1909

COMPARISON OF TWO LEVELS OF GRADE ONE FRENCH INSTRUCTION AND THEIR RELATIONSHIP TO ARITHMETIC ABILITY AND RELATED COGNITIVE FUNCTIONING

by Virginia Carver

Thesis presented to the School of Graduate Studies of the University of Ottawa in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Psychology

ANNEXE DE LA BIBLIOTHÈQUE
University of Ottawa
BIBLIOTHÈQUES
LIBRARIES
University of Ottawa

Ottawa, Canada, 1974

© Virginia Carver, Ottawa, Canada, 1974

UMI Number: DC53951

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.
In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform DC53951
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

This thesis was prepared under the supervision of H. P. Edwards, Ph.D., Associate Professor and Dean of the Faculty of Psychology of the University of Ottawa.

The writer is indebted to Catherine Casserly, M.A., Research Officer, Ottawa Roman Catholic Separate School Board, for her help in carrying out the study.

CURRICULUM STUDIORUM

Virginia P. Carver was born December 17, 1940, in Berkhamsted, England. She received the Bachelor of Arts degree from Carleton University, Ottawa, Ontario, in 1970 and the Master of Arts degree in 1972 from the same university. The title of her thesis was The Effect of Message Redundancy and Vocabulary on Infants' Perception of English Language Samples.

ABSTRACT

Recent studies regarding the effects of bilingualism on cognitive functioning in children have shown that bilinguals may be more cognitively flexible and have a greater ability to attend than unilinguals. This, in turn, may be related to the enhanced arithmetic-ability found in children enrolled in second language immersion programs.

Comparison of Grade 1 children enrolled in a French immersion program with Grade 1 children enrolled in a normal English program receiving 75 minutes of French per day on measures of arithmetic-ability, concrete operational thinking and attention, as measured by the Matching Familiar Pictures test, found no significant differences between the two groups on any of the dependent variables.

In previous studies no account was taken of initial differences in number ability, and this may account for their findings of enhanced arithmetic-ability. Further research to see if the cognitive functioning and arithmetic-ability of bilingual and unilingual children diverge in later grades is suggested.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

INTRODUCTION

The relationship between bilingualism and intellectual functioning and bilingualism and academic functioning has been the subject of a large number of studies during the last fifty years. Though many of the early studies concluded that bilingualism depressed performance, more recent studies have shown that either the performance of bilingual children is comparable to that of monolingual children or that, in some cases, intellectual and academic performance appears to be enhanced. For instance, research on bilingualism and intelligence has shown that bilinguals may have more flexible cognitive structures (Peal & Lambert, 1962; Cohen, 1969; Balkan, 1970).

Many of the more recent studies have been done with children enrolled in second language immersion programs at the primary school level, with the main focus of the research being the effect of bilingualism on academic functioning. From these studies it appears that the bilingual children perform as well academically as monolingual children; and in some areas, notably arithmetic, make significantly higher test scores.

The purpose of this study is to examine the relationship between arithmetic and two related cognitive measures, logical thinking and cognitive style, in a sample of children enrolled in a French language immersion program in comparison with a sample of children enrolled in a regular English program.

Chapter I will first examine the controversy regarding the relationship between bilingualism and intellectual and academic performance. The relationship between arithmetic and bilingualism will then be discussed, relating arithmetic performance to the possible enhancement of logical thinking and cognitive style. Chapter II will present the research design, while Chapters III and IV will present the results and a discussion of the results.

CHAPTER I

REVIEW OF THE LITERATURE

The concern about the effects of bilingualism[1] on intellectual functioning and academic achievement has a long history both in North America, with its extensive influx of non-English speaking immigrants, and in countries such as Wales and Ireland where the native language has continued to exist alongside the dominant language, English. The burgeoning of intelligence testing after the First World War provided the available instruments to assess the abilities of children who were being schooled in a second language.

The literature on the relationship between bilingualism and intellectual or academic performance divides fairly neatly into two sections. The first set of studies deals mainly with children who are immigrants or members of minority language groups. A large proportion of this research concludes that bilingualism has a depressing effect on intellectual or academic performance. The second group of studies, done mainly in the last decade, shows that there

---

1. Bilingualism is used here in a very general sense to indicate both balanced bilinguals and children in the process of acquiring a second language.

is no necessary relationship between bilingualism and lowered intelligence test scores or poor academic achievement.

In the first part of the review of the literature an attempt will be made to explain the apparently contradictory findings regarding bilingualism and intelligence. The relationship between academic achievement and bilingualism will then be discussed, relating academic achievement studies to the bilingualism and intelligence controversy.

1. Bilingualism and Intelligence

Negative Findings

Several authors have done extensive literature reviews on the relationship between bilingualism and intelligence. These reviews cover studies done during the last fifty years and have, with a few exceptions, concluded that bilingualism has a detrimental effect on verbal IQ scores. If factors such as socio-economic class are equated between bilinguals and monolinguals, differences are not normally found on non-verbal IQ.

Darcy (1953, 1963) has reviewed a number of studies done between the nineteen twenties and nineteen sixties in which she found 21 which showed bilingualism depressed intelligence test scores, seven showing no difference, and only two which found bilinguals made higher IQ scores.

Darcy's findings are supported by literature reviews of Peal and Lambert (1962), Diebold (1968), and by the Welsh studies of Jones (1966).

There are a number of reasons why most of these early studies have found bilingualism to be associated with lowered intelligence test scores in contrast to the more recent research considered in the next section. These reasons relate basically to the type of populations used. The majority of the research has been done either with immigrant children in the United States, or with minority language groups in countries such as Wales and Ireland. Stevens (1966) has pointed out that the subjects of these studies are "members of an economically, occupationally and/or culturally underprivileged stratum within the society against whose norms they were measured. . . . (p. 75)." Many studies failed to control for social class or rural/urban differences when comparing groups of bilinguals and monolinguals. In addition, the definition of bilingualism is not comparable across these studies; nor the number of years exposure to the second language.

One can conclude then, that many of these children may well have been "disadvantaged" before entering school, and the handicap of bilingualism was added to already existing linguistic and cognitive deficits.

Positive Findings

This set of studies shows that bilingualism either does not appear to affect intellectual functioning or that bilingualism may be related to enhanced intellectual functioning. Apart from social class differences in the populations studied in this section in contrast to the previous section, two other factors may have contributed to the more positive findings to be discussed. First, as Ianco-Worrall (1972) has pointed out, the consequences of bilingualism have been traditionally evaluated in terms of global scores on standardized intelligence tests, but recent studies have examined more specific aspects of cognitive and academic functioning. Secondly, a number of educators in bilingual countries such as Canada have felt that the need to acquire two languages could best be met through special total immersion programs at the primary school level.

One of the first studies to show the positive effects of bilingualism was done by Peal and Lambert (1962) with ten-year-old French-Canadian children in Montreal. It has become something of a classic in the field of bilingualism and intelligence. Peal and Lambert used as subjects all ten-year-old children from six French-speaking schools in Montreal. All these schools could be classified as approximately middle class, and all ten-year-old subjects available at the time of testing were used regardless of grade. All the children were initially tested to discover the degree of bilingualism, and were divided into two groups, bilingual and unilingual. Information was also obtained on the child's family and father's occupation. The children were given a number of intelligence tests, the Lavoie-Laurendeau Group Test of General Intelligence, Raven's Progressive Matrices, and Thurstone's Primary Mental Abilities Test. When social-class differences between the two groups were held constant, it was found that the bilingual children obtained significantly higher overall scores on both verbal and non-verbal intelligence tests. When the same two groups were matched on non-verbal IQ, the bilingual children were still reliably superior in verbal IQ.

The most interesting aspect of this study is that certain non-verbal abilities appear to be related to bilingualism. Peal and Lambert (1962) compared the scores of the bilinguals and monolinguals on the tests that measured mainly spatial-perceptual abilities versus those that measured mainly concept formation or symbolic flexibility. They found that no differences existed on the spatial-perceptual tests, but that the bilinguals were superior on concept formation or symbolic flexibility tests.

Peal and Lambert (1962) made the point in their study that it was not possible to say whether bilingualism enhanced intelligence or if the more intelligent child became bilingual, though they leaned toward the former interpretation. Support for this interpretation comes from a number of studies done both in North America and Europe. These studies show that bilingualism may be enhancing what can very broadly be described as "cognitive flexibility." For instance, Balkan (1970) studied a group of bilingual children in Switzerland, matching each bilingual with a monolingual of the same sex, social background, age and intelligence. He administered the Batterie d'Aptitudes Scolaires Collective and matched his two groups on general reasoning or the "g" factor and then examined their scores on more specific intellectual factors. He found that the bilingual children were superior on numerical ability, verbal and perceptual plasticity, particularly the latter two. By breaking down the bilinguals into subgroups, Balkan was able to show that only early or balanced bilinguals were significantly different from monolinguals.

One can conclude from Balkan (1970) and Peal and Lambert (1962) that certain specific intellectual abilities are affected by bilingualism, but that this applies only to early or balanced bilinguals. Cohen (1969) has also studied bilingualism in the context of a second-language program at the preschool and primary school level in France. Her study looked at the relationship between second-language acquisition, New Math and traditional/active teaching methods with children who were eight years old at the time of the study. Though her results were somewhat confounded in that she was unable to constitute sufficient groups to isolate the effects of all three variables, she did find that children who had acquired their second language early, i.e., at the preschool age, made intellectual gains on Raven's Progressive Matrices in comparison with late bilinguals and monolinguals. It also seemed that bilingualism contributed to more specific cognitive gains in areas such as inductive reasoning and the ability to permute elements. Such cognitive abilities are normally characteristic of older children who have reached Piaget's formal operations stage. These children could be described as more cognitively flexible, less tied to concrete modes of thinking.

To round out this survey of positive findings, two other studies should be considered which provide support for the intellectually enhancing effects of bilingualism. The first study was also done with children enrolled in a second-language program and indicates that if bilingualism has an effect on intellectual functioning, it may not only apply to children who have acquired their two languages very early in life. In Grades 4 and 5 of the St. Lambert project, a primary school French immersion program for English-speaking children, the immersion children, in comparison to two English control groups initially matched on intelligence and social class measures, proved superior on a number of creativity subtests (Lambert, Tucker, d'Anglejan & Silny, 1972).

The second study was done in Singapore by Torrance, Gowan, Wu and Aliotti (1970). They compared monolinguals and bilinguals on the Torrance Test of Creative Thinking and found bilinguals to be superior on two of the tests, elaboration and originality, though only elaboration reached significance level. The monolinguals were superior on the flexibility and fluency subtests.

Two conclusions can be drawn from all these studies. The first is that studies which show bilingualism depresses IQ scores may well be using subjects who are "disadvantaged" in comparison with the majority culture. The second conclusion is that if there is a relationship between bilingualism and superior intelligence test scores, it seems to contribute to greater "cognitive flexibility."

2. Bilingualism and Academic Achievement

The previous section considered one aspect of bilingualism, namely its effect on intelligence. However, intelligence affects academic achievement, and it is therefore not surprising that a similar dichotomy in findings exists for bilingualism and academic achievement.

Negative Findings

MacNamara (1966) has carried out a comprehensive survey of studies dealing with the relationship between bilingualism and academic achievement, specifically arithmetic and linguistic skills. Most of the studies he considers were done with immigrants in the U.S.A., and thus they have the same handicaps that depressed intellectual functioning. MacNamara concludes that bilingual children make lower scores on language and arithmetic problem-solving tests in comparison to monolinguals; but that mechanical arithmetic is not affected, presumably because it is not so heavily dependent on language ability as problem solving arithmetic is.

MacNamara's conclusions from his extensive review of the literature were supported by his own study (MacNamara, 1966) on the effect of Gaelic instruction for English-speaking children in Ireland. He found that children who were taught completely in Gaelic in comparison with children taught in English or who had received only the first few years of schooling in Gaelic, made poorer scores in problem arithmetic but not in mechanical arithmetic. Generally, he also found their standard of English language skills were below those of children of similar age in Great Britain.

MacNamara (1966) drew discouraging conclusions regarding the effect of bilingualism and bilingual instruction on academic achievement. As was the case for intelligence, this may be due in large measure to the nature of the subject population. In contrast, the studies to be considered in the next section have either found that bilingualism does not affect academic achievement or that in some areas it is related to superior test scores.

Positive Findings

The studies considered in this section are concerned with children enrolled in special second-language programs at the primary school level. The results of three major longitudinal studies, two in Ottawa and one in Montreal, will be discussed in some detail since they indicate the relationship between bilingualism and specific aspects of academic functioning. The format of these studies is much the same in that children enrolled in the pilot Grade 1 are being followed through to Grade 8, and their scores on a series of intelligence and achievement tests compared with control groups enrolled in a regular English program. In addition to the pilot Grade 1, the progress of a follow-up Grade 1 is also being assessed.

The first immersion program to be initiated was the St. Lambert project in Montreal started in 1966. The subjects in this study are being compared to two English and one French control group with whom they were equated at the beginning of Grade 1 on intelligence and socio-economic class measures. Results so far (Lambert & MacNamara, 1969; Lambert, 1970; Lambert, Just & Segalowitz, 1970; Lambert, Tucker, d'Angeljan & Silny, 1972) indicate that the children have acquired a remarkable degree of fluency in French, though they are not overall up to the standard of the French controls. Moreover, their English skills in most areas have not suffered in comparison with the English controls. In arithmetic they have equalled or bettered the controls. Specifically, both the pilot and follow-up Grade 2 immersion children made significantly higher scores in arithmetic computation but not in arithmetic problem-solving. At the Grade 4 level, both the pilot and follow-up immersion children again made significantly higher scores in arithmetic on the Lorge-Thorndike mathematics subtest.

Similar results have been found with the two Ottawa French-immersion programs run by the Ottawa Separate School Board and the Ottawa Board of Education. In the former Board's program, the pilot Grade 1 children were compared to all other children enrolled in the Board's regular program receiving 75 minutes of French per day and to a sub-group matched on age, sex, and kindergarten background. In comparison with both these groups the immersion children did as well in English skills and, as was expected, were superior in French. In arithmetic they made significantly higher scores on both the concepts and computation subtests of the Metropolitan Achievement Test (Edwards & Casserly, 1971).

In the second year, the pilot and follow-up immersion groups were compared to samples of 75-minute children chosen as being representative of all children enrolled in the Board's system. In addition, they were compared to a sub-group matched on age, sex, and intelligence. The results for the second year of the program were very similar to the initial year. Both the follow-up Grade 1 and the pilot Grade 2 immersion children were superior to the control groups in French skills. In English skills they did perform somewhat more poorly in comparison to the control groups, but this is not surprising since neither the Grade 1 nor Grade 2 immersion children were getting any instruction in English Language Arts. In arithmetic the immersion children again made significantly higher scores on the Metropolitan Achievement Test in comparison with the 75-minute children. In Grade 1 the immersion children made higher scores on both the computation and concepts subtests, while in Grade 2 they made higher scores on the concepts subtest but curiously not on computation (Edwards & Casserly, 1972).

In the Ottawa Board of Education program the results for the pilot Grade 1 children and follow-up kindergarten children were essentially the same as those found for the Separate School Board program, with the Grade 1 immersion children doing less well in English language skills but better in French. There was no difference in arithmetic between the immersion children and the children in the regular program as measured by the Metropolitan Achievement Test. When IQ differences between the follow-up immersion kindergarten and the regular kindergarten children were equalized there was no difference in arithmetic or English language skills, but the immersion children did significantly better in French (Barik & Swain, 1972).

Arithmetic and Bilingualism

One fact that seems to emerge clearly from these studies of second-language immersion programs is that there is a trend for the immersion children to do significantly better in arithmetic than children in regular programs. This was found in both Grade 1 and Grade 2 of the Ottawa Separate School Board program and in Grade 2 and Grade 4 of the St. Lambert program. Several other studies have also examined the relationship between bilingualism and arithmetic-achievement. Trevino (1968) compared Spanish-speaking and English-speaking children enrolled in a bilingual English/Spanish primary school program in the U.S.A. with English-speaking monolinguals enrolled in a regular English program.

When compared to a control group in Grades 1 and 3, using the California Achievement Test, the bilinguals made superior scores on arithmetic fundamentals in Grade 1 and arithmetic reasoning in Grade 3.

Two other studies, though done many years ago, also showed bilinguals to make higher scores in arithmetic. Rigg (1926) compared bilingual children who spoke their native language at home with monolinguals in the St. Louis, U.S.A., school system and found the bilingual children made reliably higher scores on the Woody-McCall Arithmetic Fundamentals Test.

Manuel (1935) found similar results comparing English-speaking and bilingual Mexican children in Grades 2 to 8 in Texas. On the New Stanford Reading and Arithmetic Tests, the bilingual children had higher scores in reading and arithmetic in Grade 2 and in arithmetic in Grades 3, 4, and 6. However, the bilingual children were also grade for grade between 1.6 years and 2.9 years older than the monolingual children.

