BILINGUALISM AND REASONING ABILITY

BY

Monique Marie Gabrielle Trites

B.A., The University of British Columbia, 1983
Teaching Certif., The University of British Columbia, 1984

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Education

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

October 1986

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Educational Psychology and Special Education

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date October 16, 1986

ABSTRACT

The purposes of the present study were (1) to investigate relationships between levels of verbal and non verbal reasoning ability and second language acquisition and (2) to examine the concurrent validity of the Test Of Nonverbal Intelligence (TONI) using the WISC-R and the Children's Word Finding Test as criteria and to verify the equivalence of the two forms of the TONI.

The present study was conducted because of the controversy that exists in the literature, over the relationship between bilingualism and cognitive ability. While previous research used "general intelligence" tests to investigate differences between unilinguals and bilinguals, the present study examined a more precise construct: verbal and non-verbal reasoning ability in Grade 3 French Immersion students as compared to Grade 3 Non French Immersion Students.

All students came from homes in which English was the only language spoken, had been continuously enrolled from kindergarten through the end of Grade 3 in their current academic programs, and were not attending Learning Assistance centers. The two groups came from families with high socioeconomic status.

An analysis of variance was used to compare the two groups on the WISC-R, the TONI and the Children's Word Finding Test. Correlations between scores on the three tests were calculated to verify the level of equivalence of the two forms of the TONI, and of the validity of

the TONI concurrently with the WISC-R and with the Children's Word Finding Test.

The French Immersion group scored higher than the Non-French Immersion group on the TONI A ($p = .07$) and the Children's Word Finding Test ($p = .10$) and significantly higher on the Performance Scale and Full Scale of the WISC-R. After controlling for variability on cognitive ability levels (WISC-R), the French Immersion group still scored higher than the Non-French Immersion group on the TONI A and the Children's Word Finding Test ($p = .09$).

The TONI correlated poorly with the WISC-R, as well as with the Children's Word Finding Test. The low correlation of the TONI with the WISC-R was explained by the difference in the range of the age interval of the norms of each test.

The coefficient of equivalence of the two forms of the TONI was significantly lower than that reported by the authors of the technical manual. Furthermore, an order effect was found for both French Immersion and Non-French Immersion groups, when the Form A of the TONI was administered first (This was not however, the case when the Form B of the TONI was administered first). When the TONI A was given first the correlation between the scores of the TONI A and TONI B was high for the French Immersion group but not for the Non-French Immersion group.

It was concluded that although interpretation of the results is confounded by possible initial differences between groups, the French

Immersion group demonstrated a tendency towards better verbal and non-verbal inductive and deductive reasoning ability ($p = .09$). Further, it was concluded that the two forms of the TONI are not equivalent and that there is an order effect when TONI A is administered first, but not when TONI B is administered first. Further research is needed on this test before it can be considered a reliable substitute for the WISC-R.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

ACKNOWLEDGEMENTS

I wish to thank Dr. Patricia Arlin and my graduate student colleagues and friends for their valuable comments, particularly at the proposal stage of my thesis research. I am indebted to Lydia Bergant, Suzanne Jacobsen, Cathy Nolan, and Maria Iaquinta for their help in testing and scoring.

I greatly appreciated the help of Dr. William Bruneau in clarifying my thoughts and expression of ideas through numerous discussions.

I am especially grateful to my husband for his support, help and encouragement all through the preparation of this project.

I wish also to express my appreciation to my research supervisor, Dr. Julianne Conry, for sharing her expertise and for her availability, and to my committee members: Dr. Robert Conry, who helped me with the methodology and the use of the computer, and Dr Robert Roy for his support and his perspicacious input.

A final word of appreciation goes to those Vancouver School Board students, principals, teachers, and research department officials who contributed their time and energy to the project.

CHAPTER I

INTRODUCTION

There has been a lengthy controversy concerning the relationship between bilingualism and cognitive development. Some researchers have argued that bilingualism enhances cognitive ability while others say that bilingualism hinders it. Saer (1922) and Smith (1923), for instance, studied Welsh children whose mother tongue was Irish, but who attended school in English. Saer stated that "mental confusion" occurred more often in the bilingual children than in the monoglots, and Smith concluded that bilingualism seemed to be an intellectual disadvantage. Nowadays, when French Immersion is becoming more and more popular in Canada, parents, researchers and professionals are once again interested in the question whether bilingualism enhances cognitive development or not.

There is considerable disagreement as to the answer. For example, Lambert and Peal (1962) found, on the one hand, that English-French bilingual children performed better than unilingual on a type of non-verbal test involving concept-formation or symbolic "flexibility" as defined by Ahmed (1954) as the ability to reorganize mentally the elements of a problem. English-French bilingual children also performed better than unilinguals on verbal intelligence tests. However, after the two groups had been matched on nonverbal test results and on the socioeconomic status of the parents, the authors found no significant difference between the two groups on verbal intelligence tests.

