the judgment of comparative magnitude, then problem difficulty would not be relevant at the decision stage.

It is possible using Campbell's (Campbell, 1987, Campbell & Graham, 1985) associative interference model, to argue that the interaction is caused by simultaneous presentation of the arithmetic problem with comparison number. Campbell (1987) argued that presentation of true or false answers in verification tasks distorts the problem difficulty effect and disrupts the speed and accuracy of retrieval. True answers act as priming stimuli facilitating retrieval. Alternatively, presentation of false answers activates associative links which either increase the size of the candidate set or make it more difficult to discriminate between correct and incorrect answers. This argument would be a plausible explanation for the interaction between problem

difficult and comparison split, except that Zbrodoff et al. (1990) found that split did not influence processing of the arithmetic problem.

Zbrodoff and Logan (1990) experimental design permitted an evaluation of the direction of the interaction between problem magnitude and comparison split. This was done by imposing a variety of delays between the onset of problem stimuli and the onset of answers to be verified. They hypothesized that if enough time was allowed for subjects to generate problem answers before presenting the answer for them to verify, the problem difficulty effect would be absorbed and would not interact with comparison split. Their results indicated that while split did not influence response times for addition, problem magnitude did influence comparison response times. The similarity between the effects found in the present experiments, as well as Mirsky's (1993) subtraction data, to those of Zbrodoff and Logan suggest that it is both likely and logical that the influence is one of problem difficulty effect on the comparison processes.

Restle (1970) could have settled the issues if he had examined and compared both the add and compare and the compare only conditions in his experiment, but this seems to have been an oversight on his part. He hypothesized an analog representation on which the operation and comparison were performed.

Another important and telling finding, or rather lack thereof, in the present experiments was the absence of congruity effects. This was especially interesting since congruity was strongly enhanced in the subtract and compare study (Mirsky, 1993). The hypothesis presented here is that the congruency of the arithmetical operation

with the instruction influences the impact of the markedness effect. When the operation was subtraction, an operation with which we obtain the smaller number, congruity effects were greatly enhanced. Conversely, in the addition experiments reported here the nature of the arithmetic operation acted in concert with the markedness effects causing minimal congruity effects.

Baranski and Petrusic (1992) demonstrated that the internal representation of remembered stimuli used in comparative decisions are analogical in character and of a lesser quality than their perceptual counterparts. Remembered stimuli are "more variable and less precise than the percepts they symbolically represent" (p.41). Evidence accrual theories posit that in order to achieve comparable levels of accuracy, as found in comparison performance with perceptual stimuli, internal evidence accumulators require more accruals of evidence before response criteria are met (Petrusic, 1992). Additionally, data exists suggesting that accruals of evidence for remembered stimuli will be longer, due to the fact that the memory retrieval process must be reactivated on each accrual (Moyer & Bayer, 1978). Moreover, Petrusic (1992) found evidence suggesting that semantic congruity effects occurring on each accrual behave in an accumulative way. Collectively, this information reasonably leads to the conclusion that semantic congruity effects would be amplified for remembered comparisons (Petrusic, 1992; Petrusic, Baranski & Aubin, 1995).

If evidence accrual theories are correct, they would successfully account for the interaction of problem size and comparison split found in this and other experiments (Campbell, 1987; Zbrodoff and Logan, 1990). They would also provide

viable explanations for the amplification of semantic congruity effects under relative comparison judgment tasks of remembered numerical stimuli.

Although there is considerable evidence supporting associative network models to explain and predict numerous phenomena observed in mental arithmetic performance (Ashcraft, 1982, 1992; Campbell, 1987; Campbell & Graham, 1985; Lefevre et al., 1988; Siegler et al., 1984; Winkelman & Schmidt, 1974; Zbrodoff et al., 1986) there is also strong convergent evidence indicating internal analogue representations of numerical magnitudes (Dehaene, 1992; Gellistel & Gelman, 1992; Link, 1990; McCloskey, 1992; Restac, 1970). The operate and compare paradigm introduced using subtraction (Mirsky, 1993) and validated in the current experiments using addition provides a promising means of characterizing the richness of the relationship between mental arithmetic and comparative judgments of numerical magnitudes.

References

- Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. Developmental Review, 2, 213-236.
- Ashcraft, M. H. (1985). Is it farfetched that some of us remember our arithmetic facts? Journal for Research in Mathematics Education, 16, 99-105.
- Ashcraft, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), Formal methods in developmental psychology (pp. 302-338). New York: Springer-Verlag.
- Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. Cognition, 44, 75-106.
- Ashcraft, M. H. (1990). Strategic processing in children's mental arithmetic: A review and proposal. In D.F. Bjorklund (Ed.), Children's strategies: Contemporary views of cognitive development. (pp. 185-211). Hillsdale, NJ: Erlbaum.
- Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. Journal of Experimental Psychology, 4, 527-538.
- Ashcraft, M. H., Fierman, B. A., & Bartolotta, R. (1984). The production and verification tasks in mental addition: An empirical comparison. Developmental Review, 4, 157-170.