The evidence points to an apparent relationship between bilingualism and arithmetic-achievement. This raises the question as to whether the immersion children's facility with arithmetic is an isolated phenomenon, or whether it is concomitant with some other underlying cognitive abilities. This will be discussed in the next section.

3. Arithmetic and Cognition

The relationship between arithmetic and cognition can be approached through two questions. First, what is the relationship between arithmetic achievement and cognitive development; and secondly, if certain cognitive abilities are related to arithmetic, is there any evidence that they may be related to bilingualism?

Aiken (1971), in a review of studies of mathematics achievement and intellective variables, notes the positive relationship between mathematical ability and general intelligence. However, it is apparent that the performance of bilingual children cannot be explained by generally higher intelligence test-scores since they maintained their superiority in arithmetic even when intelligence was held constant (Lambert, 1970; Lambert *et al*., 1970, 1972; Edwards & Casserly, 1972). One must therefore look more closely at the cognitive abilities that underlie arithmetic and their relationship to the more specific aspects of cognitive functioning that may be enhanced by bilingualism. The first aspect to be discussed is the relationship between Piagetian concrete operations or logical thinking, and arithmetic. Secondly, the relationship between cognitive style and arithmetic will be discussed. Because cognitive maturity in one does not necessitate maturity in the other, separate hypotheses will be generated based on the aforementioned relationships to explain differences in arithmetic-ability between immersion and 75-minute children.

Arithmetic and Concrete Operations

Skemp (1961) has defined arithmetic as the "study of the property of numbers (as opposed to number), their interrelationships, the various operations which can be done on numbers and the results of those operations (p. 46)." The work of Jean Piaget (1965) shows that the development of the understanding of number is broad based, including the concepts of "ordination, cardination, the additive and multiplicative properties of numbers, and class inclusion and seriation (Peters, 1967, p. 295)." Children begin to acquire these operations at approximately seven years of age, the time when the immersion children were showing significantly higher arithmetic scores. It may be that the immersion children were more cognitively mature than the 75-minute children in terms of their development of concrete operational thinking.

A number of studies have examined the relationship between concrete operations and arithmetic-achievement. At this stage the child begins to be able to deal in a more logical and systematic fashion with the world, though he is still concrete in his thinking. His developing cognitive structures not only affect his ability to deal with number concepts, but change his ideas of such things as causality and morality and also broaden the linguistic structures that he is capable of using. Peters (1967) has pointed out that the development of number requires a variety of skills, not all of which develop at the same time. Moreover, the organization and processes of cognition in middle and late childhood can be seen as analogous to logico-mathematical thinking (Flavell, 1963).

Because children in the early grades of primary school are in the process of acquiring these cognitive structures, it is apparent that many children are able to perform simple arithmetic operations such as addition and subtraction in a rote fashion. Nevertheless, studies generally show a positive relationship between concrete operational thinking and arithmetic. For instance, Hood (1962) and Robinson (1968) gave a series of tests related to the development of number concepts such as seriation, numeration, classification and conservation to preschool and school-age children. Hood's subjects ranged in age from 4.9 years to 8.7 years while Robinson's subjects were all in Grade 1. Though both these authors found a wide variation in the age at which concrete operations were acquired, generally there was a significant relationship between achievement in arithmetic and performance on the Piagetian tests.

Two studies by Steffe (Van Engen & Steffe, 1966; Steffe, 1968) investigated the relationship between ability to conserve number and arithmetic-ability. Van Engen and Steffe (1966) tested Grade 1 children on their understanding of number conservation and addition in both a physical and a paper-and-pencil test. They found that though almost all of the children could correctly answer the paper-and-pencil questions, no more than 50% of the children could correctly answer the questions when faced with actual objects. They concluded that many children had simply memorized addition-facts rather than abstracting the concept of the sum of two whole numbers from the physical situation. In the second study, Steffe (1968) formed four groups of Grade 1 children on the basis of their level of conservation of number and found that the three top groups performed at the same level on tests of addition-facts and addition-problems, but the lowest conservation group made the poorest arithmetic scores. It seems therefore that concrete operational thinking may be more related to problem-solving arithmetic than mechanical arithmetic.

Other studies have found correlations ranging from .35 to .53 between arithmetic and Piagetian concrete operations (Elkind, 1961; Almy, Chittenden & Miller, 1967; Goldschmid, 1967), indicating that arithmetic-achievement rests partly, but not solely, on the understanding of number concepts.

Concrete Operations and Bilingualism

If children who make high scores in arithmetic are cognitively more mature in concrete operational thinking, is there any evidence that bilingualism might enhance such cognitive development? Two studies have directly examined the relationship between bilingualism and concrete operations. Liedtke and Nelson (1968) compared a group of bilingual Grade 1 children with a control group of monolinguals on conservation of area and length. The bilingual children had acquired two languages prior to entering school. The authors found that the bilingual children were reliably superior on both conservation tests and attributed this to their "cognitive flexibility."

The second study, (Ianco-Worral, 1972), whose findings were in the opposite direction, was done with English-Afrikaans bilinguals aged 4 - 6 years and 7 - 9 years. Though the main purpose of the study was to test Leopold's theory on the earlier separation of word-sound from word-meaning in bilinguals, the children were also given a classification test based on Piaget. No differences were found between the bilingual children and a matched group of monolinguals.

Though the evidence concerning the relationship of bilingualism to the development of concrete operational thinking is not clearcut, a number of studies that have just been considered in relation to intelligence and bilingualism indicate that cognition may be enhanced by the possession of two languages. This research (Peal & Lambert, 1962; Cohen, 1969; Balkan, 1970; Torrance et al., 1970, Lambert et al., 1972) showed that bilingualism appeared to contribute to more flexible cognitive processes. For instance, Peal and Lambert's (1962) study found bilinguals did better than monolinguals on tests of symbolic flexibility or concept formation. Cohen (1969) found bilingual children made higher scores on inductive reasoning and tests of permutation of elements. Balkan (1970) found that on tests of verbal and perceptual plasticity bilinguals did better than monolinguals. Finally, Lambert et al. (1972) and Torrance et al. (1970) showed that bilingual children made higher scores on a number of intellectual creativity tests.

If one can put the pieces of the jigsaw puzzle together, it does seem there is sufficient evidence to say that bilingual children have intellectual or cognitive structures which can be generally described as more flexible than those of monolinguals. It seems their thinking may be less rigid, and also that they may earlier acquire the ability to systematically think through a number of alternative solutions. It is possible there are a number of ways in which more flexible cognitive structures could enhance the attainment of concrete operations.

Concrete Operations and Cognitive Flexibility

One can consider that one of the problems encountered by the pre-operational child is his inability to deal with more than one change in relationship. That is, he may comprehend a change in one dimension, but cannot coordinate that with a change in another dimension. Such a child is tied to the perceptually obvious aspect of the situation, and his thinking is rigid. Some evidence that supports the role of more flexible thought processes in the attainment of concrete operations is provided by cross-cultural studies.

Though some studies have found that operations such as conservation appear at approximately the same time in a number of cultures (Goodnow & Bethon, 1966), there is also evidence that concrete operational thinking is affected by lack of schooling on such measures as conservation or combinatorial thinking (Goodnow & Bethon, 1966; Greenfield, 1966). It seems that schooling creates in a child a more abstract attitude, an analytical set forcing him away from the concrete and perceptually obvious and into conceptualizing events in his head. If one can equate such abstract or analytical thinking with cognitive flexibility, one has a rationale for proposing that bilingualism enhances concrete operations.

In summary, there is a trend for children enrolled in second language programs to make higher scores on arithmetic tests in comparison with children in regular programs. Two possible hypotheses have been suggested to account for this. The first hypothesis, discussed in the preceding section, is that bilingualism is related not directly to arithmetic-achievement, but rather to concrete operational thinking which in turn contributes to higher scores in arithmetic. The second hypothesis, which will be discussed in the following section, proposes that the immersion children's higher arithmetic scores are related to cognitive style.

4. Cognitive Style and Arithmetic

A number of studies have considered academic achievement not just in terms of variables such as intelligence, but in relation to individual problem-solving style. Research by Kagan and his colleagues (Kagan, Moss & Siegel, 1963; Kagan, Rosman, Day, Albert & Phillips, 1964; Kagan, 1965; Kagan, Pearson & Welch, 1966) indicates that developmental differences in cognitive performance may be due not only to differences in cognitive structures, but also to individual differences in visual scanning and impulsive versus reflective behaviour.

One task used by Kagan et al. (1963, 1964) to measure differences in cognitive style is the Matching Familiar Figures Test. Scores on this test allow one to divide children into reflective or impulsive. If the latency to first response is above the mean and the number of errors below the mean they are reflective; if latency of response is below the mean and number of erros above the mean they are impulsive. Kagan et al. (1964) have defined reflection as "the consideration of alternative solution hypotheses (either classification or problem-solving sequences) when many alternatives are available simultaneously (p. 33)."

Kagan conjectures that one of the basic determinants of conceptual reflection versus impulsivity may be found in the early activity and attentional behaviour of the infant. In older children, he has examined classroom behaviour in terms of the ratio of time spent attending to a task divided by the time spent distracted. He found a positive relationship between amount of time spent attending and conceptual reflection. The ratio reliably discriminated impulsive from non-impulsive children, and he comments that the classroom behaviour of the children paralleled that found in their problem-solving behaviour. "The reflective child seemed capable of greater concentration; they showed less distractibility during the conceptual task (Kagan, 1964, p. 30)."

Spitler (1971) has attempted to develop a model of cognitive style and relate it to the teaching and learning of mathematics. She feels that while arithmetic computation requires impulsiveness, arithmetic problem-solving and reasoning require reflection. Cathcart and Liedtke (1969), using the M.F.F., have found a relationship between reflection and arithmetic performance. They gave an arithmetic test, two conservation tests and the M.F.F. to 46 Grade 2 children and 12 grade 3 children. There was no statistically significant difference between intelligence or cognitive style for these children, though there was a tendency for the more intelligent child to be more reflective. There was also no significant interaction between cognitive style, sex, grade, intelligence and conservation, but when the means of the arithmetic tests were examined, the reflective group was found to have significantly higher scores. This relationship held when sex, intelligence, and conservation performance were held constant.

Arithmetic and Attention

As noted previously, the dimension of reflection/impulsivity parallels ratings of attention in the classroom. There is evidence that classroom ratings of attention are also strongly related to achievement in arithmetic, adding support for the findings of Cathcart and Liedtke (1969).

Cobb (1972) rated fourth grade children on 14 classroom behaviours such as attention, talking to teacher positively, volunteering information, play, etc. Of these 14, attention was found to be the most powerful predictor of arithmetic-achievement on the Stanford Achievement Test. Lahaderne (1968) found that attention and arithmetic-achievement were significantly correlated in four sixth grade classrooms. When intelligence was partialled out, the correlation remained significant for boys, but not for girls. Finally, Attwell, Orpet and Meyers (1967) rated one hundred kindergarten children on a number of classroom behaviours and related these scores to scores attained on the California Achievement Test given to the same children on reaching Grade 5. For the 57 children left from the original sample, they found that attention was the only behavioural rating to predict each of the six areas of the C.A.T. Of the ratings that had significant correlations with arithmetic-reasoning and arithmetic-fundamentals, attention was the highest.

The relationship between arithmetic and attentive or reflective behaviour versus impulsiveness or distractable behaviour seems fairly well established. However, the author has not come across any study which directly relates such aspects of cognitive style to second-language acquisition, though there are a number of pieces of indirect evidence to indicate that second-language acquisition may be related to reflective/attentive behaviour and these will be considered next.

Bilingualism and Attention

Stevens (1966) asks what is the child doing "in operational terms, as he acquires a second language? What kinds of behaviour are being reinforced? What behaviours are being practiced (p. 78)?" He concludes that exposure to a second language at an early age will enhance the ability to receive and process incoming stimuli. He feels there is evidence that increased practice in attending should increase the child's capacity for attending, and this in turn should lead to increased cognitive gains. In support of his hypothesis, Stevens (1966) looked at the possible effect of bilingualism on psycholinguistic ability and found partial confirmation in his results for the idea that exposure to a second language, in this case French-immersion at the kindergarten and Grade 1 level, results in cognitive gains.

There are a number of other pieces of evidence from the Ottawa Separate School Board program which support Stevens' findings. As in Stevens' study, psycholinguistic ability was measured using the Illinois Test of

Psycholinguistic Abilities. The pilot immersion children were superior to the 75-minute children only on the grammatical closure subtest, but the follow-up immersion children were superior on all four subtests, auditory association, auditory reception, auditory memory, and grammatical closure (Edwards & Casserly, 1971, 1972). If Stevens' hypothesis is correct, increased ability to attend underlies the superior performance of the immersion children on this test.

There are two other tests used by the School Board which, though again not measuring attention directly, would seem to require as a major component an ability to attend and be free from distraction. The Slingerland Screening Test for Specific Language Disability was given to both pilot and follow-up 75-minute and immersion Grade 1 children. Both the pilot and follow-up immersion Grade Ones were superior to the 75-minute children on combined score for self-corrections, poor formations and total errors. However, this difference disappeared for the pilot immersion children when compared to the 75-minute group matched on age, sex, and kindergarten background. On the Pupil Behaviour Rating Scale, both the French and English teachers rated the pilot Grade 1 immersion group superior on auditory comprehension and listening; the French teacher also rated them as superior on the behaviour subtests (which also included a rating of attention). In the follow-up Grade 1, significant differences were found in favour of the immersion children on all subtests for both languages. Some of these differences disappeared on the matched-group comparisons for the English version, but not for the French version.

Though the evidence is somewhat indirect, there does seem to be indications that the immersion program has given the immersion children practice in attending and this may underlie their superior achievement in arithmetic.

In summary, this section has related arithmetic-achievement to cognitive style or the manner in which children solve problems. It has been hypothesized that the higher arithmetic scores of the immersion children are related to the dimension of reflection or their ability to bring sustained attention to bear on a task.

5. Summary and Hypotheses

The present study was intended to examine the performance of English-speaking children enrolled in a French immersion program in comparison with children enrolled in a regular English program receiving 75 minutes of French per day on measures of arithmetic-achievement, Piagetian concrete operational thinking, and cognitive style.

Evidence has been presented to indicate that children enrolled in second-language immersion programs do significantly better on standardized arithmetic tests in comparison with children enrolled in regular programs. Two possible explanations were suggested to account for the higher arithmetic scores of the immersion children. The first explanation suggested that bilingualism enhanced "cognitive flexibility," and this manifested itself in greater cognitive maturity in logical or concrete operational thinking and thus the children's understanding of number concepts. The second explanation suggested that the immersion children were more reflective or attentive and this facilitated arithmetic-performance.

On the basis of the evidence presented in the review of the literature, the major part of the study will examine the relative performance of immersion versus 75-minute children on arithmetic-achievement and cognition. However, the study will also try to answer the question which cognitive measures best predict arithmetic-performance, per se, for the two groups as a whole.

With these two aims in mind, the following null hypotheses are presented:

1. There will be no significant difference between the mean score of the immersion children and the mean score of the 75-minute children on the arithmetic computation and concepts subtests and on the composite arithmetic score of the Metropolitan Achievement Test.

2. There will be no significant difference between the mean scores of the immersion children and the mean scores of the 75-minute children on the following tests of concrete operational thinking: conservation of substance, all and some, seriation, numeration, conservation of weight, classification of squares and circles, multiple seriation; or on the total score for these tests combined.

3. There will be no significant difference between the error score and average latency to first response on the Matching Familiar Figures Test for the immersion children in comparison with the 75-minute children.

4. If program, immersion versus 75-minute, is used as a moderator variable (Zedeck, 1971) in a stepwise multiple regression equation with arithmetic achievement as the criterion and measures of concrete operational thinking and cognitive style as predictors, there will be no significant difference between the multiple R for the immersion program and the multiple R for the 75-minute program.

5. With program used as a moderator variable in a stepwise multiple regression, there will be no significant difference between the separate multiple Rs for each program and the multiple R for the total group.

6. If a stepwise multiple regression is computed for the total group (N=116) and for two randomly chosen subgroups (n=58), there will be no significant difference between the multiple R for the total group and the multiple Rs for each of the subgroups compared to the total group and to each other.

7. If the B weights and constant from each random subgroup are applied to the scores of the other random subgroup and multiple Rs computed, there will be no significant difference between the cross validation multiple Rs compared to each other and to the multiple R for the total group.

CHAPTER II

RESEARCH DESIGN

This is a descriptive study designed to provide information regarding the differences between a French immersion program and a regular English program on measures of arithmetic-achievement and related cognitive indices. This chapter presents a description of the subjects, followed by a presentation of the measuring instruments and the procedure for carrying out the testing. The last section of the chapter discusses the statistical analysis of the data.