Barik and Swain's (1976) longitudinal study found no difference over a three year period (Grade 1 to 3) between regular English program students and French Immersion students with respect to I.Q. measure or subtest scores for children from grade 1 to 3 when the scores had been adjusted for initial I.Q. differences. They classified bilingual subjects as "high" French achievers and "low" French achievers, and looked at the increase of their I.Q. over the three year period. High achievers were those who scored at the top third on four tests of French performance; low achievers were those who scored at the bottom third on the same four tests of French performance. The analysis of the results showed a significant increase of the I.Q. of the high French achievers group over the three year period, while there was no significant increase for the low French achievers.

These findings suggest that bilingualism may increase the I.Q. of those bilinguals who have attained a certain level of proficiency in their second language. However, the authors do not take a stand and conclude by saying that "the results of the analysis suggest that the relationship between bilingualism and cognitive functioning is a complex one" and that "the issue warrants further investigation." (Barik and Swain, 1976, p. 261)

The opposing findings on the relationship between bilingualism and cognitive functioning can in part be explained by sampling and methodological differences in the studies, such as the type of test to measure cognitive ability, the matching of the subjects on socioeconomic status, the definition of bilingualism, and so on.

The influence of socioeconomic status on I.Q. has been demonstrated in several studies. However, in most earlier studies involving bilinguals and in some of more recent ones, unilinguals and bilinguals have not been matched on the socioeconomic status of their parents. This methodological step is especially important for studies involving French Immersion children because they generally come from higher socioeconomic background than do children in regular classrooms (Trites and Price, 1980).

Another major problem in this kind of research may be the fairness of the tests used regarding the language and culture of the population concerned. Because intelligence tests are language-oriented, they might be unfair to French Immersion students when given in English before these students are formally taught English language arts. Nielsen (1983) conducted research to see whether the WISC-R test, when administered in English to French Immersion grade 2 students, puts them at a disadvantage. She compared results on WISC-R tests given in English to grade 2 French Immersion students with those of grade 2 students from the regular English program. Although the French Immersion group mean was higher, analysis of the 12 subtests and three IQ scores revealed no significant differences between the two groups. The conclusion was that the WISC-R test, when administered in English to French Immersion students does not put them at a disadvantage.

One solution to the problem of the fairness of the test used, has been to administer a translation of the test. In her review of the effect of bilingualism upon the measurement of intelligence, Darcy

(1953) reported on studies using translations of standard intelligence tests. She said that the results of subjects tested in their second language were lower than their results when tested in their first language. However, she suggested that "the translation of a standardized test is not an equivalent to the test in the language in which it was originally standardized." (1953, p. 25)

Another solution to the problem of fairness has been to use nonverbal tests as a measure to compare the general intelligence of bilingual and unilingual children. However, there is only a moderate correlation between these tests and verbal intelligence tests such as the WISC-R or the Binet. Lambert and Tucker (1972) used the Raven's Progressive Matrices test as a nonverbal I.Q. test. However, the purpose of the test is to measure reasoning capacity and it yields only percentile scores. Items do not test verbal or mathematical ability; they were constructed to deal with visual spatial reasoning alone. Raven's intention was to provide a reliable estimate of a person's capacity to think clearly when allowed to work steadily, to form comparisons and reason by analogy (Raven, 1976). Although correlations between the Raven's Progressive Matrices and the WISC-R range from +.54 to +.86 (Raven, 1976), the concurrent validity of the test with standard intelligence tests has not been convincingly established. The Raven's Progressive Matrices cannot be used as a substitute for the WISC-R since the latter offers much greater versatility in its assessments of mental activity (Mehrota, 1968); the Raven and the WISC-R test have been shown not to be equivalent (Holmes, 1981)

The use of intelligence tests to compare cognitive functioning between the two groups as traditional studies have done, presents another difficultly. The variety of definitions for the concept of intelligence and the variety of tests used to measure it, have caused confusion. Furthermore, abilities measured by intelligence tests seem too varied, easily to permit the discovery of consistent differences between bilingual and unilingual children. Memory, visual perception, visual motor integration, organization, knowledge of basic information, problem solving in social context, as well as conceptual and logical reasoning, are only a few examples of the abilities involved in standard intelligence tests. When comparing two groups using this kind of test, differences in results on sub-sections of the tests, both positive and negative, may balance each other out. Thus, differences among final overall results may not be significant.