- Baird, J.C. & Noma, E. (1975). Psychophysical study of numbers. I. Generation of numerical responses. Psychological Research, 38, 189-207.
- Banks, W., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. Journal of Experimental Psychology: Human Perception and Performance, 2, 435-447.
- Banks, W. P., Mermelstein, R., & Yu, H. K. (1982). Discriminations among perceptual and symbolic stimuli. Memory & Cognition, 10, 265-278.
- Baranski, J. V., & Petrusic, W. M. (1992). The discriminability of remembered magnitudes. Memory & Cognition, 20(3), 254-270.
- Baroody, A. J. (1984). A reexamination of mental arithmetic models and data: A reply to Ashcraft. Developmental Review, 4, 148-156.
- Baroody, A. J. (1985). Mastery of basic number combinations: Internalization of relationships or facts? Journal for Research in Mathematics Education, 16, 83-98.
- BMDP Statistical Software Manual (1990). Berkeley, CA: University of California Press.
- Buckley, P.B., & Gillman, C.B. (1974). Comparisons of digits and dot patterns. Journal of Experimental Psychology, 103, 1131-1136.
- Campbell, J. I. D. (1987). Network interference and mental multiplication. Journal of Experimental Psychology: Learning, Memory and Cognition, 13, 109-123.
- Campbell, J. I. D. (1990). Retrieval inhibition and interference in cognitive arithmetic. Canadian Journal of Psychology, 44, 445-464.

- Campbell, J. I. D. (1992). Network-Interference model of arithmetic memory: Developments since Campbell & Oliphant. Personal Communication.
- Campbell, J. I. D. & Clark, J. (1989). Time course of error priming in number-fact retrieval: Evidence for excitatory and inhibitory mechanisms. Journal of Experimental Psychology, 15, 920-929.
- Campbell, J. I. D., & Graham, D. (1985). Mental multiplication Skill: Structure, process, and acquisition. Canadian Journal of Psychology, 39, 338-366.
- Campbell, J. I. D., & Oliphant, M. (1992). Representation and retrieval of arithmetic facts: A network-interference model and simulation. In J. I. D. Campbell (Ed.), The nature and origins of mathematical skills (pp. 331-363). Holland: Elsevier.
- Cattell, J. M. (1902). The time of perception as a measure of differences in intensity. Philosophische Studien, 19, 63-68.
- Collins, A.M., & Loftus, E.F. (1975). A spreading activation theory of semantic processing. Psychological Review, 82, 407-428.
- Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. Perception & Psychophysics, 45, 557-566.
- Dehaene, S. (1992). Varieties of numerical abilities. Cognition, 44, 1-42.
- Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. Journal of Experimental Psychology: General, 122(3), 371-396.

- Dehaene, S., Dupoux, E., & Mehler, J. (1990). 's numerical comparison digital? Analogical symbolic effects in two-digit number comparison. Journal of Experimental Psychology: Human Perception and Performance, 16, 626-641.
- Donders, F. C. (1969). On the speed of mental processes. Acta Psychologica, 30, 412-431.
- Fechner, G. T. (1966). Elements of psychophysics (H. E. Adler, Trans.). New York: Harper & Row. (Original work published 1860).
- Gallistel, C. R., and Gelman (1992). Preverbal and verbal counting and computation. Cognition, 44, 43-74.
- Geary, D., Widaman, K., & Little, T. (1986). Cognitive addition and multiplication: Evidence for a single memory network. Memory & Cognition, 14, 478-487.
- Groen, G., & Parkman, J. (1972). A chronometric analysis of simple addition. Psychological Review, 79, 329-343.
- Hamann, M. S., & Ashcraft, M. H., (1985). Simple and complex mental addition across development. Journal of Experimental Child Psychology, 40, 49-72.
- Hellstrom, A. (1979). Time errors and differential sensation weighting. Journal of Experimental Psychology: Human Perception and Performance, 5(3), 460-477.
- Henmon, V. A. C. (1906). The time of perception as a measure of differences in sensation. Archives of Philosophy, Psychology & Scientific Methods, 8, 1-75.
- Hinrichs, J.V., Yurko, D.S., & Hu, J.M. (1981). Two-digit number comparison: Use of place information. Journal of Experimental Psychology: Human Perception and Performance, 7, 890-901.

Holyoak, K.J. (1978). Comparative judgments with numerical reference points.

Cognitive Psychology, 10, 203-243.

Jamieson, D., & Petrusic, W. (1975). Relational judgments with remembered stimuli.

Perception & Psychophysics, 18, 373-378.