1. Subjects

The subjects in this study were 116 Grade 1 children at three schools under the jurisdiction of the Ottawa Roman Catholic Separate School Board; 38 of these children were enrolled in the regular English program receiving 75 minutes of French per day, while the remaining 78 were enrolled in an immersion program where approximately 70% of the instruction was in French. At this point in time the immersion children were receiving no instruction in English Language Arts. Instruction in arithmetic was through the medium of French. Of the 38 children enrolled in the 75-minute program, 14 were females and 24 males. In the case of the immersion children, 43 were females and 35 males. The average age for the immersion and 75-minute groups was 74.1 months and 74.7 months, respectively. There was no significant difference in age between the two groups.

The three schools were chosen specifically to obtain immersion and 75-minute classes that were comparable in socio-economic background. To this end, all the three schools finally chosen were from similar middle-class areas. However, it should be noted that since not every school in the Separate School Board system has an immersion class, children enrolled in the program tended to be drawn from a wider catchment area than would be the case with children in the 75-minute classes. Two other factors governed the choice of these schools. The first, and most important, was the availability of an area for individual testing. Because of the shortage of space in many schools, this was not easily obtained. The second factor was the necessity of not conflicting with longitudinal testing that was being carried out in a number of schools as part of the overall assessment of the French language program.

In those schools with immersion programs, the number of children in the 75-minute classes tended to be smaller, and this was so in all three schools selected. In the case of School I, there were 31 immersion children and 21 75-minute children. In School II, there were 24 children in the immersion class and eight in the 75-minute class. In School III, there were 24 children in the immersion class and 9 in the 75-minute class. One child in the School III immersion class was not tested because of a lengthy period of absence from school due to sickness; otherwise, all the children were tested.

The disparity between the sizes of the immersion and 75-minute classes in Schools II and III was compensated for by the fact that in both schools Grades 1 and 2 were combined; so, in fact, the ratio of children to teacher was similar for both immersion and 75-minute classes.

2. Program

At the time that this study was carried out, the revised program of French instruction providing children with either immersion or 75 minutes of French per day was in its third year of operation in the Separate School Board system. Because Ottawa is the capital of a bilingual country and has residents speaking both official languages, it has felt increasingly the need to produce a bilingual population. It was in this context that the Separate School Board revised their French language program to provide French immersion at the primary level.

For both the immersion and 75-minute children, French is taught in an eclectic manner. The 75-minute children receive Art, Music, Social Studies and, of course, Second Language Instruction through the medium of French.

The immersion children receive all but one of their major subjects--Art, Music, Mathematics, Physical Education, Science, and Social Studies--in French. Only Religious Education is given in English. The immersion children have normal Second Language Instruction which includes such content as phonation exercises, basic linguistic structures and pattern drills. In addition they receive regular instruction in French Language Arts which combines a global, phonic and syllabication approach to reading. At this stage there is no instruction in English Language Arts (Game, 1972).

Though both the 75-minute and immersion children receive instruction in the major subject areas using both written and oral presentation, there is certainly, initially, somewhat more emphasis on oral presentation for the immersion children.

With few exceptions, such as transfers from other school boards, children enrolled in the Separate School Board's Grade 1 immersion and 75-minute programs have attended half-day French, half-day English kindergarten at the five-year-old level. Thus their contact with the French language is presumed to be similar at the beginning of Grade 1.

3. Measuring Instruments

The following section describes the measuring instruments used in this study, the apparatus employed, and the scoring system. The first half describes the measuring instruments used to obtain background information on the children, followed by a description of the instruments used to test the null hypotheses.

(a) Metropolitan Readiness Test. This test is administered by the Separate School Board to all children prior to entering Grade 1 who are present on the day of testing. Test scores were made available to the author by the School Board. Of the 116 children in the immersion and 75-minute programs combined, scores were available on 99. The other 17 children were either absent on the day of testing, repeating a grade, or transfers from another school board. The 17 were approximately evenly divided between the immersion and 75-minute groups. Where test scores were not available, the mean score from either the immersion or 75-minute group was used in the statistical analysis.

The M.R.T. consists of six subtests and a total score. The subtests are word meaning, listening, matching, alphabet, numbers, and copying. However, for the purposes of this research only the number subtest and total score were used. These two scores were chosen for a number of reasons. Since the object of the study was to examine children's arithmetic-ability and understanding of number, the number subtest was an obvious choice to use as a co-variate. In addition, several studies have shown that the number subtest score and the total score are the best predictors of first grade achievement (Hagenson, 1968; Akers, 1970). In addition, Scott (1969) has found a significant correlation between the number subtest and the Piagetian tests of seriation.

(b) Draw-A-Man Test. It was decided to administer a short, non-verbal measure of intelligence in order that any differences between the two groups could be statistically controlled. The Draw-A-Man Test was considered appropriate for the purpose because of its length, and also because of its non-threatening nature and appeal to children. Harris (1963) reports reliability correlations ranging from .60 to .86 with intervals between drawings of up to three months. Interscorer reliability commonly exceeds .90 (Harris, 1963).

The test was administered as a group test, each child being given a blank sheet of paper and asked to "Draw a man, draw the best man you can." There was no time limit on the test. Scoring was in accordance with the instructions in the revised manual (Harris, 1963). Each drawing was scored independently by the author and another scorer after consideration of the scoring criteria given by Harris. Interscorer correlations were .97 for the immersion group and .98 for the 75-minute group.

(c) Social Class. Because information concerning a child's social background, such as parents' education and occupation, is not now available to either individual schools or the School Board, it was decided to try and obtain this information through a questionnaire mailed to parents. An example of the questionnaire and covering letter can be found in Appendix 1.

The questionnaire was divided into two halves. The first half requested the occupation of both the mother and father; the second half listed levels of education ranging from primary school through post-graduate work and requested both parents to indicate the highest level of education attained.

The questionnaire was sent with a covering letter stressing that its completion was completely voluntary and anonymous. Approximately ten days after the questionnaire was mailed, a follow-up telephone call was made to each child's parents asking that they complete and return the questionnaire. Of the questionnaires mailed, 67, or 57%, were returned. Fourteen of these were from parents of 75-minute children and 54 from parents of immersion children. This disparity in returns presumably reflects the greater interest in the French language program of the School Board on the part of parents of immersion children.

A separate score was computed for occupation and education which was then added to provide a total score. Education was divided into ten levels, two points being allowed for each level completed. The educational level of each parent was scored separately and the two scores averaged. The distribution by educational level for the 75-minute and immersion parents can be seen in Table 1.

Occupations were assigned to one of seven different classes according to the scale devised by Blishen (1958). The points allotted for each occupational level ranged from two for those in Class 7, the lowest class, to 14 for those in Class 1, the highest class. If both parents were employed, their scores were averaged in the same manner as for education. The distribution by occupation for immersion and 75-minute groups can be seen in Table 2. The final score for social class was the total of the averaged scores for education and occupation.

The measuring instruments used to test for academic achievement and cognitive differences between the immersion and 75-minute children will be discussed in the following paragraphs.

Table 1

Levels of Education of Parents - 75-Minute and Immersion Groups

	75-minute		Immersion	
	Father	Mother	Father	Mother
Grades 1 - 8	2	3	5	2
Grades 9 - 10	2	1	8	8
Grades 11 - 12	3	3	9	21
Grade 13		2	6	2
Partial Community College	1		3	6
Completed Community College		3	2	6
Partial B.A.	3		3	4
Completed B.A.		1	8	1
Partial Post Graduate			4	2
Completed Post Graduate	1		4	1
Total	12	13	52	53

Table 2

Occupational Level of Parents - 75-Minute and Immersion Groups

	75-minute		Immersion	
	Father	Mother	Father	Mother
Class I			3	
Class II	2	2	20	5
Class III	2	1	10	9
Class IV			5	4
Class V	5		7	3
Class VI	2	1	5	1
Class VII	1	1	2	
Total	12	5	52	22

(d) Metropolitan Achievement Test Primary Battery I. This test consists of five subtests, word knowledge, word analysis, reading, arithmetic concepts and arithmetic computation, of which only the last two, arithmetic concepts and computation, were administered. Scores were computed for each subtest and also added together to provide a composite score. All three raw scores were used in the statistical analysis.

The M.A.T. was chosen for a number of reasons. A major reason for selecting this particular instrument was its use in the three longitudinal studies being carried out on second-language immersion programs in Montreal and Ottawa, thus allowing comparison to be more easily made between the findings of this study and those of previous research.

The M.A.T. also provides separate sections for concepts and computation. Because some studies show that bilingualism affects only problem-solving or arithmetical concepts (MacNamara, 1966), it is important to have separate scores for these two areas.

The content of the two subtests, concepts and computation, covers those areas taught in the Grade 1 curriculum of the Separate School Board. The first section, concepts, is presented verbally by the examiner; thus it is independent of reading skill. This, of course, is particularly important when testing the French immersion children who do not get any instruction in English Language Arts. The manual reports reliability coefficients above .90 for the two arithmetic subtests combined.

(e) Matching Familiar Figures Test. This test consists of a standard picture and six alternatives, one of which is identical to the standard. From the six alternatives, the subject must pick out the picture that is identical in every respect to the standard. There are 12 such matching tasks in the M.F.F., but this was considered too long for subjects in Grade 1, therefore only the first six tasks were administered.

Each subject is timed to the nearest half second from the presentation of the standard to his first choice. If the subject chooses incorrectly, he is told to try again. After six incorrect responses, the next task is presented. Full instructions for the administration of the M.F.F. are given in Appendix 2. Two scores are computed, the average time to first choice for all six tasks and the total number of errors.

(f) Tests of Concrete Operational Thinking. Seven Piagetian tests were administered individually to each child. The tests chosen were those which appeared to tap most closely the child's understanding of number. According to Piaget (1965), the construction of the concept of number requires the synthesis of the operations of classification and seriation. Thus a range of tests which tapped both these abilities was administered. The purpose of administering as many as seven tests was twofold. First, the range of tests would give one a better idea of the strengths and weaknesses of the immersion children in comparison with the 75-minute children. Secondly, the literature does not indicate which of the Piagetian tests are the best predictors of arithmetic-achievement at this age level, and this was a secondary purpose of this study.

Generally speaking, the format followed in the test administration was not taken directly from Piaget, but from a variety of authors who have standardized Piaget's rather flexible questioning technique. The source of the format for each of the seven tests is indicated. The full test schedule can be found in Appendix 3.

(i) Conservation of Substance. This test examined the child's ability to conserve substance in the face of a shape transformation. The format used was based on Rothenberg (1969). For each change in shape the subject was asked two questions; one concerning the equality of the two pieces of Play-Doh, and one concerning their difference. In addition, the subject was required to explain his answer to each question.

(ii) <u>All and Some</u>. This test examined the child's understanding of class inclusion and use of the words "all" and "some." The format used was taken from Kofsky (1968) and involved four questions regarding the classification of nine blocks which differed by colour and shape. For instance, the subject had to differentiate between questions such as, "Are all the triangles red?" versus "Are all the reds triangles?"

(iii) <u>Seriation</u>. In this test the child was required to seriate a set of ten slats which increased in size in half-inch steps. He was then required to insert another set of nine slats into the original set. This second set differed in size from the original set by a quarter of an inch. The format was taken from Elkind (1964).

(iv) <u>Numeration</u>. This test examined the child's ability to coordinate ordinal and cardinal numbers using the ten graded slats from the previous test and a plastic doll. The format of the test was also taken from Elkind (1964).

(v) <u>Conservation of Weight</u>. The format of this test was identical to that of conservation of substance, except the child was examined on his understanding of weight in the face of a shape transformation.

(vi) Classification of Squares and Circles. This test looked at the child's ability to form hierarchical classes and to deal with a class containing only one element. The test required the child to form a series of two groups out of wooden blocks that differed in shape, size and colour. The format was taken from Lovell, Mitchell and Everett (1968).

(vii) Multiple Seriation. This test required the child to coordinate the operations of seriation and classification by forming a matrix out of 16 blocks of four different shades of green and four different sizes. The format was taken from Lovell, Mitchell and Everett (1968).

The subject's answers on the seven Piagetian tests were scored in two ways. Initially, using the criteria described by Piaget (Piaget, 1965; Piaget *et al.*, 1970), for each test the subject was assigned to one of three stages on the basis of his answers to the questions. Those subjects assigned to Stage I gave no indication of concrete operational thinking. Stage II children were transitional. That is, they exhibited some indication of concrete operational thinking but their answers were inconsistent such as giving conservation answers to some questions; when pushed their ability to think logically broke down. Finally, Stage III children manifested concrete operational thinking consistently in their answers within each test.

In addition, a numerical score was assigned to each test. For each correct answer on each test, a score of 1 was assigned. The scores for the tests ranged from a maximum of 4 for each of the conservation tests to 14 for the numeration test.

All the answer sheets were initially scored by the author. Twenty-five percent of the total was then picked at random and rescored independently by the author and another scorer. Interscorer correlations for each subtest can be found in Table 3.

4. Procedure

The study was carried out in two phases. The first phase involved the administration in May of the individual tests, concrete operational thinking and cognitive style; followed in June by the administration of the group tests, Metropolitan Achievement Test and Draw-A-Man Test.

(a) Individual Testing. Because of space shortages in all the schools, it was necessary for both experimenters to test in the same room. However, in the case of Schools II and III these were very large rooms, and in the case of School I testing was done at either end of the hallway. It did not appear that the children were distracted by the other testing in progress or could hear the answers of the other subjects.

Table 3

Interscorer Correlations for Concrete Operational Thinking Tests

Test	r
Substance	.90
All and Some	.94
Seriation	.99
Numeration	.99
Weight	1.00
Squares and Circles	.97
Multiple Seriation	.99

The individual testing involved the administration of the seven Piagetian concrete operational tests and the Matching Familiar Figures Test. Each child was seated at a table across from E, the testing materials being placed on another table or chair behind E. Each child was tested in two approximately half-hour sessions, one of which took place before recess and the other after recess. Half the children in each class were tested in the morning and half in the afternoon. All the testing was carried out by the same two experimenters, the author and an assistant. Each tester examined the same number of children from each class, evenly divided between morning and afternoon and balanced as nearly equally as possible by sex. Assignment of subjects to tester and time of testing session was done randomly.

Because the number of tests involved precluded administering them in every combination of order, the order of administration was the same for each child. This was as follows: conservation of substance, all and some, seriation, numeration, conservation of weight, classification of squares and circles, multiple seriation, and the Matching Familiar Figures Test.

(b) Group Testing. The group testing was carried out by the author in June during normal classroom time. Prior to the administration of the two subtests of the M.A.T., the children were given the practice items to ensure that they understood how to mark their answers. Administration of the two subtests was done in accordance with the instructions given in the M.A.T. manual. After the administration of the M.A.T., the children were given the Draw-A-Man Test. No time limit was set for this test. However, no child took longer than 15 minutes.

5. Statistical Analyses

This section outlines the statistical analyses used to test for significant differences between the 75-minute and immersion subjects. The first section discusses the statistical analyses of the three measures used to obtain background information on the children. This is followed by presentation of the statistical analyses used in testing hypotheses 1 to 7.

In order to determine if there were any significant differences between the two groups on the Metropolitan Readiness Test, Draw-A-Man Test, and social class rating, a separate, one-way, fixed effects analysis of variance was carried out on each of the measures as follows.

(a) Metropolitan Readiness Test. Raw scores from the number subtest, and the total score for all six subtests combined, were each analyzed in a one-way, fixed effects analysis of variance for the total sample ($N = 116$). In addition, to equate the two groups, 14 females and 24 males were randomly picked from the immersion group using a table of random numbers and an analysis of variance computed for the number and total scores with equal Ns. Though it appears that the F-test is quite robust with respect to violation of homogeneity of variance if Ns are equal, it was decided to apply Hartley's F-Max test to both the variances for the equal and unequal Ns (Kirk, 1968). In neither case was it necessary to reject the null hypothesis of no difference in homogeneity of variance between the two groups using a p level of .05.

(b) Draw-A-Man Test. Two scores were used from the D.A.M., raw scores and standard scores. The raw scores were converted to standard scores using the conversion table in the manual (Harris, 1963). Separate one-way, fixed effects analyses of variance were computed for the raw scores and for the standard scores for equal and unequal Ns in the same manner as for the Metropolitan Readiness Test. The F-Max test applied to raw scores and standard scores for equal and unequal Ns indicated that there was no significant difference in homogeneity of variance between the two groups.

(c) Social Class Rating. The total of the averaged scores for education and occupation was analyzed using a one-way, fixed effects analysis of variance for both the total sample and for the equal N sample. Application of Hartley's F-Max test revealed no significant differences in homogeneity of variance between the two groups for either equal or unequal Ns.

In order to test null hypotheses 1 to 3, multivariate analyses of variance were computed first using the composite score on the Metropolitan Achievement Test, the total score for the seven Piaget subtests, and the error and latency scores on the M.F.F.; and second using the concept and computation subtest scores on the M.A.T., the seven Piaget subtest scores and the two scores on the M.F.F. These two multivariate analyses of variance yielded both a univariate F for each of the dependent variables as well as a multivariate F.