Generally, studies comparing bilinguals and unilinguals did not involve students with learning problems unless research had directly to do with learning disabilities. The choice of subjects for the samples who have average or above-average academic success reduces the chances of finding differences on intelligence tests between the samples. For the preceeding reasons, it seems to be more appropriate to compare bilinguals and unilinguals on specific abilities such as inductive and deductive reasoning, which might be developed through the acquisition of a second language.

Lambert and Tucker (1972) said that bilingual children from French Immersion schools have developed "linguistic detective skills" which they define as "an attentive, patient, inductive concern with

words, meanings and linguistic regularities" (1972, p.208). The authors based this inference on the better-than-expected results of bilinguals on tests of French word discrimination, listening comprehension, and decoding. It seems natural that children develop such skills in French Immersion situation where all subjects are taught in French from Kindergarten to the end of grade 3. Teachers use gesture, mimes and pictures to communicate their message; the children have to pay very close attention to these clues, and use their inductive and deductive reasoning ability in order to understand the meaning of what is said.

The goal of the study of reasoning ability here proposed was to demonstrate whether those who acquire a second language develop superior reasoning ability. It was also to contribute to the historic dialogue concerning the possible correlation between second language learning and intelligence. This study should offer a new perspective to investigators on how best to test for differences in cognitive ability between bilinguals and unilinguals. Testing for reasoning skills as proposed here emphasizes specific thinking processes rather than learned facts.

Statement of the Problem

The goal of this study was to compare levels of reasoning ability at the end of grade 3 of bilingual children and unilingual children matched on socioeconomic background.

Bilinguals children were defined as students registered in early French Immersion since Kindergarten and coming from anglophone families where the only language used in the home was English.

Unilingual children were defined as students registered in a regular English program since Kindergarten and coming from anglophone families where the only language used in the home was English .

Reasoning ability was measured according to results obtained by the students on a nonverbal reasoning ability test (the Test of Nonverbal Intelligence) and on a verbal reasoning test (the Children's Word Finding Test), both tests involving inductive and deductive reasoning.

There was one independent variable, the condition bilingual or unilingual. There were two dependent variables, the level of verbal reasoning and the level of nonverbal reasoning.

The objective of the study was to serve the theoretical purpose of answering the question whether there is a relationship between bilingualism and the level of verbal and nonverbal reasoning ability.

Furthermore, the use of the TONI as a nonverbal reasoning test led to verifying the equivalence of Form A and Form B of the test and to establish its concurrent validity with the WISC-R and with the Children's Word-Finding Test. This was important to do because no study had been done on the concurrent validity of the TONI with the WISC-R for normal populations.

Although it is called an intelligence test, the TONI's items focus exclusively on problem solving as in the Raven's Progressive Matrices. The authors pointed out that their test of nonverbal intelligence measures a small piece of the construct of intelligence (manual p.25). The concurrent validity of the T.O.N.I. with the WISC-R has been established only with very small samples (N=10 to 16) of special education groups (eg; deaf, learning disabled, mentally retarded) and ranges from +.46 to +.95 9 (manual, p. 14). For normal groups, as opposed to special education groups, scores on the TONI are compared only to group I.Q. (Otis Lennon Mental Ability Test). Reviewers of the test in the last Mental Measurement Yearbook said that "T.O.N.I. needs much more empirical evidence of concurrent validity than has been presented to date" (Mayo, 1985, p.1581).

Comparing the correlations of the TONI with the WISC-R and the TONI with the Children's Word Finding Test shed some light on the meaning of the results on the TONI for normal groups.

In summary, the purpose of the study was to answer three questions:

1. Is there a difference on the level of reasoning ability between grade 3 French Immersion students and grade 3 Non-French Immersion students, and if there is one, in which direction?

2. What is the concurrent validity of the TONI test with the WISC-R test and with the Children Word Finding Test? Are the

validity coefficients different in magnitude between French Immersion and Non-French Immersion?

3. Are the two forms of the TONI test equivalent?

CHAPTER II

REVIEW OF THE LITERATURE

This Chapter reviews the pertinent literature concerning bilingualism and cognitive ability. It begins by briefly reviewing earlier research. It then proceeds to outline studies that may have important implications on the variables used in this study; in particular, the definition of bilingualism, the socioeconomic status, the social context, the kinds of tests used, and the number of books owned by the family. The Chapter concludes by reviewing the most recently completed research.

Bilingualism and Cognitive Development and Ability, Early Studies

The central question of this study requires prior investigation of a broader issue: that of the role of language in the development of cognition. This broader problem interested, among others, both Piaget (1923) and Whorf (1956). Although Piaget argued that language has only a minimal causal role in cognitive development, Whorf suggested that perception of the environment and its mental representation depend essentially on language. Vigotsky (1962) also devoted considerable time to studying the relationship between language and cognition. He saw language as having a leading role in the development of cognition. In brief these researchers may be categorized in two "schools": first Whorf's school, which claims that language precedes and is necessary to cognitive development,

and Piaget's school, which defends the view that language is not closely related to cognitive development.