Johnson, D. M. (1939). Confidence and speed in the two-category judgment. Archives

of Psychology, 241, 1-52.

Kellog, W. N. (1931). Time of judgment in psychometric measures. American Journal

of Psychology, 43, 65-86.

Kohler, W. (1923). Zur Analyse des Sukzessivvergleichs und der Zeitfehler [On the

analysis of successive comparison and of the time-order error]. Psychologische

Forschung, 4, 115-175.

Lefevre, J., Bisanz, J., & Mrkonjic, L. (1988). Cognitive arithmetic: Evidence for

obligatory activation of arithmetic facts. Memory & Cognition, 16, 45-53.

Link, S. (1990). Modeling imageless thought: The relative judgment theory of

numerical comparisons. Journal of Mathematical Psychology, 34, 2-41.

Luce, R. D. (1986). Reaction times. New York: Oxford University Press.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from

acquired dyscalculia. Cognition, 44, 107-157.

Miller, K., Perlmutter, M., & Keating, D. (1984). Cognitive arithmetic: Comparison

of Operations. Journal of Experimental Psychology: Learning, Memory, and

Cognition, 10, 46-60.

2 of/de 2

PM-1 3 1/2 x 4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

- Mirsky, C.A. (1993). On the time to subtract and compare: Non-Additivity in mental arithmetic. Unpublished honours thesis, University of Ottawa, Ottawa, Ontario.
- Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. Perception & Psychophysics, 13, 180-184.
- Moyer, R. S., & Bayer, R. H. (1976). Mental comparisons and the symbolic distance effect. Cognitive Psychology, 8, 228-246.
- Moyer, R., & Landauer, T. (1967). Time required for judgements of numerical inequality. Nature, 215, 1519-1520.
- Moyer, R., & Landauer, T. (1973). Determinants of reaction time for digit inequality judgments. Bulletin of the Psychonomic Society, 1, 167-168.
- Munsterberg, H. (1894). Studies from the Harvard psychological laboratory: A psychometric investigation of the psychophysics law. Psychological Review, 1, 45-51.
- Pachella, R. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), Human information processing: Tutorials in performance and cognition (pp.41-82). Hillsdale, NJ: Erlbaum.
- Parkman, J. M. (1971). temporal aspects of digit and letter inequality judgments. Journal of Experimental Psychology, 91, 191-205.
- Parkman, J. M. (1972). Temporal aspects of simple multiplication and comparison. Journal of Experimental Psychology, 95, 437-444.
- Parkman, J. M., and Groen, G. J. (1971). Temporal aspects of simple addition and comparisons. Journal of Experimental Psychology, 89, 335-342.

- Petrusic, W.M. (1992). Semantic congruity effects and theories of the comparison process. Journal of Experimental Psychology: Human Perception and Performance, 18, 962-986.
- Petrusic, W.M., Baranski, J.V., & Aubin (1995). Internal psychophysics: Binary and quaternary relational judgments with remembered and perceived magnitudes. Under review.
- Poltrock, S. E. (1989). A random walk model of digit comparison. Journal of Mathematical Psychology, 33(2), 131-162.
- Restle, F. (1970). Speed of adding and comparing numbers. Journal of Experimental Psychology, 83, 274-278.
- Rule, S. J. (1969). Equal discriminability of number. Journal of Experimental Psychology, 79, 35-39.
- Schneider, B. Parker, S., Ostrosky, D., D., Stein, D., & Kanow, G. (1974). A scale for the psychological magnitude of numbers. Perception & Psychophysics, 16, 43-46.
- Sekuler, R., Rubin, E. & Armstrong, R. (1971). Processing numerical information: A choice time analysis. Journal of Experimental Psychology, 90, 75-80.
- Shepard, R.N., Kilpatrick, D.W., & Cunningham, J.P. (1975). The internal representation of number. Cognitive Psychology, 1, 82-138.
- Siegler, R. (1988). Strategy choice procedures and the development of multiplication skill. Journal of Experimental Psychology: General, 117, 258-275.

- Siegler, R. & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), Origins of cognitive skills (pp. 229-287). Hillsdale, NJ: Erlbaum.
- SPSSX Users Guide (2nd ed.). (1986). Chicago, IL: McGraw-Hill.
- Stazyk, E. H., & Ashcraft, M. H. (1981). Mental addition: A test of three verification models. Memory and Cognition, 9, 185-196.
- Stazyk, E. H., Ashcraft, M. H., & Hamann, M. (1982). A network approach to mental multiplication. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 320-335.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. Acta Psychologica, 30, 276-315.
- Suppes, P., & Groen, G. J. (1967). Some counting models for first grader performance data on simple addition facts. In J. M. Scandura (Ed.), Research in mathematics education. Washington, D. C.: National Council of Teachers.
- Svenson, O. (1975). Analysis of time required by children for simple addition. Acta Psychologica, 39, 289-302.
- Wald, A. (1945). Sequential tests of statistical hypotheses. The Annals of Mathematical Psychology, 16, 117-186.
- Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. Ergonomics, 3, 189-230.
- Winkelman, J., & Schmidt, J. (1974). Associative confusions in mental arithmetic. Journal of Experimental Psychology, 102, 734-736.