For each set of scores, a multivariate analysis of variance was carried out for both the total sample ($N = 116$) and for the equalized sample ($N = 76$). For each set of scores, Hartley's F-Max test was applied to both equal and unequal Ns. In only one case, for the scores on the seriation test, was it found necessary to reject the null hypothesis of no significant difference in homogeneity of variance between the two groups. The null hypothesis was rejected for both equal and unequal Ns. However, Kirk (1968) has stated that the F-test is robust with respect to violation of this assumption if Ns are equal, and it was therefore considered safe to carry out the analysis of variance on this test.

The univariate analyses of variance for the Metropolitan Readiness Test, Draw-A-Man and Social Class Rating were carried out using a SOUPAC analysis of variance program; the multivariate analysis of variance employed a MANOVA and ANOVA program.

On the seven tests of concrete operational thinking, in addition to the analysis of variance, a chi square test was also applied to the data. The children were divided into two groups within each program, those in Stages I and II versus those in Stage III and a 2 X 2 χ^2 carried out.

In order to test null hypothesis 4 of no significant difference between the stepwise multiple Rs for the immersion and 75-minute programs, separate multiple Rs were computed for each group. Each multiple R was then converted to a z-score and the standard error of the difference between two zs obtained using a table of t with p set at .05 (Guilford, 1965).

To test null hypothesis 5 required the computation of a stepwise multiple regression initially on the total sample (N = 116). This multiple R was then converted to a z score. Comparison between the total multiple R and the multiple R for the immersion and 75-minute group separately was made using the following formula (Dayhaw, 1969; McInnis, 1972):

$$Rs = \frac{1 - R_t^2}{\sqrt{N_s - 1}} \sqrt{\frac{N_t - N_s}{N_t - 1}}$$

where R_t = multiple R for the total group;

N_t = number in the total group;

N_s = number in the subgroup.

This formula gives one the standard error of the difference between a multiple correlation coefficient of a subgroup and the multiple correlation coefficient of a total group containing the subgroup (McInnis, 1972).

In order to test null hypothesis 6, it was necessary to divide the total sample into two equal random subgroups using a table of random numbers. Each subgroup had N = 58 subjects. Separate stepwise multiple Rs were computed for each random subgroup. These multiple Rs were then converted to z scores and the standard error of the difference between two zs computed to test for significant differences between the two multiple Rs. The two random subgroup multiple Rs were then compared separately to the total multiple R using the formula given to test null hypothesis 6 for the standard error of the difference between a multiple correlation coefficient of a subgroup and the multiple correlation coefficient of a total group containing the subgroup.

In order to test null hypothesis 7, it was necessary to cross-validate the multiple Rs obtained on each of the two subgroups. In order to do this, the B weights, plus a constant, from the stepwise multiple regression of one random subgroup were applied to the scores of the other random subgroup and vice versa using the formula $X' = B_1X_1 + B_2X_2 \ldots\ldots B_nX_n + C$ (Guilford, 1965). B weights or unstandardized weights, plus a constant, were used rather than beta weights because they were being applied to raw scores. The new predicted scores for each subgroup were then correlated with the actual arithmetic scores to provide a new multiple R.

To test for differences between the cross-validated multiple Rs each was converted to a z-score and the standard error of the difference between two zs computed (Guilford, 1965). In addition, the difference between each multiple R and the total multiple R was calculated using the formula for the standard error of the difference between a multiple correlation coefficient of a subgroup and the multiple correlation coefficient of a total group containing the subgroup. As well as testing separately the difference between the total group multiple R and the cross-validation groups multiple R, a comparison was made between the total group multiple R and the average of the two cross-validation multiple Rs by averaging the two z-scores for the two cross-validation multiple Rs.

All the stepwise multiple Rs were computed using SOUPAC Stepwise Multiple Regression program.

6. Additional Analyses

In order to statistically control differences that might exist between the immersion and 75-minute groups, it was decided to use the number and total score from the Metropolitan Readiness Test and the raw score from the Draw-A-Man Test as covariates in a multivariate analysis of covariance. Originally it was also intended to use social class as a covariate, but because of the small number of returns, especially for the 75-minute group, this was not considered feasible.

Because the number score of the M.R.T. forms part of the total M.R.T. score, these two scores were used separately as covariates, but each in conjunction with the raw score on the Draw-A-Man Test. Thus each dependent variable was used twice in an analysis of covariance, once with M.R.T. number score and Draw-A-Man raw score as covariates, and the second time with M.R.T. total score and Draw-A-Man raw score as covariates. Correlations between each of the covariates and each of the dependent variables are shown in Tables 4, 5, and 6.

As was the procedure for the multivariate analysis of variance, the multivariate analysis of covariance was

Table 4

Correlations between Scores on the Metropolitan Readiness Test - Total Score and Dependent Variables Used in an Analysis of Covariance

Dependent Variable	75-minute (N=38)	Immersion (N=38)	Immersion (N=78)	Total (N=76)	Total (N=116)
M.A.T. Total	.70***	.69***	.63***	.69***	.65***
M.A.T. Computation	.62***	.61***	.49***	.61***	.54***
M.A.T. Concepts	.69***	.65***	.63***	.68***	.65***
Concrete Ops. Total	.60***	.53***	.61***	.58***	.61***
Substance	.45***	.35*	.46***	.41***	.46***
All & Some	.29	.15	.22	.23*	.24*
Seriation	.45***	.19	.46***	.36***	.46***
Numeration	.58***	.41**	.43***	.51***	.46***
Weight	.53***	.14	.23*	.36***	.34***
Squares & Circles	.29	.44***	.44***	.36***	.39***
Multiple Seriation	-.00	.36*	.38***	.14	.25***
M.F.F. Error	-.48***	-.29	-.26*	-.36***	-.33***
M.F.F. Latency	.13	.20	.22	.15	.18

* $p < .05$
** $p < .02$
*** $p < .01$

Table 5

Correlations between Scores on the Metropolitan Readiness Test - Number and Dependent Variables Used in an Analysis of Covariance

Dependent Variable	75-minute (N=38)	Immersion (N=38)	Immersion (N=78)	Total (N=76)	Total (N=116)
M.A.T. Total	.58***	.60***	.53***	.59***	.55***
M.A.T. Computation	.53***	.63***	.50***	.57***	.50***
M.A.T. Concepts	.55***	.49***	.49***	.53***	.51***
Concrete Ops. Total	.50***	.32	.37***	.44***	.41**
Substance	.39*	.11	.31*	.27*	.33***
All & Some	.31	.10	.27*	.21	.28***
Seriation	.37*	.10	.43***	.26*	.41***
Numeration	.50***	.29	.41***	.41***	.43***
Weight	.42***	-.00	.16	.22	.24*
Squares & Circles	.13	.28	.36***	.22	.30***
Multiple Seriation	.07	.34*	.34***	.19	.26***
M.F.F. Error	-.48***	-.32	-.18	-.36***	-.26***
M.F.F. Latency	.28	.11	.23	.19	.24*

* $p < .05$
** $p < .02$
*** $p < .01$

Table 6

Correlations between Scores on the Draw-A-Man Test and Dependent Variables Used in an Analysis of Covariance

Dependent Variable	75-minute (N=38)	Immersion (N=38)	Immersion (N=78)	Total (N=76)	Total (N=116)
M.A.T. Total	.33*	.22	.29*	.27*	.27***
M.A.T. Computation	.30	.13	.21	.21	.20*
M.A.T. Concepts	.32	.27	.30***	.30***	.29***
Concrete Ops. Total	.40*	.51***	.47***	.47***	.44***
Substance	.35*	.31	.08	.34***	.14
All & Some	.16	.04	-.16	.12	-.10
Seriation	.37*	.34*	-.09	.39***	.07
Numeration	.26	.34*	.08	.32***	.11
Weight	.39*	.37*	.16	.38***	.21*
Squares & Circles	.26	.40*	.09	.36***	.13
Multiple Seriation	-.02	.14	.03	.05	-.00
M.F.F. Error	-.34*	-.19	-.35***	-.24*	-.31***
M.F.F. Latency	.21	.31	.00	.22	.02

* $p < .05$
** $p < .02$
*** $p < .01$

carried out separately first on the total score of the M.A.T. , the total score of the Piaget tests, and the two scores from the M.F.F.; secondly, on the subtest scores of the M.A.T., the seven Piaget subtests and the two M.F.F. scores. Each of the multivariate analyses of covariance also yielded univariate Fs for each dependent variable.

All the multivariate analyses of covariance were carried out using a MANCOVA and ANOVA program.

In order to examine the pattern of intercorrelations between the test scores, a matrix of intercorrelations was computed for each subject group separately.

The review of the literature had indicated that there might be a relationship between children classified as reflective on the Matching Familiar Figures Test and arithmetic-ability. Therefore it was decided to examine this relationship, as well as the relationship between reflection/impulsivity and concrete operational thinking. On the basis of his score on the M.F.F., each subject was assigned to one of four groups as follows: those with below average error scores and below average latency scores (fast-accurate); those with below average error scores and above average latency scores (reflective); those with above average error scores and below average latency scores (impulsive); finally, those with above average error scores and above average latency scores (slow-inaccurate). These four groups were used as independent variables in a simple one-way analysis of variance using as dependent variables the two arithmetic subtest scores on the M.A.T., the composite arithmetic score on the M.A.T., the scores on the seven concrete operational thinking subtests, plus the total of the seven concrete operational scores.

For all statistical analyses, a p level of .05 was used.

CHAPTER III

RESULTS

This chapter begins by presenting the results of the analyses of variance carried out on the Metropolitan Readiness Test, the Draw-A-Man Test, and Social Class Rating. This is followed by a presentation of the results of the major hypotheses 1 to 7.

(a) Metropolitan Readiness Test. Tables 7 to 10 present the results of the analyses of variance of the total scores and number subtest scores. On neither of these were any significant differences found between the immersion and 75-minute groups.

(b) Draw-A-Man Test. From Tables 11 to 14, it can be seen that there were highly significant differences for both raw scores ($F = 9.4009$, $p < .01$ $n = 76$; $F = 7.1342$, $p < .01$ $n = 116$) and standard scores ($F = 12.6931$, $p < .01$ $n = 76$; $F = 11.0958$, $p < .01$ $n = 116$), both in favour of the immersion group.

(c) Social Class Rating. No significant differences were found between the two groups on social class. These results are shown in Tables 15 and 16.

Table 7

Analysis of Variance Summary Table:
Metropolitan Readiness Test -
Total Score
(N=76)

Source	SS	df	MS	F
Between Groups	428.13	1	428.13	3.0945
Within Groups	10,239.95	74	138.38	
Total	10,668.08	75		

Table 8

Analysis of Variance Summary Table:
Metropolitan Readiness Test -
Number Subtest
(N=76)

Source	SS	df	MS	F
Between Groups	22.23	1	22.22	1.9386
Within Groups	848.43	74	11.47	
Total	870.65	75		

Table 9

Analysis of Variance Summary Table:
Metropolitan Readiness Test -
Total Score
(N=116)

Source	SS	df	MS	F
Between Groups	68.22	1	68.22	0.4525
Within Groups	17,185.02	114	150.75	
Total	17,253.24	115		

Table 10

Analysis of Variance Summary Table:
Metropolitan Readiness Test -
Number Subtest
(N=116)

Source	SS	df	MS	F
Between Groups	1.04	1	1.04	0.0778
Within Groups	1,526.60	114	13.39	
Total	1,527.64	115		

Table 11

Analysis of Variance Summary Table:
Draw-A-Man Test - Raw Score
(N=76)

Source	SS	df	MS	F
Between Groups	345.32	1	345.32	9.4009***
Within Groups	2,718.18	74	36.73	
Total	3,063.50	75		

* $p < .05$
** $p < .02$
*** $p < .01$

Table 12

Analysis of Variance Summary Table:
Draw-A-Man Test - Standard Score
(N=76)

Source	SS	df	MS	F
Between Groups	2,901.63	1	2,901.63	12.6931***
Within Groups	16,916.32	74	228.60	
Total	19,817.95	75		

* $p < .05$
** $p < .02$
*** $p < .01$

Table 13

Analysis of Variance Summary Table:
Draw-A-Man Test - Raw Score
(N=116)

Source	SS	df	MS	F
Between Groups	533.06	1	533.06	7.1342***
Within Groups	8,520.02	114	74.74	
Total	9,053.08	115		

* $p < .05$
** $p < .02$
*** $p < .01$

Table 14

Analysis of Variance Summary Table:
Draw-A-Man Test - Standard Score
(N=116)

Source	SS	df	MS	F
Between Groups	2,454.43	1	2,454.43	11.0958***
Within Groups	25,217.16	114	221.20	
Total	27,671.59	115		

* $p < .05$
** $p < .02$
*** $p < .01$

Table 15

Analysis of Variance Summary Table:
Social Class Rating
(N=76)

Source	SS	df	MS	F
Between Groups	80.36	1	80.36	1.7849
Within Groups	1,800.79	40	45.02	
Total	1,881.15			

Table 16

Analysis of Variance Summary Table:
Social Class Rating
(N=116)

Source	SS	df	MS	F
Between Groups	124.10	1	124.10	2.5885
Within Groups	3,116.40	65	47.94	
Total	3,240.50	66		

(d) Metropolitan Achievement Test, Piaget Tests, and M.F.F. Tables 17 to 20 present the results of the multivariate analyses of variance for the total scores and subtest scores for equal and unequal Ns. First, it can be seen that when the total scores on the M.A.T. and the Piaget subtests were used with the two M.F.F. scores in a multivariate analysis of variance there were no significant differences between the two groups for each of the variables considered separately, or for the variables taken together.

When the 11 subtest scores were examined separately, Tables 19 and 20 show that the two groups differed significantly on only one subtest, seriation. This difference was apparent for both equal and unequal Ns ($F = 5.2187$, $p < .05$ $n = 76$; $F = 4.6996$, $p < .05$ $n = 116$). Overall there was no significant difference between the two groups.

A multivariate analysis of covariance was performed for the total scores and then for the subtest scores using the M.R.T. and Draw-A-Man Test scores as covariates. When the effects of readiness and intelligence were taken into account, the significant differences between the two groups on seriation disappeared for both equal and unequal Ns. In addition, no other significant differences were found between the two groups on any of the variables taken separately, or on the overall multivariate Fs. These results are presented in Tables 21 to 28.