Because these theories are in some respects opposed, various researchers have subsequently conducted empirical studies with bilingual children to test the viability of each. Bain (1975) attempted to reconcile these two schools in his study of six year old bilinguals and unilinguals. The subjects' task involved the discovery of rules and their generalization to other situations. He found that at the preoperational stage the difference between bilinguals and unilinguals in discovery time was significant in favour of the bilinguals; however, at the concrete operation stage the bilingual group had a slight but non-statistically significant advantage. There were no differences in transfer time at either stage. He concluded by saying:

"The differences in discovery time were seen as support for Vigotsky's view that specific kinds of language experience give unique direction to development. The lack of difference in transfer time was seen as support for Piaget's view that developmental level of operative structures determine performance limits of previously learnt material." (1975, p.13).

The introduction of intelligence tests in the 1920's provided another impetus for study, especially of differences in results of bilinguals and unilinguals on these tests, and of any possible correlation between intelligence and bilingualism.

Before 1960, most such studies reported lower results on IQ tests for bilinguals than for unilinguals. Darcy's review (1953) is valuable as a catalogue of earlier research ordered in three categories: studies which found a "favorable effect", studies which found an "unfavorable effect", and studies which found "no effect" of bilingualism upon measurement of intelligence. Following are two characteristic studies of the ones
she presents in her review.

In Wales, Saer (1923) was the first to publish research on bilingualism and intelligence. He tested 1400 children each year from the age of 7 to the age of 11 with the Stanford-Binet test. He administered a non-standardized Welsh translation to children whose mother tongue was Welsh and who were attending an English school. When he compared the rural bilingual group to the rural or urban unilingual group, he found the unilingual group to be superior. However, when he compared the urban bilingual group to the urban unilingual group he found no difference between the two. Darcy reported that he concluded his study by stating that "children who become bilingual at an early age, by learning the second language during their play and in contact with other children, have an advantage over those who learn the second language at school" (Darcy, 1953, p.25). His comments changed when he gave a vocabulary test to the same subjects and found that the mean range of vocabulary of unilinguals was higher than that of the bilinguals in both English and Welsh. He explained this finding by stating that "mental confusion" occurs in the bilingual children more often than

in the unilinguals. These findings do not support Whorf's school of thought, since Saer found no difference between the two groups with regard to cognitive ability. In addition, methodological problems - such as the use of a translation in Welsh of the Stanford-Binet that had not been standardized in that language - prevent us from arriving at a clear conclusion on the problem.

Pintner (1932) compared the results of 430 children from three elementary schools using the Pintner-Cunningham Primary Mental Test and the Pintner Primary Non-Language Test. In order to classify children as bilingual or unilingual, only their surnames or the judgements of their teachers were used. This way of classifying unilinguals and bilinguals is likely inaccurate. Pintner obtained disparate results: in one school the bilingual group scored higher than the unilingual, in the second one the unilingual group scored higher than the bilingual group and in the third school he found no difference. In this experiment, it is important to note that subjects were not matched on socioeconomic status and that the level of bilingualism was not controlled at all.

Among the 33 studies she reviews, Darcy recorded only two which found bilingualism to be associated with a higher level of intelligence. In those studies, she argues, variables such as socioeconomic status, sex and degree of bilingualism were not adequately controlled. Eleven studies found no difference between unilinguals and bilinguals on intelligence.

Darcy argued that in most cases verbal I.Q. measures in a second language were not adequate measures of the I.Q. of bilinguals, since a) bilinguals most often scored higher on nonverbal measures than on verbal measures, and b) that differences between those two measures were larger for bilingual than for unilingual groups. Darcy argues that bilinguals suffer from a "language handicap" when tested with verbal intelligence tests. As we report later, Nielsen (1983) found that bilingual children, defined as French Immersion students, are not disadvantaged by intelligence tests administered in their first language.

From these earlier studies, it is difficult to say whether there is a positive or negative correlation between bilingualism and cognition level because of the methodological problems described earlier in these studies.

In trying to answer questions of this kind, researchers are faced with another problem: they have to deal with human subjects. From a conceptual point of view, a true experiment could not be performed to establish a causal relationship between bilingualism and cognitive ability. For even if such an experiment were possible, innumerable confounding variables would prevent researchers from arriving at a clear conclusion. As MacNab (1979) points out, "most studies are correlational and give no indication of the cause of associations found between language skills and cognitive ability" (1979, p. 250).