- Woodrow, H. (1933). Weight discrimination with a varying standard. American Journal of Psychology, 45, 391-416.
- Woods, S. S., Resnick, L. B., & Groen, G. J. (1975). An experimental test of five process models for subtraction. Journal of Educational Psychology, 67, 17-21.
- Zbrodoff, N. J. & Logan, G. (1986). On the autonomy of mental processes: A case study of arithmetic. Journal of Experimental Psychology: General, 115, 118-130.
- Zbrodoff, N. J., & Logan, G. (1990). On the relation between production and verification tasks in the psychology of simple arithmetic. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 83-87.

Appendix A

Table A-1

Stimulus Set for Experiment 1

Addend Split								
0 (Ties)			1			2		
Problem	Sum	Comparison Number	Problem	Sum	Comparison Number	Problem	Sum	Comparison Number
1+1	2	1						
1+1	2	3	1+2	3	2			
2+2	4	3	1+2	3	4	1+3	4	3
2+2	4	5	2+3	5	4	1+3	4	5
3+3	6	5	2+3	5	6	2+4	6	5
3+3	6	7	3+4	7	6	2+4	6	7
4+4	8	7	3+4	7	8	3+5	8	7
4+4	8	9	4+5	9	8	3+5	8	9
5+5	10	9	4+5	9	10	4+6	10	9
5+5	10	11	5+6	11	10	4+6	10	11
6+6	12	11	5+6	11	12	5+7	12	11
6+6	12	13	6+7	13	12	5+7	12	13
7+7	14	13	6+7	13	14	6+8	14	13
7+7	14	15	7+8	15	14	6+8	14	15
8+8	16	15	7+8	15	16	7+9	16	15
8+8	16	17	8+9	17	16	7+9	16	17
9+9	18	17	8+9	17	18			
9+9	18	19						
Total Pairs		18			16			14

Note. Simple and Compound comparisons were presented in both addend orders (e.g., 2+4, 4+2), with both instructions ("larger", "smaller") and in both presentation orders (e.g., 7 2+4, 2+4 7).

Appendix A

Table A-2

Stimulus Set for Experiment 2

Simple and Compound Comparisons								
Addend Split		Comparison Split		Addend Split		Comparison Split		
1		1	3	2	4	1	3	
Problem	Sum ^a	Comparison Number		Problem	Sum	Comparison Number		
1+2	3	2						
1+2	3	4	6	1+3	4	3	1	
2+3	5	4	2	1+3	4	5	7	
2+3	5	6	8	2+4	1+5	6	5	3
3+4	7	6	4	2+4	1+5	6	7	9
3+4	7	8	10	3+5	2+6	8	7	5
4+5	9	8	6	3+5	2+6	8	9	11
4+5	9	10	12	4+6	3+7	10	9	7
5+6	11	10	8	4+6	3+7	10	11	13
5+6	11	12	14	5+7	4+8	12	11	9
6+7	13	12	10	5+7	4+8	12	13	15
6+7	13	14	16	6+8	5+9	14	13	11
7+8	15	14	12	6+8	5+9	14	15	17
7+8	15	16	18	7+9		16	15	13
8+9	17	16	14	7+9		16	17	19
8+9	17	18						
Total ^b		16	14	14	10			

Note. Simple and Compound comparisons were presented in both addend orders (e.g., 2+4, 4+2), with both instructions ("larger", "smaller") and in both presentations orders (e.g., 7 vs. 2+4, 2+4 vs. 7).

^aSums of addition problems listed represent the comparison number stimuli used in the comparable simple comparison condition (e.g., the corresponding simple comparison to 7+8 vs. 14 is 15 vs. 14).

^bTotal refers to the total number of comparison pairs at each addend split at comparison splits 1,3.

Appendix A

Table A-3

Adding by Ones Stimulus Set used in Experiment 2

Comparison ^a		
Compound	Simple	Comparison Digit
2 + 1	3	2
2 + 1	3	4
3 + 1	4	3
3 + 1	4	5
4 + 1	5	4
4 + 1	5	6
5 + 1	6	5
5 + 1	6	7
6 + 1	7	6
6 + 1	7	8
7 + 1	8	7
7 + 1	8	9
8 + 1	9	8
8 + 1	9	10
9 + 1	10	9
9 + 1	10	11
Total Pairs		16

^aThe columns beneath this heading show the numerical formats for one of the comparison magnitudes in a comparison pair for compound and simple comparisons.

END

0 4 - 0 6 - 9 | 6

FIN