Table 17

Multivariate Analysis of Variance Summary Table: Composite Score on Metropolitan Achievement Test, Total Score Piaget Subtests, and M.F.F. Error and Latency Scores (N=76)

Variable[a]	Mean Square	Univariate F	p less than
1	45.80	0.4461	0.5063
2	134.22	1.9766	0.1640
3	0.01	0.0012	0.9727
4	231.00	0.3232	0.5715

df for hypothesis = 1
df for error = 74

F ratio for multivariate test of equality of mean vectors = 0.6928, df = 4 and 71, p less than 0.5994

a Variable 1 - M.A.T. Composite Score
2 - Piaget Total Score
3 - M.F.F. Error
4 - M.F.F. Latency

Table 18

Multivariate Analysis of Variance Summary Table: Composite Score on Metropolitan Achievement Test, Total Score Piaget Subtests, and M.F.F. Error and Latency Scores
(N=116)

Variable	Mean Square	Univariate F	p less than
1	0.15	0.0015	0.9694
2	58.58	0.8419	0.3608
3	4.72	0.3716	0.5434
4	883.85	1.3601	0.2460

df for hypothesis = 1
df for error =114

F ratio for multivariate test of equality of mean vectors = 0.9472, df = 4 and 111, p less than 0.4397

a Variable 1 - M.A.T. Composite Score
2 - Piaget Total Score
3 - M.F.F. Error
4 - M.F.F. Latency

Table 19

Multivariate Analysis of Variance Summary Table: Metropolitan Achievement Subtest Scores, Piaget Subtest Scores and M.F.F. Error and Latency Scores (N=76)

Variable[a]	Mean Square	Univariate F	p less than
1	5.80	0.2159	0.6436
2	20.01	0.6158	0.4352
3	2.58	0.7778	0.3807
4	0.64	0.8780	0.3519
5	13.47	5.2187	0.0253*
6	11.07	0.8999	0.3460
7	1.59	0.4525	0.5033
8	2.58	0.7900	0.3770
9	0.47	0.2318	0.6317
10	0.01	0.0012	0.9727
11	231.00	0.3232	0.5715

df for hypothesis = 1
df for error = 74

F ratio for multivariate test of equality of mean vectors = 0.6354, df = 11 and 64, p less than 0.7761

* $p < .05$
** $p < .02$
*** $p < .01$

a Variables 1 - M.A.T. Computation
2 - M.A.T. Concepts
3 - Conservation of Substance
4 - All and Some
5 - Seriation
6 - Numeration
7 - Conservation of Weight
8 - Squares and Circles
9 - Multiple Seriation
10 - M.F.F. Error
11 - M.F.F. Latency

Table 20

Multivariate Analysis of Variance Summary Table: Metropolitan Achievement Subtest Scores, Piaget Subtest Scores and M.F.F. Error and Latency Scores
(N=116)

Variable[a]	Mean Square	Univariate F	p less than
1	3.86	0.1534	0.6961
2	2.70	0.0782	0.7803
3	1.65	0.4980	0.4819
4	0.05	0.0497	0.8241
5	12.98	4.6996	0.0323*
6	0.07	0.0046	0.9460
7	1.46	0.4361	0.5104
8	0.73	0.2128	0.6456
9	2.32	1.0467	0.3084
10	4.72	0.3716	0.5434
11	883.85	1.3601	0.2460

df for hypothesis = 1
df for error =114

F ratio for multivariate test of equality of mean vectors = 0.8751, df = 11 and 104, p less than 0.5669

* $p < .05$
** $p < .02$
*** $p < .01$

a Variable 1 - M.A.T. Computation
2 - M.A.T. Concepts
3 - Conservation of Substance
4 - All and Some
5 - Seriation
6 - Numeration
7 - Conservation of Weight
8 - Squares and Circles
9 - Multiple Seriation
10 - M.F.F. Error
11 - M.F.F. Latency

Table 21

Multivariate Analysis of Covariance Summary Table: Metropolitan Achievement Test Composite Score, Piaget Total Score and M.F.F. Error and Latency Scores with Draw-A-Man Test and Metropolitan Readiness Test Total Scores (N=76)

Variable[a]	Mean Square	Univariate F	p less than
1	32.44	0.5957	0.4428
2	4.79	0.1132	0.7375
3	10.30	1.0639	0.3058
4	1,277.32	1.8799	0.1749

df for hypothesis = 1
df for error = 72

F ratio for multivariate test of equality of mean vectors = 0.6099, df = 4 and 69, p less than 0.6569

a Variable 1 - M.A.T. Composite Score
2 - Piaget Total Score
3 - M.F.F. Error
4 - M.F.F. Latency

Table 22

Multivariate Analysis of Covariance Summary Table: Metropolitan Achievement Test Composite Score, Piaget Total Score and M.F.F. Error and Latency Scores with Draw-A-Man Test and Metropolitan Readiness Test Number Scores
(N=76)

Variable[a]	Mean Square	Univariate F	p less than
1	30.68	0.4596	0.5001
2	3.62	0.0759	0.7838
3	11.33	1.1927	0.2784
4	1,408.64	2.1135	0.1504

df for hypothesis = 1
df for error = 72

F ratio for multivariate test of equality of mean vectors = 0.6765, df = 4 and 69, p less than 0.6106

a Variable 1 - M.A.T. Composite Score
2 - Piaget Total Score
3 - M.F.F. Error
4 - M.F.F. Latency

Table 23

Multivariate Analysis of Covariance Summary Table: Metropolitan Achievement Test Composite Score, Piaget Total Score and M.F.F. Error and Latency Scores with Draw-A-Man Test and Metropolitan Readiness Test Total Scores (N=116)

Variable[a]	Mean Square	Univariate F	p less than
1	69.04	1.2380	0.2683
2	6.29	0.1657	0.6848
3	28.25	2.6461	0.1067
4	1,055.94	1.6532	0.2012

df for hypothesis = 1
df for error = 112

F ratio for multivariate test of equality of mean vectors = 1.0023, df = 4 and 109, p less than 0.4098

a Variable 1 - M.A.T. Composite Score
2 - Piaget Total Score
3 - M.F.F. Error
4 - M.F.F. Latency

Table 24

Multivariate Analysis of Covariance Summary Table: Metropolitan Achievement Test Composite Score, Piaget Total Score and M.F.F. Error and Latency Scores with Draw-A-Man Test and Metropolitan Readiness Test Number Scores
(N=116)

Variable[a]	Mean Square	Univariate F	p less than
1	111.77	1.8762	0.1735
2	15.56	0.3638	0.5477
3	32.58	3.1007	0.0810
4	1,239.25	2.0023	0.1599

df for hypothesis = 1
df for error = 112

F ratio for multivariate test of equality of mean vectors = 1.3097, df = 4 and 109, p less than 0.2710

a Variable 1 - M.A.T. Composite Score
2 - Piaget Total Score
3 - M.F.F. Error
4 - M.F.F. Latency

Table 25

Multivariate Analysis of Covariance Summary Table: Metropolitan Achievement Test Subtest Scores, Piaget Subtest Scores and M.F.F. Error and Latency Scores with Draw-A-Man and Metropolitan Readiness Test Total Scores
(N=76)

Variable[a]	Mean Square	Univariate F	p less than
1	7.66	0.4449	0.5070
2	7.67	0.4234	0.5174
3	0.32	0.1175	0.7328
4	0.18	0.2484	0.6198
5	2.83	1.2665	0.2642
6	0.87	0.0940	0.7601
7	1.31	0.4498	0.5047
8	0.34	0.1220	0.7279
9	1.15	0.5597	0.4569
10	10.30	1.0639	0.3058
11	1,277.32	1.8779	0.1749

df for hypothesis = 1
df for error = 72

F ratio for multivariate test of equality of mean vectors = 0.4862, df = 11 and 62, p less than 0.9050

a Variable 1 - M.A.T. Computation
2 - M.A.T. Concepts
3 - Conservation of Substance
4 - All and Some
5 - Seriation
6 - Numeration
7 - Conservation of Weight
8 - Squares and Circles
9 - Multiple Seriation
10 - M.F.F. Error
11 - M.F.F. Latency

Table 26

Multivariate Analysis of Covariance Summary Table: Metropolitan Achievement Test Subtest Scores, Piaget Subtest Scores and M.F.F. Error and Latency Scores with Draw-A-Man and Metropolitan Readiness Test Number Scores (N=76)

Variable[a]	Mean Square	Univariate F	p less than
1	8.33	0.4524	0.5034
2	6.27	0.2714	0.6041
3	0.21	0.0735	0.7872
4	0.17	0.2379	0.6272
5	2.92	1.2785	0.2619
6	0.77	0.0778	0.7812
7	1.11	0.3682	0.5460
8	0.22	0.0780	0.7809
9	1.33	0.6622	0.4185
10	11.33	1.1927	0.2784
11	1,408.64	2.1135	0.1504

df for hypothesis = 1
df for error = 72

F ratio for multivariate test of equality of mean vectors = 0.5438, df = 11 and 62, p less than 0.8656

a Variable 1 - M.A.T. Computation
2 - M.A.T. Concepts
3 - Conservation of Substance
4 - All and Some
5 - Seriation
6 - Numeration
7 - Conservation of Weight
8 - Squares and Circles
9 - Multiple Seriation
10 - M.F.F. Error
11 - M.F.F. Latency

Table 27

Multivariate Analysis of Covariance Summary Table: Metropolitan Achievement Test Subtest Scores, Piaget Subtest Scores and M.F.F. Error and Latency Scores with Draw-A-Man and Metropolitan Readiness Test Total Scores (N=116)

Variable[a]	Mean Square	Univariate F	p less than
1	25.19	1.4157	0.2366
2	10.30	0.5294	0.4684
3	0.30	0.1105	0.7403
4	0.19	0.2203	0.6398
5	11.02	4.7788	0.0309
6	2.42	0.2158	0.6432
7	0.03	0.0087	0.9261
8	0.03	0.0101	0.9201
9	2.55	1.2118	0.2734
10	28.25	2.6461	0.1067
11	1,055.94	1.6532	0.2012

df for hypothesis = 1
df for error = 112

F ratio for multivariate test of equality of mean vectors = 1.1071, df = 11 and 102, p less than 0.3636

a Variable 1 - M.A.T. Computation
2 - M.A.T. Concepts
3 - Conservation of Substance
4 - All and Some
5 - Seriation
6 - Numeration
7 - Conservation of Weight
8 - Squares and Circles
9 - Multiple Seriation
10 - M.F.F. Error
11 - M.F.F. Latency

Table 28

Multivariate Analysis of Covariance Summary Table: Metropolitan Achievement Test Subtest Scores, Piaget Subtest Scores and M.F.F. Error and Latency Scores with Draw-A-Man and Metropolitan Readiness Test Number Scores (N=116)

Variable[a]	Mean Square	Univariate F	p less than
1	36.92	2.1306	0.1472
2	19.50	0.8806	0.3501
3	0.10	0.0330	0.8563
4	0.11	0.1305	0.7187
5	9.45	3.9260	0.0500
6	4.93	0.4309	0.5130
7	0.00	0.0002	0.9900
8	0.00	0.0005	0.9828
9	3.08	1.4657	0.2286
10	32.58	3.1007	0.0810
11	1,239.25	2.0023	0.1599

df for hypothesis = 1
df for error =112

F ratio for multivariate test of equality of mean vectors = 1.1831, df = 11 and 102, p less than 0.3082

a Variable 1 - M.A.T. Computation
2 - M.A.T. Concepts
3 - Conservation of Substance
4 - All and Some
5 - Seriation
6 - Numeration
7 - Conservation of Weight
8 - Squares and Circles
9 - Multiple Seriation
10 - M.F.F. Error
11 - M.F.F. Latency

The adjusted and unadjusted means for the Metropolitan Achievement Test, the Piaget subtests and the M.F.F. are shown in Tables 29 to 34. Except for the unadjusted equal N scores, there seems to be a trend for the means of the 75-minute group to be higher on the Metropolitan Achievement Test. On the Piaget subtests, there is a trend for the unadjusted mean scores of the immersion children to be higher, but this trend is not evident for the adjusted mean scores.

In addition to analyzing the Piagetian subtests using numerical scores, both groups of children were divided into those who exhibited concrete operational thinking (Stage III) versus those who did not (Stages I and II), and a X^2 computed for each of the subtests. In one subtest, multiple seriation, there were no Stage III children, therefore comparisons were made between Stage I and Stage II children. Table 35 shows that the two groups differed on only one subtest, classification of squares and circles, the number of Stage III immersion children being significantly greater than the number of Stage III 75-minute children ($X^2 = 5.682$, $p < .02$). On the seriation subtest, the difference between the two groups almost reached significance level in favour of the immersion children ($X^2 = 3.45$, $p < .10$).

Table 29

Means and Adjusted Means for 75-minute and Immersion Groups on Metropolitan Achievement Test (N=76)

Test	Immersion		75-minute	
	Mean	Adj. Means	Mean	Adj. Means
M.A.T. - Composite Score	43.03	41.57[a] 41.55[b]	41.47	42.93 42.95
M.A.T. - Computation Score	19.71	19.08 19.10	19.16	19.79 19.77
M.A.T. - Concepts Score	23.34	22.52 22.49	22.32	23.14 23.17

a Adjusted for effect of Metropolitan Readiness Test - Number and Draw-A-Man Test.

b Adjusted for effect of Metropolitan Readiness Test - Total and Draw-A-Man Test.

Table 30

Means and Adjusted Means for 75-minute and Immersion Groups on Metropolitan Achievement Test (N=116)

Test	Immersion		75-minute	
	Mean	Adj. Means	Mean	Adj. Means
M.A.T. - Composite Score	41.40	40.72[a] 40.87[b]	41.47	42.87 42.56
M.A.T. - Computation Score	18.77	18.49 18.56	19.16	19.73 19.58
M.A.T. - Concepts Score	22.64	22.34 22.32	22.32	23.14 22.97

a Adjusted for effect of Metropolitan Readiness Test - Number and Draw-A-Man Test.

b Adjusted for effect of Metropolitan Readiness Test - Total and Draw-A-Man Test.

Table 31

Means and Adjusted Means for 75-minute and Immersion Groups on Concrete Operational Thinking Test (N=76)

Test	Immersion		75-minute	
	Mean	Adj. Means	Mean	Adj. Means
Substance	2.16	1.92[a] 1.90[b]	1.79	2.03 2.04
All & Some	3.08	3.04 3.04	2.89	2.94 2.94
Seriation	4.84	4.63 4.63	3.99	4.21 4.22
Numeration	9.39	8.91 8.90	8.63	9.12 9.13
Weight	1.76	1.49 1.48	1.47	1.75 1.76
Squares & Circles	3.58	3.34 3.32	3.21	3.45 3.47
Multiple Seriation	3.29	3.23 3.24	3.45	3.51 3.50
Total Score	28.10	26.54 26.51	25.45	27.01 27.04

a Adjusted for effect of Metropolitan Readiness Test - Number and Draw-A-Man Test.

b Adjusted for effect of Metropolitan Readiness Test - Total and Draw-A-Man Test.

Table 32

Means and Adjusted Means for 75-minute and Immersion Groups on Concrete Operational Thinking Test (N=116)

Test	Immersion Mean	Immersion Adj. Means	75-minute Mean	75-minute Adj. Means
Substance	2.04	1.98[a] 1.99[b]	1.79	1.92 1.89
All & Some	2.94	2.95 2.95	2.89	2.88 2.86
Seriation	4.71	4.68 4.70	3.99	4.06 4.02
Numeration	8.58	8.45 8.49	8.63	8.90 8.81
Weight	1.71	1.63 1.65	1.47	1.64 1.61
Squares & Circles	3.38	3.32 3.34	3.21	3.33 3.30
Multiple Seriation	3.15	3.13 3.14	3.45	3.49 3.46
Total Score	26.96	26.20 26.30	25.45	27.01 26.81

a Adjusted for effect of Metropolitan Readiness Test - Number and Draw-A-Man Test.

b Adjusted for effect of Metropolitan Readiness Test - Total and Draw-A-Man Test.

Table 33

Means and Adjusted Means for 75-minute and Immersion Groups on Matching Familiar Figures Test
(N=76)

Test	Immersion		75-minute	
	Mean	Adj. Means	Mean	Adj. Means
M.F.F. - Error Score	7.26	7.66[a] 7.64[b]	7.24	6.84 6.86
M.F.F. - Latency Score	56.18	53.34 53.56	59.67	62.52 62.29

a Adjusted for effect of Metropolitan Readiness Test - Number and Draw-A-Man Test.

b Adjusted for effect of Metropolitan Readiness Test - Total and Draw-A-Man Test.

Table 34

Means and Adjusted Means for 75-minute and Immersion Groups on Matching Familiar Figures Test (N=116)

Test	Immersion		75-minute	
	Mean	Adj. Means	Mean	Adj. Means
M.F.F. - Error Score	7.67	7.91[a] 7.88[b]	7.24	6.74 6.80
M.F.F. - Latency Score	53.79	53.36 53.55	59.67	60.55 60.17

a Adjusted for effect of Metropolitan Readiness Test - Number and Draw-A-Man Test.

b Adjusted for effect of Metropolitan Readiness Test - Total and Draw-A-Man Test.

Table 35

Chi Square for Number of Children at Stages I and II versus Stage III by Program

Test	Stages I & II		Stage III		Level of Signif.
	Imm.	75-min.	Imm.	75-min.	
Substance	25	26	13	12	ns
All & Some	22	28	16	10	ns
Seriation	12	20	26	18	ns
Numeration	33	33	5	5	ns
Weight	23	26	15	12	ns
Squares & Circles	24	33	14	5	$p<.02$
Multiple Seriation	12	8	26	30	ns

The studies considered in the review of the literature indicated that children classified as reflective on the M.F.F. might do better in arithmetic and academic achievement than impulsive children. This proposition found partial support in the results of the analyses of variance with significant differences found between the four groups (fast-accurate, slow-accurate, fast-inaccurate, slow-inaccurate) on the concepts subtest of the M.A.T. and on numeration. Further examination of the mean scores using Scheffé's test of multiple comparisons (Keith, 1972) indicated that the slow-accurate (reflective) group mean was significantly higher than that of the slow-inaccurate group on the concepts subtest and significantly higher than both the slow-inaccurate and fast-inaccurate (impulsive) groups on numeration. Results of the analyses of variance are presented in Tables 36 to 46.