The study of Peal and Lambert (1962) were the first to give empirical evidence that bilingualism might be positively correlated with cognitive development. In Montreal Peal and Lambert (1962) selected 364 subjects from French schools of the Catholic School Board. In order to distinguish unilinguals from bilinguals, they administered 4 tests to the subjects (Word Association Test, Word Detection Test, Peabody Picture Vocabulary Test, Subjective Self-Rating Score). From these, the authors obtained a group of 89 "balanced bilinguals" and 75 monolinguals. Subjects who did not obtain the necessary score to be classified in one category or the other were eliminated. A "balanced bilingual" was defined as a person who is equally skilled in French and English. However, as Cummins (1976) remarked, referring to Peal and Lambert study (1962), "the criterion for balance has been quite lenient". A child could be clearly more proficient in English and still be classified as balanced. Macnamara (1966) criticized Peal and Lambert's method of selecting bilinguals, arguing that the use of linguistic measures for so doing invalidates any linguistic comparison between bilingual and unilingual groups. After the classification of the subjects as bilinguals or unilinguals, Peal and Lambert (1962) tested the two groups on intelligence, attitude and achievement. Three measures of intelligence were used: a verbal measure, the Lavoie-Laurendeau Group Test of General Intelligence in French, a nonverbal measure: the Raven Progressive Matrices Test and the subtests from the Thurstone Primary Mental Abilities translated in French which drew least directly on verbal ability (1962, p. 9). For the statistical

analysis, the subjects were matched on sex, age, and socioeconomic status.

Results showed that the bilingual group performed better on nonverbal subtests involving concept-formation and mental reorganization, called by the authors "mental flexibility," while the two groups performed similarly on nonverbal tests requiring only perceptual speed. Peal and Lambert offered an hypothesis for explaining this result. They said that compound bilinguals (those who learned both languages at the same time) develop more flexibility in thinking because they have constantly to switch from one code to another, and to make use of two different perspectives. The bilinguals appeared to perform better on verbal tests. However, when the two groups were matched on nonverbal ability as well as on age, sex, and socioeconomic status, there was no significant difference between them on verbal intelligence. The authors concluded that bilinguals appear to have a more "diversified set" of mental abilities than the monolinguals. It is unclear, however to what aspects or criteria of "diversity" the authors appealed in making this judgment.

This study brought new life to research on intelligence and bilingualism. It gave insight into the importance and the role of variables not previously considered in this field of research. The most recent studies in the field, are preceded by findings on the following variables: bilingualism, socioeconomic status, social context, fairness of the tests used, and the number of books at home.

Bilingualism

Bilingualism was the least well defined variable in the earlier studies. In some studies, the only necessary condition of classification as a bilingual was simply to have been in contact with two languages in the home (Tsushima and Hogan, 1975), or even to have a surname which sounded other than English (Pintner, 1937). At the other end of the scale, are somewhat more precise operational definitions of bilingualism, as in Peal and Lambert's study (1962).

While nonrigorous definitions of bilingualism such as the former cannot be expected to produce clear results, precise measures of linguistic balance may introduce a bias into the comparison of the bilingual and unilingual groups. In the Peal and Lambert study 200 subjects out of 364 were rejected on the basis that they could not unambiguously be classified as either monolingual or bilingual. It is then possible that selected balanced bilinguals had different cognitive abilities than the rejected bilinguals.

In order to find out if the method of selecting bilinguals used by Peal and Lambert (1962) invalidates linguistic comparisons between bilinguals and unilinguals , Cummins (1975) examined the verbal and nonverbal ability of balanced and non-balanced bilinguals in one of his studies and found no difference between the two groups on these abilities. On the other hand, there were significant differences on measures of divergent thinking between balanced and non-balanced bilinguals. He concluded by saying that the criterion

of balance in recent studies including the Peal and Lambert (1962) study had been lenient and that "a high level of language learning aptitude or intelligence is unlikely to have been necessary to attain the criterion of balance" and that "there is no evidence that the procedures used to select the bilingual samples in subsequent studies biased the comparison between bilingual and monolingual samples." (1976, p.17). The biais refers to results on cognitive ability tests.

Bilinguals can be balanced and still attain a low level of expertise in both of their languages. Cummins (1976) suggested that bilinguals must attain an adequate level in their first language as well as in their second language in order to avoid cognitive disadvantage. Some studies (Skutnabb-Kangas and Toukomaa, 1976) have used balanced bilinguals who still obtained lower results than unilinguals in either of their languages when achievement tests were administered in both of their languages. Cummins (1976 and 1979) proposed two hypotheses to explain lower results of bilinguals in intelligence tests in studies where bilinguals were approximately as competent in both languages: developmental interdependence, and minimum threshold level of competence in second language. The interdependence theory suggests that the level of competence in a second language is a function of competence in the first language. The threshold hypothesis states that there may be a first threshold level of linguistic competence which bilingual children must attain in their second language in order to avoid cognitive disadvantage. Then, they must attain a second threshold of linguistic competence

in their second language in order to benefit from the advantage of being bilingual from the point of view of their cognitive functioning. Concerning the threshold hypothesis, Cummins made the inference that studies which obtained lower results for bilinguals than for unilinguals on cognitive ability, involved bilinguals who had not attained the first threshold level in their second language. Studies finding no difference between bilinguals and unilinguals on cognitive ability involved bilinguals who had not attained the second threshold level.

This work on definitions of bilingualism, and Cummins' hypotheses, suggest the importance, in a study such as the present one, of a sample of bilinguals who have at least the following characteristics: a) are approximately as skilled in both of their languages, b) are competent in their first language before coming to school, and c) have attained an adequate level of competence in their second language so that they are not being disadvantaged by cognitive ability tests. Such a population invites investigation of a correlation between bilingualism and cognitive ability.

The present study was arranged so as to involve a random (as opposed to a selected sample of "balanced bilinguals") sample of bilingual subjects from French Immersion schools which would be as comparable as possible to the monolingual sample.The sample was chosen at the end of grade 3 on the assumption that at that point in their education, their English and French language levels would be approximately equivalent. In fact, starting in grade 4, early French Immersion children spend 50% of their time being taught in French

and 50% of their time being taught in English. This teaching structure presumes that from grade 3 on, students are equally skilled in English and in French. However, this has not been demonstrated by any study.

Although Cummins' theory has not been yet supported by empirical research and although the threshold levels are not operationally defined, Cummins' hypotheses bear some implications for the choice of bilingual subjects in the present study. In this study the bilingual subjects came from unilingual English homes and attended French Immersion school from Kindergarten to grade 3. In the Vancouver School Board (V.S.B.), the percentage of time spent in French in early immersion classes is 100% from Kindergarten to the end of grade 2. In grade 3, 80% of time is spent in French while the remaining 20% is spent on English language arts. It is assumed that at the end of grade 3, children are proficient enough in their second language to encounter no cognitive disadvantage of the kind described by Cummins.

Socioeconomic Status

A body of evidence now shows how essential it is to control for the variable of socioeconomic background in any study of cognitive ability. The importance of the socioeconomic background on student aspirations, cognitive development, as well as achievement, has been well documented in the United-States (Coleman, 1966; Blau

and Duncan, 1967) as well as in Canada (Harvey, 1974; Laxer et al, 1974).

Burns (1982) conducted research in Northern Ontario in order to study French Immersion parents' socioeconomic backgrounds - with regard to income, occupation, and education - as compared to the general population. He found that French Immersion children represent an elite cohort; their parents have a higher income and a more extensive formal education than the general population. Furthermore, professionals are heavily represented among French Immersion parents. Although the conclusions of Burns' study are limited to the Northeastern Region of Ontario, Burn reported that the Toronto School Board has conducted the same kind of study, and has found that French Immersion pupils were from families of a higher socioeconomic status than pupils in regular English programs.

Considering the results of the studies on socioeconomic background, it appears important to match bilinguals with unilinguals on the socioeconomic status of their parents, or to make statistical adjustments in the data analyses. The socioeconomic variable should refer to levels of formal education as well as income or occupation.

Social Context

Here, "social context" has to do with the way the second language is viewed by society. Lambert (1975) made the difference

between "additive" and "subtractive" bilingualism. In additive bilingualism, the bilingual is adding a socially relevant language to his first language at no cost to his first language competence. This is the case in Canadian French Immersion schools. In subtractive bilingualism, the bilingual is replacing his first language by his second one, more socially accepted. Lambert (1977) associated the studies in which bilinguals scored lower than unilinguals on cognitive ability tests with a social context of subtractive bilingualism. He associated the majority of studies in which bilinguals scored higher than unilinguals on cognitive ability tests with a social context of additive bilingualism.

The social context of Canadian French Immersion schools - which leads to an additive bilingualism - should not be the source of disadvantage for bilinguals in a study of bilingualism and cognitive ability.

Fairness of the Tests Used

Earlier studies often reported that bilinguals were tested in their second language (Darcy, 1953). Moreover, they were mostly from a culture other than the one in which they were living. Darcy pointed out that some bilinguals suffered from a language handicap.

On the basis of these experiences, one might ask whether it is fairer to test French Immersion children in their first language, English, or in their second language, French.