Table 36

Analysis of Variance Summary Table for Metropolitan Achievement Test-Composite under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	472.89	3	157.63	1.6183
Within Groups	10,909.41	112	97.41	
Total	11,382.30	115		

Table 37

Analysis of Variance Summary Table for Metropolitan Achievement Test-Computation under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	25.05	3	8.35	0.3284
Within Groups	2,847.71	112	25.43	
Total	2,872.76	115		

Table 38

Analysis of Variance Summary Table for Metropolitan Achievement Test-Concepts under Four Levels of Matching Familiar Figures Test Scores
(N=116)

Source	SS	df	MS	F
Between Groups	288.55	3	96.18	2.9478*
Within Groups	3.654.32	112	32.63	
Total	3,942.87	115		

* $p < .05$
** $p < .02$
*** $p < .01$

Table 39

Analysis of Variance Summary Table for Concrete Operational Thinking-Total under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	394.99	3	131.67	1.9414
Within Groups	7,595.86	112	67.82	
Total	7,990.85	115		

Table 40

Analysis of Variance Summary Table for Concrete Operational Thinking-Substance under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	8.38	3	2.79	0.8431
Within Groups	370.99	112	3.31	
Total	379.37	115		

Table 41

Analysis of Variance Summary Table for Concrete Operational Thinking-All & Some under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	1.28	1	0.43	0.4570
Within Groups	104.45	112	0.93	
Total	105.73	115		

Table 42

Analysis of Variance Summary Table for Concrete Operational Thinking-Seriation under Four levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	10.03	3	3.34	1.1781
Within Groups	317.88	112	2.84	
Total	327.91	115		

Table 43

Analysis of Variance Summary Table for Concrete Operational Thinking-Numeration under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	158.44	3	52.81	4.0266***
Within Groups	1,469.04	112	13.12	
Total	1,627.48	115		

*$p < .05$
**$p < .02$
***$p < .01$

Table 44

Analysis of Variance Summary Table for Concrete Operational Thinking-Weight under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	18.04	3	6.01	1.8431
Within Groups	365.42	112	3.26	
Total	383.46	115		

Table 45

Analysis of Variance Summary Table for Concrete Operational Thinking-Squares & Circles under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	9.53	3	3.18	0.9317
Within Groups	382.04	112	3.41	
Total	391.57	115		

Table 46

Analysis of Variance Summary Table for Concrete Operational Thinking-Multiple Seriation under Four Levels of Matching Familiar Figures Test Scores (N=116)

Source	SS	df	MS	F
Between Groups	1.69	3	0.06	0.2494
Within Groups	253.12	112	2.26	
Total	254.81	115		

(e) Multiple Regression. In order to test hypothesis 4 which stated in the null form that if program was used as a moderator variable, the multiple R for the 75-minute group would not differ significantly from the multiple R for the immersion group using arithmetic-achievement as the criterion and the other cognitive measures as predictors, separate stepwise multiple Rs were computed. Table 47 shows the best predictors which entered into the regression equation for the 75-minute and immersion groups separately and combined. Numeration was the most powerful predictor in all four cases, while multiple seriation entered into the prediction equation for both the equal N immersion and unequal N immersion groups as well as the total group.

Zedeck (1971) has defined one type of moderator variable as a qualitative or quantitative variable that influences a predictor-criterion relationship. Thus he states that a third variable (the moderator):

> is used as a basis for splitting the total group into subgroups. That is, r_{xy} for Group A is significantly different from r_{xy} for Group $\bar{A}$ (not A), where A is a qualitative or quantitative descriptive variable (p. 296).

Zedeck has described three methods for demonstrating "differential validities among qualitative subgroups or predictability subgroups (p. 303)." These methods are:

Table 47

Best Predictors for Predicted Metropolitan Achievement Test-Composite Score Using Program as a Moderator Variable

Group	Best Predictor Name r	2nd Best Predictor Name R	3rd Best Predictor Name R
75-minute	Numeration r = .66	Weight R = .73	Squares & Circles R = .74
Immersion (N=38)	Numeration r = .56	Multiple Seriation R = .74	M.F.F. (Error) R = .76
Immersion (N=78)	Numeration r = .53	Multiple Seriation R = .62	Substance R = .67
Total (N=116)	Numeration r = .58	Substance R = .64	Multiple Seriation R = .67

> (a) Significantly different coefficients between subgroups; that is r_{xy} for Subgroup 1 is significantly different from r_{xy} for Subgroup 2; (b) patterns of validity coefficients reveal some coefficients to be significantly different from zero; that is, r_{xy} for Subgroups 1 and 2 is significantly different from zero, but r_{xy} for Subgroup 3 is not significantly different from zero; and (c) the validity coefficient of a subgroup is different from the validity coefficient of the total group; that is r_{xy} for Subgroup 1 is greater than r_{xy} for the total group, whereas r_{xy} for Subgroup 2 is equal to or less than r_{xy} for the total group (p. 303).

In order to test method (a), separate multiple Rs were computed for the two immersion groups ($N = 38$ and $N = 78$) and for the 75-minute group. These multiple Rs are shown in Table 48. Since the F for the multiple Rs did not reach significance when the third predictor entered the equation, the multiple R cutting off after two predictors was used. Table 48 shows program was not acting as a moderator in the manner described by Zedeck (1971) since there was no significant difference between the two multiple Rs. In addition, when method (b) was applied, it was found that all the multiple Rs differed significantly from zero.

In order to test hypothesis 5 of no significant difference between the multiple Rs for the 75-minute and immersion groups in comparison with the total group, the standard error of the difference between a multiple correlation coefficient of a subgroup and the multiple correlation coefficient of a total group containing the

Table 48

Multiple Rs, z Scores, Standard Error of the Difference and Levels of Statistical Significance for Immersion and 75-minute Comparisons

Group	N	R	z	t	df	p
75-minute	38	.73***	.9287			
Immersion (N=38)	38	.74***	.9505	.0912	74	ns
Immersion (N=78)	78	.62***	.7250	.9951	114	ns

* $p < .05$
** $p < .02$
*** $p < .01$

subgroup was computed for the following: total group versus 75-minute group; total group versus immersion (N = 38); total group versus immersion (N = 78). Table 49 gives the results of these calculations. It can be seen that both the 75-minute group and the reduced immersion group differ significantly from the total group, again indicating that program is not acting as a moderator variable.

In order to test null hypothesis 6 of no significant difference between the multiple Rs for two randomly constituted subgroups, and between each subgroup and the total group, the multiple Rs for each subgroup were first converted into z scores and the standard error of the difference between two z scores computed. Table 50 shows the predictors entering into the equations for each random subgroup. Table 51 shows that there was no significant difference between the two random subgroup multiple Rs. To test the second part of the hypothesis, the standard error of the difference between a multiple correlation coefficient of a subgroup and the multiple correlation coefficient of a total group containing the subgroup was computed for each subgroup comparison. Table 52 shows that there was no significant difference between either subgroup and the total group.

Table 49

Multiple Rs, z Scores, Diff/Rs and Levels of Statistical Significance for 75-minute and Immersion Groups Compared with Total Group

Group	N	R	z	Diff/Rs	p
Total Group	116	.64	.7582		
75-minute	38	.73	.9287	2.13	.05
Immersion (N=38)	38	.74	.9505	2.41	.02
Immersion (N=78)	78	.62	.7250	0.86	ns

Table 50

Best Predictors for Predicted Metropolitan Achievement Test-Composite Score Using Random Subgroups (n=58)

Group	Best Predictor Name r	2nd Best Predictor Name R	3rd Best Predictor Name R
Subgroup A	Numeration .58	M.F.F. (error) .64	Multiple Seriation .68
Subgroup B	Substance .59	Numeration .68	Multiple Seriation .71

Table 51

Multiple Rs, z Scores, Standard Error of the Difference and Levels of Statistical Significance for Subgroup Comparisons

Group	N	R	z	t	df	p
Subgroup A	58	.68***	.8291			
Subgroup B	58	.71***	.8872	.304	114	ns

* $p < .05$
** $p < .02$
*** $p < .01$

Table 52

Multiple Rs, z Scores, Diff/Rs and Levels of Statistical Significance for Random Subgroup Comparisons with Total Group

Group	N	R	z	Diff/Rs	p
Total Group	116	.67	.8107		
Subgroup A	58	.68	.8291	0.36	ns
Subgroup B	58	.71	.8872	1.48	ns

In order to test the final hypothesis, the constant and B weights from each of the random subgroups were applied to the scores of the other subgroup in a cross-validation design, and new multiple Rs computed between the new predicted arithmetic scores and the actual arithmetic scores. Table 53 shows that when the multiple Rs were converted to z scores, there was no significant difference between the z scores. Table 54 shows that the cross-validation multiple Rs, both when compared separately and averaged, did not differ from the multiple R for the total group.

(f) Intercorrelation Matrices. Though these were not used in the testing of any hypotheses, intercorrelations were computed between each of the dependent variables in order to examine any differences in the pattern of intercorrelations between the two groups. Tables 55, 56, and 57 present these intercorrelations. It seems that generally the intercorrelations for the 75-minute group are slightly higher, which is not unexpected since the 75-minute children are probably a more heterogeneous group than the immersion children. One interesting pattern that appears to emerge is for the intercorrelations between the concrete operational tests involving classification, i.e., All and Some, Squares and Circles, and Multiple Seriation, to be more highly correlated with arithmetic-achievement for the immersion children than for the 75-minute children. Equally,

Table 53

Multiple Rs, z Scores, Standard Error of the Difference and Levels of Statistical Significance for Sub-group Comparisons

Group	N	R	z	t	df	p
Cross Validation B to A	58	.62***	.7250			
Cross Validation A to B	58	.63***	.7414	0.086	114	ns

* $p < .05$
** $p < .02$
*** $p < .01$

Table 54

Multiple Rs, z Scores, Diff/rs and Levels of Statistical Significance for Cross Validation Subgroup Comparisons with Total Group

Group	N	R	z	Diff/Rs	p
Cross Validation B to A	58	.62	.7250	1.65	ns
Cross Validation A to B	58	.63	.7414	1.34	ns
Average of Cross Validation Groups			.7332	1.50	ns
Total Group	116	.67	.8107		

Table 55

Intercorrelation Matrix of Dependent Variables
75-minute (N=38)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. M.A.T. Composite	-												
2. M.A.T. Comput.	.93	-											
3. M.A.T. Concepts	.94	.75	-										
4. Concrete Op. Thinking-Total	.69	.62	.67	-									
5. Substance	.51	.42	.54	.78	-								
6. All & Some	.01	.07	-.05	.30	.25	-							
7. Seriation	.53	.50	.48	.74	.48	.18	-						
8. Numeration	.68	.60	.66	.84	.48	.08	.61	-					
9. Weight	.59	.54	.57	.80	.83	.25	.45	.53	-				
10. Squares & Circles	.24	.17	.27	.57	.40	.17	.21	.38	.49	-			
11. Multiple Seriation	.18	.21	.13	.23	.05	-.00	.20	.14	-.05	-.13	-		
12. M.F.F. (Error)	-.41	-.29	-.47	-.39	-.35	-.17	-.29	-.41	-.33	-.22	.21	-	
13. M.F.F. (Latency)	.20	.22	.15	.20	.14	-.14	.15	.37	.03	-.06	.03	-.54	-

r - .325 $p < .05$
r - .413 $p < .01$

Table 56

Intercorrelation Matrix of Dependent Variables
Immersion (N=78)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. M.A.T. Composite	-												
2. M.A.T. Comput.	.88	-											
3. M.A.T. Concepts	.92	.64	-										
4. Concrete Op. Thinking-Total	.71	.56	.72	-									
5. Substance	.42	.32	.43	.66	-								
6. All & Some	.22	.11	.27	.24	.20								
7. Seriation	.24	.11	.31	.44	.36	.14	-						
8. Numeration	.53	.44	.52	.68	.25	.19	.39	-					
9. Weight	.28	.16	.32	.55	.66	.23	.20	.24	-				
10. Squares & Circles	.37	.35	.33	.52	.38	.07	.24	.26	.17	-			
11. Multiple Seriation	.40	.40	.35	.36	.19	.24	.17	.20	-.06	.37	-		
12. M.F.F. (Error)	-.32	-.19	-.36	-.44	-.20	.02	-.24	-.33	-.20	-.14	-.04	-	
13. M.F.F. (Latency)	.16	.04	.22	.19	.09	.03	.18	.37	.24	.00	.12	-.30	-

$r = .232 \quad p < .05$
$r = .302 \quad p < .01$

Table 57

Intercorrelation Matrix of Dependent Variables
Immersion (N=38)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. M.A.T. Composite	-												
2. M.A.T. Comput.	.92	-											
3. M.A.T. Concepts	.93	.70	-										
4. Concrete Op. Thinking-Total	.68	.52	.73	-									
5. Substance	.45	.32	.51	.70	-								
6. All & Some	.22	.22	.20	.22	.03	-							
7. Seriation	.21	.17	.21	.43	.08	.15	-						
8. Numeration	.57	.41	.63	.80	.36	.06	.45	-					
9. Weight	.17	.01	.31	.62	.69	.17	.15	.28	-				
10. Squares & Circles	.40	.36	.38	.54	.21	-.08	-.02	.33	.13	-			
11. Multiple Seriation	.54	.55	.46	.41	.20	.13	.01	.10	.01	.34	-		
12. M.F.F. (Error)	-.29	-.30	-.22	-.20	-.02	-.22	-.02	-.33	-.03	-.27	.05	-	
13. M.F.F. (Latency)	.11	-.03	.21	.22	.04	-.22	.24	.31	.25	.00	-.05	-.35	-

$r = .325 \quad p < .05$
$r = .418 \quad p < .01$

conservation of weight seems to be more highly correlated with arithmetic for the 75-minute children. Curiously, this difference does not show up for conservation of substance.

This chapter has presented the results of hypotheses 1 to 7. The next chapter discusses these results and presents a summary of the results and conclusions to be drawn from them.

CHAPTER IV

DISCUSSION

This chapter discusses the results of the statistical analyses presented in the previous chapter in the light of the predictions derived from the review of the literature.

Because all other hypotheses stemmed from it, the results of hypothesis 1 will be discussed first. This hypothesis stated in the null form that there would be no difference between the scores of the immersion children and the scores of the 75-minute children on the Metropolitan Achievement Test. This null hypothesis was accepted for all comparisons made both with or without the effects of intelligence and readiness removed. However, Table 29, p. 86, shows that for equal Ns there was a trend for the unadjusted means of the immersion group to be higher, but for unequal Ns and adjusted means, the trend was reversed.

First, the difference in trends between equal and unequal Ns will be discussed. An analysis of variance using male and female as the independent variables and the three scores on the Metropolitan Achievement Test as dependent variables indicated that there were no significant differences between males and females. This is in keeping with the findings of Edwards and Casserly (1972). However, within the 75-minute group the ratio of boys to girls was higher, and the analysis of variance by sex indicated that there was a trend for the mean scores in arithmetic for the boys to be slightly higher than that of the girls. This is the probable explanation of the reversal in trends for unequal Ns. Thus the results for the equal Ns where the male/female ratio is the same for both groups are considered to reflect more accurately differences between the two programs, at least in arithmetic.

There are a number of possible reasons why no differences were found between the immersion and 75-minute groups, in spite of consistent differences found by Edwards and Casserly (1971, 1972) in the two years prior to this study. The first reason relates to the homogeneity of the immersion and 75-minute children in social background. All three schools were specifically chosen to minimize any difference in social-class background between the two groups, and this was confirmed by finding no significant differences in social-class ratings. This information was not available to Edwards and Casserly (1971, 1972) and it is possible that their immersion and 75-minute samples may have differed on this variable. Of course, social-class ratings are very gross measures, and it is likely that more subtle measures of home environment might more clearly differentiate between parents who enroll their children in immersion programs versus those who do not.

Both this study and Edwards and Casserly (1971, 1972) found IQ differences between the two groups, the immersion children having significantly higher mean IQ scores. For instance, Edwards and Casserly (1971) found the Lorge Thorndike mean IQs for the pilot immersion and 75-minute groups to be 105 and 100, respectively; for the follow-up Grade 1 immersion and 75-minute children, the mean IQs were 112 and 105, respectively. This compares with a Draw-A-Man standard score IQ of 109 for the immersion children and 99 for the 75-minute children. However, Edwards and Casserly (1971, 1972) found differences both with and without the effect of intelligence being removed through covariance, and it is possible that intelligence differences, per se, did not contribute very much to the differences in arithmetic-achievement that were found. For instance, the Lorge Thorndike IQ in Edwards and Casserly's study only correlated .22 with arithmetic-achievement on the M.A.T. In the present study the correlations ranged from .22 to .33 for Draw-A-Man standard IQs with the Metropolitan Achievement Test. These correlations are generally lower than those found for the readiness scores.

It would seem more pertinent to examine the relationship between the readiness scores on the Metropolitan Readiness Test and arithmetic-achievement. Tables 4 and 5,

p. 56 and 57, show that these range from .49 to .70, generally being higher for the M.R.T. total score than for M.R.T. number score. It is obvious then, that any differences in readiness would contribute substantially to differences in arithmetic-achievement at the end of the first grade. It may be that the differences in arithmetic-achievement found by Edwards and Casserly were due to initial readiness differences that existed at the beginning of Grade 1.

Thus, though the findings of this study do not provide evidence that immersion programs depress academic achievement, they show no acceleration on the part of the immersion children, at least at the Grade 1 level. This, of course, is similar to the findings of Barik and Swain (1972) and Lambert _et al_. (1969), Lambert only finding differences in arithmetic-achievement at the later grades.

Because the two groups did not show any real differences in arithmetic-achievement, it is not unexpected to find none in concrete operational thinking. The slightly higher mean scores shown by the immersion children disappear when account is taken of readiness and intelligence, and in this they follow the trend of the arithmetic scores. In fact, like the arithmetic mean scores, there seems to be a reversal in trend in that overall the mean scores for the 75-minute children are slightly higher when adjusted for readiness and intelligence, though the difference is very slight.