Nielsen (1983) asked this question regarding the use of the WISC-R test among grade 2 French Immersion students in B.C. She found that "use of the WISC-R does not appear to be disadvantageous for use with higher SES French immersion children whose main language of the home is English." She matched her two groups on the socioeconomic status of the father; however her subjects were all from a higher socioeconomic class. She indicated the limitations of her study's results by stating that her two groups of children were "quasi-matched." By this she meant that there were possible initial differences other than socioeconomic status of the father between the French Immersion and the regular program group. The results of her study do not give any indication on how French Immersion children would perform if they had been continuously in a regular program. It only shows that they do not score lower than Non-French Immersion children when administered the WISC-R test in English. Once again in this most recent study, the limited range of socioeconomic background of the subjects the investigator had to choose from, prevents a clear conclusion on the effect of bilingualism upon cognitive functioning.

D'Anglejan et al (1978) studied the ability of groups of learners of English as a second language to solve problems of deductive reasoning in their native and second language. They found that their subjects took more time to process the information in their second language than in their first language whatever their proficiency level was in that second language. It is possible that,

even though those subjects were bilinguals, their first language was still dominant.

On this subject, Cummins (1976) indicated that "balance" does not imply complete equilinguality and that the quantitative procedure used to select "balanced bilinguals" was designed only to eliminate those who were more dominant in one of their two languages. This statement implies that children in French Immersion remain dominant in their first language. This has also been my observation in French Immersion schools.

This research suggests that it might be more fair to test grade 3 French Immersion children in their first language, English, especially if the tests were oral tests.

Number of Books at Home

Furthermore, the summary report of the British Columbia Reading Assessment (Jeroski, 1984, p.41) reported that the number of books in the homes of the students was an effective predictor of student achievement in reading. A significant difference between the two groups on this variable would influence the results of students on the tests used in this study. Therefore, this information was collected and analysed.

The Latest Studies

Since the Peal and Lambert experiment (1962), research on bilingualism and cognitive ability may be summarized under three heads:
general cognitive ability, linguistic flexibility, and divergent or creative thinking. No researchers have studied the relationship between bilingualism and verbal and non-verbal reasoning ability.

Because of its longitudinal aspect the study conducted by Barik and Swain (1976) may eventually show a causal relationship between bilingualism and cognitive ability. However, some inevitable confounding variables such as social milieu have prevented the authors from making such a link (MacNab, 1979). Barik and Swain collected IQ data (using the Otis-Lenon Mental Ability Test) over a 5 year period with pupils from Kindergarten through grade 4 in a French Immersion school. Considering Grade 1 to 3 only, the Immersion group scored significantly higher than the comparison group over the three-year period on both classification and analogies but not on the following of verbal directions where there was no significant difference between the two groups. Because there was an initial difference between the two groups' IQ, the authors adjusted the scores for initial IQ and found no significant increase or decrease in the IQ of the two groups over the three-year period. Finally, the authors noted that the study did not involve balanced bilinguals already bilingually competent, but rather pupils moving toward bilingualism. For this reason, they asked the question

whether those pupils had attained the Cummins' threshold level (1976). According to Cummins' hypothesis, if they had, they would avoid cognitive disadvantage, or might potentially, score higher on cognitive functioning tests. To answer the question whether those pupils had attained the Cummins' threshold level (1976), Barik and Swain (1976) tested the French Immersion pupils on their proficiency in French using the Test de rendement en francais. From the results of this test, they formed two new groups: high French achievers and low French achievers. They then compared the increase of both groups' individuals' IQ over a three-year period (K-3). The scores of the low French achievers remained unchanged while the score of the high French achievers increased significantly from Kindergarten to grade 3. This study gives some support to Cummins' threshold hypothesis. However, nothing here proved that low French achievers failed to reach the threshold and there is no evidence that they have experienced cognitive disadvantage as their IQ remained unchanged from Kindergarten to grade 3.

Other authors have used Piagetian tasks instead of intelligence tests to investigate the relation between bilingualism and cognition. Liedke and Nielson (1968) observed better performance among bilinguals on Piagetian concept formation tasks after controlling for age, sex, socioeconomic status and intelligence. Bain (1975) reported that bilinguals at the preoperational stage showed an advantage on unilinguals at tasks requiring discovery of rules after being matched on socioeconomic status, IQ, and school grades . However, at the concrete operation stage there was no

significant difference between the bilingual group and the monolingual group.

Some of the recent studies were more oriented towards linguistic abilities of bilinguals and their metalinguistic awareness. Ianco-Worral (1972) studied compound bilinguals in South Africa and found that they attended more to the semantic than to the phonetic relation between words. Monolinguals interpreted more similarities between words more often according to their acoustic rather than their semantic characteristics. Because bilinguals have two codes for each object, they tend to conceptualize things and events in terms of their general properties rather than their linguistic symbols. Ben-Zeev (1977) studied middle-class Hebrew-English bilinguals and lower-class Spanish-English bilinguals and matched them for IQ with monolinguals. She found that bilinguals had less vocabulary and took longer to respond on a word association task, but showed more advanced processing of verbal material, and better discrimination in making perceptual distinctions.