When children are divided on the basis of stages of concrete operational thinking, Stages I and II versus Stage III, there are more immersion children at Stage III on all of the subtests except multiple seriation. This difference reached significance for classification of squares and circles. These results, though, should be viewed with some caution since they would probably disappear if account was taken of intellectual and readiness differences. Rather, they can be viewed as reflecting the fact that the immersion children as a group are not representative of all children enrolled in the Separate School Board system. Apart from any initial preferences on the part of parents, children with learning disabilities or children who are slow learners are unlikely at this stage to be enrolled in an immersion program.

If one can forget for a moment initial readiness and intelligence differences that may exist between the two groups, several points can be made about the stage differences in concrete operational thinking that exist between the two groups. The first point involves the classification of squares, and multiple seriation subtests. It is rather curious to find that in squares and circles the immersion children are ahead while the reverse is true for multiple seriation since these two tests appear to involve similar abilities. Both involve the child's ability to make size and colour discriminations and appropriately group objects on these two dimensions.

On the seriation subtest, the difference between the number of Stage III immersion and 75-minute children did not quite reach significance. Nevertheless, it is interesting to look at the breakdown of children for each of the three stages on this test, which was the easiest of all the seven subtests. Table 35, p. 92, shows that out of a total of 76 children, 44 had reached Stage III in seriation. Table 58 shows the breakdown of children at each stage by program. It can be seen that while only one immersion child was still at Stage 1, eight 75-minute children were still at this stage, indicating that a far higher proportion of 75-minute children were unable to seriate ten slats by the end of Grade 1. Of the total of nine children in both groups who were at Stage I seriation, only two had reached Stage III in any other test indicating a rather general slowness in cognitive development.

Neither the error nor latency scores on the Matching Familiar Figures Test differentiated between the immersion and 75-minute children. For both the adjusted and unadjusted mean scores, the 75-minute children could be described as being more "reflective" since they made fewer errors and had

Table 58

Concrete Operational Thinking - Seriation Subtest, Numbers of Subjects by Stage and Program (N=76)

Program	Stage I	Stage II	Stage III
75-minute	8	12	18
Immersion	1	11	26
Total	9	23	44

a longer latency to first response than the immersion children. This is a somewhat odd finding in the light of the relationship between arithmetic, concrete operational thinking, and "reflection" to be discussed next.

Kagan _et al_. (1963, 1964) felt that individual differences in cognitive style might be relevant to subjects such as mathematics. Spitler (1971) went one step farther and hypothesized that while impulsiveness might be necessary for arithmetic computation, arithmetic problem-solving requires reflection. The relationship between arithmetic and reflection was confirmed by Cathcart and Liedtke (1969). Similar results were found in this study in that slow-accurate (reflective) children made significantly higher mean scores on the M.A.T. concepts subtest in comparison with slow-inaccurate children; in the numeration subtest they made significantly higher scores than the slow-inaccurate and fast-inaccurate groups.

Table 59 shows the mean scores of the four groups on the various arithmetic and concrete operational thinking tests. Not surprisingly, the children who make few errors on the M.F.F. also have higher mean scores on arithmetic and other cognitive measures, but this finding is most marked for the slow-accurate children. The exception to this is on the three subtests, all and some, squares and circles, and

Table 59

Mean Scores for Metropolitan Achievement Test Scores and Concrete Operational Thinking Test Scores for Four Levels of Matching Familiar Figures Test Scores (N=116)

Test	Fast Accurate	Slow Accurate	Fast Inaccurate	Slow Inaccurate
M.A.T. Composite	41.65	44.07	40.86	36.83
M.A.T. Comput.	18.77	19.45	18.92	17.75
M.A.T. Concepts	22.88	24.62	21.96	19.08
Concrete Ops. Total	26.65	29.41	25.16	24.25
Substance	1.96	2.38	1.70	2.00
All & Some	3.12	2.90	2.86	2.83
Seriation	4.46	4.97	4.28	4.17
Numeration	8.69	10.48	7.74	7.33
Weight	1.42	2.03	1.32	2.42
Squares & Circles	3.65	3.38	3.30	2.58
Multiple Seriation	3.35	3.31	3.23	2.92

multiple seriation, where the fast-accurate children have higher mean scores. These three tests all involve classificatory skills. The differences in the mean scores are of course very slight, but it is, nevertheless, interesting that there does seem to be some kind of pattern to the differences.

The fact that "reflective" children have higher mean scores in concrete operational thinking, even if only one, numeration, reached significance lends weight to the comment by Peters (1967) that the child who is free from distraction and displays an analytical attitude will be more likely to achieve conservation of number before an age-mate of the same general intelligence. However, one anomaly remains unexplained, though it only relates to trends rather than significant differences. It appears that prior to adjustment for intelligence and readiness, the immersion children have higher mean scores on concrete operational thinking, but the 75-minute children are generally more "reflective."

Two aspects of multiple regression need to be considered. First, the differences between the two programs and, secondly, the reliability of the predictors. Three methods were used to see if program was acting as a moderator variable in the manner described by Zedeck (1971). None of these methods indicated that the 75-minute group differed from the immersion group and thus program was operating as a moderator variable. The two validity coefficients for the groups did not differ from each other significantly when the standard error of the difference between two z scores was calculated. When the standard error of the difference between a subgroup and a total group of which the subgroup is a part was calculated for each subgroup in comparison with the total group, both the 75-minute and the small immersion subgroup differed significantly from the total group. However, the large immersion group did not differ significantly from the total group, but, since the N was larger, one would hesitate to use this as evidence that program was operating as a moderator variable. Finally, all multiple Rs were significantly different from zero.

In the classical validation model, only the relationship between predictor and criterion is examined, with factors that might intervene between the two behaviours being ignored (Zedeck, 1971). The use of moderators has involved both discrete and continuous variables. However, in this study the operation of a discrete variable has been considered. With such discrete variables a "heterogeneous population is divided into more homogeneous subgroups. A complex aggregate of individuals is simplified, and there is no interaction between the moderator and predictor-criterion score combination (Zedeck, 1971, p. 305)." It is apparent, however, from the results of this study that when Ss were divided on the basis of program, the two subgroups thus formed were not more homogeneous than the group as a whole. That is, the total group was not in fact a heterogeneous aggregate of individuals, but as the majority of results in this study have shown, remarkably homogeneous.

Though the variables that entered into the multiple regression equation for the two groups were not identical, they had a number of similarities. Both had numeration as the most powerful predictor. In addition, both had a variable that involved ability to classify by size and colour; multiple seriation for the immersion group and squares and circles for the 75-minute group. If one compares the large immersion group to the 75-minute group, they also both had a conservation subtest in the equation, conservation of weight and substance, respectively. These two tests are very highly correlated, at least in this study.

The second aspect to be considered is the reliability of the predictors considered for the group as a whole. In order to test this out, the total group was divided into two random subgroups and separate multiple regressions run for each subgroup. The B weights and constant from each subgroup were then applied to the scores of the other subgroup and new multiple Rs computed. Using the three tests applied to the previous multiple Rs (that is, did they differ significantly from zero, did they differ significantly from each other, and did they differ significantly from the total group), no significant differences were found either for the original random subgroups or the cross-validation subgroups. The two random subgroups were very close in terms of their multiple Rs, .68 versus .71 for the final multiple Rs. In addition, they had two predictors in common, numeration and multiple seriation. The "drop" in multiple Rs when cross-validation was carried out was very minimal, from .68 to .62 and from .71 to .63.

These predictor multiple Rs are all reasonably high, and along with the individual correlations between concrete operational thinking and arithmetic, provide an indication that a strong relationship does exist between the two. It is apparent that numeration is the most highly related to arithmetic-achievement; followed, it seems, by multiple seriation. Of course, the arithmetic-achievement score was the combined score for both concepts and computation. But if one looks at the intercorrelations for the two subgroups separately, it is apparent that concepts is related more highly to logical thinking than is computation. This provides endorsement for the idea that the early grades children can carry out mechanical arithmetic without truly understanding the nature of numbers.

Studies considered in the review of the literature showed that bilingualism was associated with both positive and negative effects on intelligence or cognitive functioning and academic achievement. This research, though it has not found the expected enhancement in arithmetic and related cognitive functioning,has indicated that even though the immersion children are being taught almost exclusively in French, they are able to perform as well as the 75-minute children.

This chapter has discussed the results of the study. The next section will summarize these results and consider conclusions that can be drawn from them.

SUMMARY AND CONCLUSIONS

This section summarizes the major findings of the study and puts forward a number of conclusions arising out of these findings.

1. No significant differences were found between the mean scores of the immersion children and the mean scores of the 75-minute children on either arithmetic-computation, arithmetic concepts, or composite arithmetic score of the Metropolitan Achievement Test.

2. On tests of concrete operational thinking, a significant difference was found between the mean score of the 75-minute children and the mean score of the immersion children on seriation. However, this difference disappeared when the effects of intelligence and readiness were removed. In addition, when the results of the tests were considered in terms of the number of children at Stages I and II versus the number of children at Stage III within each group, significantly more immersion children were found to be concrete operational on classification of squares and circles than were 75-minute children.

3. No significant difference was found between the mean scores of the 75-minute children and the mean scores of the immersion children on the error and latency measures of the Matching Familiar Figures Test.

4. Children classified as "reflective" on the Matching Familiar Figures Test were found to make significantly higher mean scores on the arithmetic concepts subtest of the Metropolitan Achievement Test in comparison with children classified as slow-inaccurate; and on the numeration subtest they made significantly higher scores than fast-inaccurate and slow-inaccurate children.

5. If separate multiple Rs were computed for the 75-minute group and the immersion group, and compared to each other and to the multiple R for the total group, it was found that program did not act as a moderator variable.

6. No significant difference was found between the multiple Rs for the two random subgroups and between the multiple R for each random subgroup separately and the multiple R for the total group.

7. No significant difference was found between the two cross-validation multiple Rs or between each cross-validation multiple R separately and the multiple R for the total group.

In general, the major conclusion from this study is that there are no differences between the 75-minute and immersion children, though a number of trends in mean scores were evident. Contrary to evidence cited in the review of the literature, the immersion program did not enhance arithmetic achievement as measured by standardized arithmetic tests.

A number of reasons were suggested for the lack of expected differences in arithmetic achievement. The possibility was raised that the pilot and follow-up immersion and 75-minute groups studied by Edwards and Casserly (1971, 1972) may have differed initially in arithmetic ability or at least in school readiness. It was apparent from the present study that readiness scores correlated more highly with arithmetic achievement than did intelligence. Though readiness scores were available for the children in this study, it would obviously be better to use arithmetic-achievement as measured by the Metropolitan Achievement Test at the end of one grade to statistically control differences at the end of the next grade. This would allow one to answer more accurately the question whether French language immersion facilitates arithmetic-ability or whether the superiority of the immersion children in previous studies has been due to initial differences in arithmetic-ability.

Other questions that might be considered are whether any superiority in arithmetic on the part of the immersion children is due to some inherent relationship between arithmetic and the structure of language (Piaget, 1965); or whether arithmetic taps relatively language-free abilities and thus does not penalize the immersion children as would a test of English-language abilities. Certainly, the majority of the concrete operational tests were language bound in this study, with the exception perhaps of seriation where E actually demonstrated what was required. But in this respect, comprehension difficulties did not seem to be apparent in the immersion children.

Though the scores of the immersion children were generally higher on concrete operations than those of the 75-minute children, this difference disappeared when adjustments were made for initial differences in readiness and intelligence. This trend reflected that found for arithmetic-achievement. To this extent the study was not able to clearly answer the question whether, if differences were found between the two groups on arithmetic-achievement, was this, in turn, reflected in differences in cognitive operations. This obviously needs further investigation, both by following up this population of children and with other populations of children in immersion programs. However, it is the opinion of the author that if concrete operational tests are used that are closely allied to arithmetic-abilities, any enhancement of one ability is likely to be concomitant with enhancement of the other ability.

Another aspect for consideration is in what ways the home environment of the immersion and 75-minute children may differ in terms of stimulation, encouragement of learning, development of achievement motivation or other measures that have been found to be related to school achievement. This information was not available either for this study or for the pilot and follow-up children in Edwards and Casserly's (1971, 1972) study. It is possible that these variables may have been more operative in the early years of the French-language immersion program, but that now a wider range of parents may be enrolling their children in French immersion. However, this is just speculation. Examination of these variables and their contribution to any differences found between immersion and 75-minute groups should be considered in a future study.

It is apparent from examination of the predictor variables that enter into the multiple regression equations for the two program groups, and also for the two random subgroups, that numeration and multiple seriation seem to be the most reliable predictors of arithmetic-achievement. No other studies have been found by the author which use this combination of cognitive variables in a multiple regression equation to predict arithmetic-achievement. Validation on another population is indicated. It is possible, given the age-related nature of Piaget tests, that these two tests would only be predictive of arithmetic-ability at a certain narrow age level. For instance, at the age level of the subjects in the present study, which is at the point of entering concrete operational thinking.

In conclusion, this study supports the findings of Edwards and Casserly (1971, 1972) to the extent that French immersion does not seem to retard the academic or cognitive abilities of children in comparison with those enrolled in a regular program. However, this may be because the children enrolled in the French immersion program do not represent the wide range of ability levels found in the School Board as a whole. Further investigation of the effect of French immersion on cognitive abilities such as those measured by Piagetian tests and particularly follow-up of this group of children to determine whether their cognitive abilities continue to remain the same or diverge is recommended.

REFERENCES

Aiken, L. R. Intellective variables and mathematics achievement: Directions for research. Journal of School Psychology, 1971, 9(2), 201-213.

Akers, J. C. A predictive study of the Metropolitan Readiness Test. (Doctoral dissertation, Oklahoma State University) Ann Arbor, Mich.: University Microfilms, 1970, No. 70-21518.

Almy, M., Chittenden, E., & Miller, P. Young children's thinking. New York: Teacher's College, Columbia University, 1967.

Attwell, A. A., Orpet, R. E., & Meyers, C. E. Kindergarten behaviour ratings as a predictor of academic achievement. Journal of School Psychology, 1967, 6(1), 43-46.

Balkan, L. Les effets du bilinguisme Française-Anglais sur les aptitudes intellectuelles. Brussels: Aimov, 1970.

Barik, H. C., & Swain, M. Bilingual education project: Interim report on the spring 1972 testing. Toronto: Bilingual Education Project Modern Language Centre, O.I.S.E., 1972.

Blishen, B. R. The construction and use of an occupational class scale. Canadian Journal of Economics and Political Science, 1958, 24(4), 519-530.

Cathcart, W. G., & Liedtke, W. Reflectiveness/impulsiveness and mathematics achievement. The Arithmetic Teacher, 1969, December, 563-567.

Cobb, J. A. Relationship of discrete classroom behaviours to fourth grade academic achievement. Journal of Educational Research, 1972, 63(1), 74-80.

Cohen, R. Contribution a l'étude differentielle de l'efficacité des methodes pedagogigues. Unpublished doctoral dissertation, University of Caen, 1969.

Darcy, N. T. A review of the literature on the effects of bilingualism upon the measurement of intelligence. Journal of Genetic Psychology, 1953, 82, 21-57.

Darcy, N. T. Bilingualism and the measurement of intelligence: Review of a decade of research. Journal of Genetic Psychology, 1963, 103, 259-282.

Dayhaw, L. T. Manuel de statistique. Ottawa: Editions de l'Université d'Ottawa, 1969.

Diebold, R. The consequences of early bilingualism in cognitive development and personality formation. In E. Norbeck, D. Price-Williams, & W. M. McCord (Eds.), The study of personality. Toronto: Holt, Rinehart & Winston, 1968.

Edwards, H. P., & Casserly, M. C. Research and evaluation of the French program 1970-1971 annual report. Ottawa: The Roman Catholic Separate School Board, 1971.

Edwards, H. P., & Casserly, M. C. Research and evaluation of second language programs 1971-1972 annual report. Ottawa: The Roman Catholic Separate School Board, 1972.

Elkind, D. Children's discovery of the conservation of mass, weight and volume: Piaget replication study II. Journal of Genetic Psychology, 1961, 98, 219-227.

Elkind, D. Discrimination, seriation and numeration of size and dimensional differences in young children: Piaget replication study VI. Journal of Genetic Psychology, 1964, 104, 275-296.

Flavell, J. The developmental psychology of Jean Piaget. Toronto: Van Nostrand Reinhold, 1963.

Game, A. Second language programs. In H. P. Edwards & M. C. Casserly, Research and evaluation of second language programs annual report 1971-1972. Ottawa: The Roman Catholic Separate School Board, 1972.

Goldschmid, M. L. Different types of conservation and non-conservation and their relation to age, sex, IQ and vocabulary. Child Development, 1967, 38(4), 1229-1246.

Goodnow, J., & Bethon, G. The effects of schooling and intelligence. Child Development, 1966, 37, 574-578.

Greenfield, P. M. On culture and conservation. In J. S. Bruner, R. R. Olver, & P. M. Greenfield (Eds.), Studies in cognitive growth. New York: Wiley, 1966.

Guilford, J. P. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1965.

Hagenson, S. L. The relationship of first grade readiness and achievement scores based on sex, race and age. (Doctoral dissertation, University of Southern Mississippi) Ann Arbor, Mich.: University Microfilms, 1968, No. 68-2935.

Harris, D. B. Children's drawings as measures of intellectual maturity. New York: Harcourt, Brace & World, 1963.

Hood, B. H. An experimental study of Piaget's theory of the development of number in children. British Journal of Psychology, 1962, 53(3), 273-286.

Ianco-Worral, A. D. Bilingualism and cognitive development. Child Development, 1972, 43, 1390-1400.

Jones, W. R. Bilingualism in Welsh education. Cardiff: University of Wales Press, 1966.

Kagan, J. Reflection-impulsivity and reading ability in primary grade children. Child Development, 1965, 36, 609-629.

Kagan, J., Rosman, B., Day, D., Albert J., & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. Psychological Monographs, 1964, 78 (1, Whole No. 578), 1-37.

Kagan, J., Moss, H. A., & Siegel, I. E. Psychological significance of styles of conceptualization. Society for Research in Child Development Monographs, 1963, 28 (86, Whole No. 2), 73-112.

Kagan, J., Pearson, L., & Welch, L. Conceptual impulsivity and inductive reasoning. Child Development, 1966, 37, 583-594.

Keith, V. Design and analysis in experimentation. Ottawa: University of Ottawa Press, 1972.

Kirk, R. E. Experimental design: Procedures for the behavioural sciences. Belmont, Cal.: Brooks/Cole, 1968.

Kofsky, E. A. A scalogram study of classification development. In I. E. Siegel & F. H. Hooper (Eds.), Logical thinking in children. London: Holt, Rinehart & Winston, 1968.

Lahaderne, H. M. Attitudinal and intellectual correlates of attention: A study of four sixth grade classrooms. Journal of Educational Psychology, 1968, 59, 320-324.

Lambert, W. E. St. Lambert elementary school, St. Lambert, Que. French immersion. Unpublished report of Dr. W. E. Lambert, 1970.

Lambert, W. E., Just, M., & Segalowitz, N. Some cognitive consequences of following the curricula of early school grades in a foreign language. Monograph Series on Languages and Linguistics. J. E. Altus (Ed.), 1970, No. 23.

Lambert, W. E., & MacNamara, J. Some cognitive consequences of following a first grade curriculum in a second language. Journal of Educational Psychology, 1969, 60, 86-96.

Lambert, W. E., Tucker, G. R., d'Angeljan, A., & Silny, F. Cognitive and attitudinal consequences of bilingual schooling: The St. Lambert project through grade 5. Unpublished Report, 1972.

Liedtke, W., & Nelson, D. L., Concept formation and bilingualism. Alberta Journal of Educational Research, 1968, 14, 225-232.

Lovell, K., Mitchell, B., & Everett, I. R. An experimental study of the growth of some logical structures. In I. E. Siegel & F. H. Hooper (Eds.), Logical thinking in children. London: Holt, Rinehart & Winston, 1968.

MacNamara, J. Bilingualism and primary education: A study of Irish experience. Edinburgh: University of Edinburgh Press, 1966.

Manuel, H. T. A comparison of Spanish-speaking and English-speaking children in reading and arithmetic. Journal of Applied Psychology, 1935, 19, 189-202.

McInnis, C. E. Does either a moderator model or a multiplicative model improve prediction over the usual multiple regression model in the context of an intensive French language training program. Unpublished doctoral dissertation, University of Ottawa, 1972.

Peal, E., & Lambert, W. E. The relationship of bilingualism to intelligence. Psychological Monographs, 1962 76 (27, Whole No. 546).

Peters, D. L. Task variations and individual differences in Piaget's conservation of number. Merrill-Palmer Quarterly of Behaviour and Development, 1967, 13(4), 295-308.

Piaget, J. The child's conception of number. London: Routledge & Kegan Paul, 1965.

Piaget, J., & Inhelder, B. The early growth of logic in the child. London: Routledge & Kegan Paul, 1970.

Rigg, M. Some further data on the language handicap. Journal of Educational Research, 1928, 19, 252-256.

Robinson, I. C. The acquisition of quantitative concepts in children. (Doctoral dissertation, University of Southern California) Ann Arbor, Mich.: University Microfilms, 1968, No. 68-1692.

Rothenberg, B. B. Conservation of number among four and five year old children: Some methodological considerations. Child Development, 1969, 40, 383-406.

Scott, R. Social class, race, seriating and reading readiness: A study of their relationship at the kindergarten level. Journal of Genetic Psychology, 1969, 115, 87-96.

Skemp, R. R. Reflective intelligence and mathematics. British Journal of Psychology, 1961, 31, 45-55.

Spitler, G. J. An investigation of various cognitive styles and the implications for mathematics education. (Doctoral dissertation, Wayne State University) Ann Arbor, Mich.: University Microfilms, 1971, No. 71-17317.

Steffe, L. P. The relationship of conservation of numerousness to problem solving abilities of first grade children. The Arithmetic Teacher, 1968, January, 47-52.

Stevens, R. P. Bilingual experience and psycholinguistic ability. McGill Journal of Education, 1966, 1(2), 74-87.

Torrance, E. P., Wu, J. J., Gowan, J. C., & Aliotti, N. C. Creative functioning of monolingual and bilingual children in Singapore. Journal of Educational Psychology, 1970, 61, 72-75.

Trevino, B.A. An analysis of the effectiveness of a bilingual program in the teaching of mathematics in primary grades. (Doctoral dissertation, University of Texas) Ann Arbor, Mich.: University Microfilms, 1968, No. 68-10,893.

Van Engen, H., & Steffe, L. First grade children's concept of addition of natural numbers. Research and Development Centre for Learning and Re-education, Technical Report No. 4, Madison, Wisconsin: University of Wisconsin, 1966.

Zedeck, S. Problems with the use of "moderator" variables. Psychological Bulletin, 1971, 76(4), 295-310.

PARENT QUESTIONNAIRE AND COVERING LETTER

THE OTTAWA ROMAN CATHOLIC
SEPARATE SCHOOL BOARD
140 Cumberland K1N 7G9
Ottawa, Ontario

LE CONSEIL DES ECOLES
SEPAREES CATHOLIQUES D'OTTAWA
339 Wilbrod K1N 6M4
Ottawa, Ontario

May 23, 1973

Dear Parents:

As you may already know, the second language programs of this School Board are undergoing development and are being evaluated over an eight-year period by a research team formed through the cooperation of the School Board, the University of Ottawa, and the Ontario Ministry of Education. Within this context, we wish at this time to solicit your cooperation in replying to the enclosed questionnaire.

There is evidence that a child's scholastic test scores may be affected by, among other things, the educational level and occupation of the parents. We are, therefore, asking you to help us by completing the attached questionnaire. Each questionnaire has been number coded to ensure complete anonymity of the information obtained from you.

Though completion of the questionnaire is, of course, voluntary, it is important that we obtain this information so that we can more fully interpret the test results used in our research on second language programs. The results of this research will be discussed at one of the regular meetings for parents.

The questionnaire should be returned in the enclosed stamped, self-addressed envelope by June 10, 1973. If you would like any further information, please call Miss Casserly, Ottawa Separate School Board, at 237-5660 Ext. 227. We are grateful for your cooperation.

Yours truly,

H. P. Edwards, Ph.D.,
Project Director

M. C. Casserly, M.A.,
Research Officer,
Second Language Learning
Evaluation

VS/db

APPENDIX 1

PARENT QUESTIONNAIRE
1973

I. OCCUPATION (Please be specific.):

Father's occupation________________________________

Mother's occupation________________________________

II EDUCATION (Please indicate by a check mark the highest level attained.)

		Father	Mother
Elementary School Grades	1 - 8	____	____
High School Grades	9 -10		
	11-12	____	____
	13	____	____
Community College or equivalent (e.g. Teachers' College, School of Nursing)			
	(1) partial	____	____
	(2) completed	____	____
Bachelor degree			
	(1) partial	____	____
	(2) completed	____	____
Post graduate education			
	(1) partial	____	____
	(2) completed	____	____

APPENDIX 2

MATCHING FAMILIAR FIGURES TEST ADMINISTRATION

"I am going to show you a picture of something you know and then some pictures that look like it. You will have to point to the picture on this bottom page (point) that is just like the one on this top page (point). Let's do some for practice." E shows practice items and helps the child to find the correct answer. "Now we are going to do some that are a little bit harder. You will see a picture on top and six pictures on the bottom. Find the one that is just like the one on top and point to it."

E will record latency to first response to the half-second, total number of errors for each item and the order in which the errors are made. If S is correct, E will praise. If wrong, E will say, "No, that is not the right one. Find the one that is just like this one (point)." Continue to code responses (not times) until child makes a maximum of six errors or gets the item correct. If incorrect, E will show the right answer.

APPENDIX 3

TEST SCHEDULE - CONCRETE OPERATIONAL THINKING

Conservation of Substance

1. Produce four clay balls, two the same size.

Q. "Can you choose two that have the same amount of plasticine in them, that are the same size?"

2. If the child replies that they are all different, say:

"True, in colour they are different, but I'm thinking of size, of amounts of plasticine. Which two are equal?"

DISCONTINUE EXPERIMENT IF CHILD CANNOT CHOOSE THEM.

Q. "Is there the same amount of clay in this ball as in that one?"

ANSWER: Yes _____ CHILD MUST AGREE OR DISCONTINUE

No _____

3. "Now watch what I do." (Roll the ball into a sausage.)

Q. "Is there the same amount in the sausage as in the ball?"

ANSWER: Yes _____

No _____

Why ______________________________

Q. "Does one have more clay?" IF YES: "Which one?" "How can you tell."

ANSWER Yes _____

No _____

Why ______________________________

4. Roll the sausage some more.

Q. "Is there the same amount of clay in the sausage as in the ball?" "Why?"

ANSWER: Yes _____

No _____

Why ______________________________

Q. "Does one have more clay?" IF YES: "Which one?" "Why?"

ANSWER: Yes _____

No _____

Why ______________________________

APPENDIX 3

All and Some

Materials: 4 blue squares, 2 blue triangles, 3 red triangles - 5 cm.

1. Spread materials out in front of child and ask him to find all the different colours and shapes.

Q. "Can you show me all the reds, can you show me all the blues, can you show me all the triangles, can you show me all the squares?"

2. Then muddle materials up again, not in neat piles.

Q. "Are all the reds triangles?" "Why?"

ANSWER: Yes _____

No _____

Why _____

Q. "Are all the triangles red?" "Why?"

ANSWER: Yes _____

No _____

Why _____

Q. "Are all the squares blue?" "Why?"

ANSWER: Yes _____

No _____

Why _____

Q. "Are all the blues square?" "Why?"

ANSWER: Yes _____

No _____

Why _____

APPENDIX 3

Seriation

1. Demonstrate seriation of staircase.

Q. "Now watch what I am going to do. I am making a staircase."

2. Break up the staircase.

Q. "Now can you make a staircase just like I did, just like mine?"

ANSWER:

3. If child is unable to seriate 10, remove five smallest.

Q. "Now see if you can make me a staircase."

4. If child correctly seriates 5, break up staircase and give child 7 slats.

ANSWER:

5. If child correctly seriates 7, break up staircase and give child 10 slats.

ANSWER:

6. Produce second series.

Q. "Can you put these slats that have been left out in their right places in the staircase? Can you put them where they fit in the staircase?"

ANSWER:

If the child doesn't understand that the slats should go in between original, prompt once: "Make one big, long staircase. Put the new stairs in between the old staircase." However, do not continue prompting <u>ad infinitum</u>.

Child prompted re fitting second series: Yes ______

No ______

Numeration

Materials: 10 slats and plastic farmer

1. Place the farmer on bottom staircase and ask child how many stairs farmer has to climb to reach each stair.

Q. "If the farmer starts at the bottom of the staircase, how many stairs does he have to climb to reach this stair, and this stair, etc. etc.?"

ANSWER: Correct ____

Incorrect ____

Q. "How many stairs does the farmer have to climb to reach this stair (4)?"

ANSWER: 4 ____
3 ____
Other ____

Q. "How many stairs does the farmer have to climb to reach this stair (7)?"

ANSWER: 7 ____
6 ____
Other ____

2. Point to second stair of disarranged slats.

Q. "How many stairs does the farmer have to climb to reach this stair if the staircase is together like it was before?" "How do you know?"

ANSWER: 2 ____ Reseriates to 2 ____
Other ____ Reseriates beyond 2 ____
Doesn't reseriate ____
Answers from memory ____

WHY ________________________________

Reconstruction suggested: "How could you find the answer - remake the staircase?"

ANSWER: 2 ____ Reseriates to 2 ____
Other ____ Reseriates beyond 2 ____
Doesn't reseriate ____

WHY ________________________________

Conservation of Weight

1. Produce four clay balls, two the same size.

Q. "Can you choose two that weigh the same, that are the same weight?"

2. If the child cannot choose two that weigh the same:

Q. "Can you make two the same weight?"

3. IF CHILD CANNOT AGREE ON TWO HAVING THE SAME WEIGHT, DISCONTINUE EXPERIMENT.

Q. "Do they weigh the same now?"

ANSWER: Yes ____ DISCONTINUE IF CHILD SAYS NO.

No ____

4. Flatten one ball into a pancake.

Q. "Are they the same weight now?" "How can you tell."

ANSWER: Yes ____

No ____

Why ____

Q. "Does one weigh more now?" "How can you tell?"

ANSWER: Yes ____

No ____

Why ____

5. Flatten pancake more.

Q. "Are they the same weight now?" "How can you tell?"

ANSWER: Yes ____

No ____

Why ____

Q. "Does one weigh more now?" "How can you tell?"

ANSWER: Yes ____

No ____

Why ______________________________________

Classification of Squares & Circles

Materials: 4 large blue squares, 4 small blue squares
3 large blue circles, 3 small blue circles,
1 large red circle.

PART II materials: 1 large red square, 1 small red square,
1 small red circle.

1. Put all the first materials in disarray in front of the <u>S</u> and ask him to put them in groups.

Q. "Put together the ones you think go with each other, put them into groups."

ANSWER:

Q. "Can you put them into two piles that are different from each other? Can you put them into two different groups? Why did you put them like that--why do you think they go together?"

ANSWER: Colour ____
Shape ____
Size ____

Reason ____

Q. "Can you find another way to put them into two groups. You have found one way, can you find another way?" "Why?"

ANSWER: Colour ____
Shape ____
Size ____

Reason ____

Q. "Let's see if you can find a third way to put them into two groups. You have found two ways. Is there another way you can make them into two groups?" "Why&"

ANSWER: Colour ____
Shape ____
Size ____

Reason ____

2. If the child cannot find a third way, introduce 1 large red square, 1 small red square, 1 small red circle.

Q. "Can you find another way to put them into two groups?" "Why?"

ANSWER: Colour ____
Shape ____
Size ____

Reason ________________________________

3. If the S makes a grouping by colour, then take away the three shapes introduced above and ask if he can make a group by colour with only 1 large red circle.

Q. "Can you put them into two groups by colour now?"

ANSWER: Yes ____
No ____

4. If the child does not make a colour grouping when 3 new shapes are introduced above, suggest it.

Q. "How about putting them into groups of different colours?"

5. If child agrees that this is O.K. then take away 3 extra shapes and ask child if he can make groups with only 1 large red circle.

Q. "Can you put them into two groups by colour now?"

ANSWER: Yes ____
No ____

Multiple Seriation

Material: There are a total of four different colours ordered from A(lightest) to D(darkest). Within each colour there are four different sizes ordered from 1 to 4.

1. Present the 16 blocks in a random arrangement. Ask the child to arrange them as he thinks best. If the child does not seem to understand what is required, ask him to put them in order; if he still hesitates, suggest that he could group them.

2. "Can you put all the light ones together, can you put all the dark ones together?"

3. "Can you put all the big ones together, can you put all the small ones together?"

4. If the child groups the blocks by only one dimension (size or colour), suggest that he group them by the other dimension using words as above. "Can you put all the . . . together?"

5. If the child is still unable to do the task, arrange the blocks in one dimension of colour, i.e. lightest colour ordered in size. If the child is still unable to construct the matrix, do the second dimension of size and ask him to fill in the rest of the table.

ANSWER:	a1, a2, a3, a4	(a) orders blocks with no trouble
	b1, b2, b3, b4	(b) needs prompting
	c1, c2, c3, c4	(c) E seriates one dimension
	d1, d2, d3, d4	(d) E seriates two dimensions
		(e) unable to complete task