Several studies have shown that bilinguals are better at analysing linguistic input than are monolinguals. Cummins (1978) gave a test to bilinguals and monolinguals in which participants had to use their deductive powers to determine if contradictory statements were true, false or neither. Bilinguals scored significantly better than monolinguals. Another study from Cummins and Mulcahy (1978) found that Ukrainian-English bilingual children were better able to analyse ambiguities in sentence structure than a group of anglophones matched for IQ, socioeconomic status and

school. The significance of these findings for the present study is that bilinguals seem to attend more to the logical aspect of sentences, and semantic aspect of words, and also to do better on problems involving various forms of deductive logic than unilinguals. In solving the word-finding problems in this study, subjects had to attend to the conceptual characteristics of the words to be found, and to use inductive and deductive logic. The results of Ianco-Worral (1972) and Cummins (1978) supported the hypothesis that bilinguals might obtain a higher score on the Children Word-Finding Test than unilinguals.

The Torrance Test of Creative Thinking has been used by various researchers to compare bilinguals and monolinguals (Landry, 1974; Carringer, 1974; Cummins, 1975; Torrance, Gowan, Wu, and Aliotti, 1970). Most of them found bilinguals to be superior compared to monolinguals on the flexibility, fluency, elaboration and originality scales. Torrance et al (1970) found that their sample of Chinese-English, whose level of bilingualism was not given, scored better than unilinguals on originality and elaboration scales, but lower than unilinguals on fluency and flexibility scales. This result might be explained by interference between the two languages, were the subjects not balanced bilinguals.

In summary, these various studies suggest that the experience of learning a second language is associated with differences in cognitive ability, linguistic perception, analytic power and verbal expression.

However, none of these studies compared how well bilinguals did, as compared to unilinguals on deductive and inductive reasoning. Only one group of researchers (D'Anglejan, Gagnon, Hafez, Tucker, Winsberg 1978) have done a study that examines the ability of groups of English-French bilinguals to solve problems of deductive reasoning (syllogism) in their native, English, and second language, French. Three groups of bilinguals from the Canadian Forces were formed in accordance to their level of proficiency in the second language. The results showed that the bilinguals succeeded in reasoning better in their first language than in their second language, and this independently of their level of bilingualism. They took a longer time to process information and solve problems in their second language, all the while making more mistakes. It is difficult to generalize these results to French Immersion students since the learning of the second language was not in an immersion situation, and as a result, the level of bilingualism may have been different. Furthermore, the study of D'Anglejan et al focused on differences in problem-solving in first and second language and not on the possible differences in problem solving ability of bilinguals compared to monolinguals.

***

In conclusion, the results of the research were still inconclusive regarding the hypothesis that bilinguals may have higher cognitive ability than monolinguals. From a linguistic point of view, it has been demonstrated that bilinguals attend more to the

semantic than the phonetic characteristics of the words, and that they score generally higher on the fluency and flexibility scales of the Torrance Test of Creative Thinking as well as on the imagination and originality scales. Some studies seem to indicate that bilinguals have better deductive reasoning.

The research findings are valuable from a methodological point of view. Researchers have been able to demonstrate that some variables, such as bilingualism, socioeconomic status, social context, language of the test administered, and number of books owned by the family, play an important role in assessing the relationship between bilingualism and cognitive ability. They have also been able to show how such variables could be controlled in order to obtain less ambiguous results. However, the human context in which these studies are done prevent us from controlling every possible confounding variable. Subjects taken from French Immersion schools represent a self-selected group. The decision of parents to send their children to French Immersion school or regular English program depends on unknown variables which might also influence the results of the studies.

In the present study, important variables such as socioeconomic status, level of bilingualism, number of books at home and so on have been controlled to investigate the possible differences between bilinguals and unilinguals on measures of nonverbal and verbal reasoning ability.

Our reason for studying reasoning ability rather than cognitive ability, as measured by general IQ tests, is that reasoning ability as defined here is a more specific ability than the broad concept of general cognitive ability. Although reasoning ability is part of cognitive ability, it is a narrower concept than that of cognitive ability. Moreover, the literature indicates that a relationship may exist between reasoning ability and the acquisition of a second language.

Research Questions

In the present study, the question was posed whether there was a difference in verbal and nonverbal reasoning ability between Grade 3 French Immersion students and Grade 3 Non-French Immersion students.

A nonverbal reasoning test involving perceptual problem solving tasks: the Test Of Nonverbal Intelligence (TONI) was used to measure nonverbal reasoning ability.

A verbal reasoning test involving vocabulary problem solving tasks (the Children's Word Finding Test) was used to measure verbal reasoning ability.

The concurrent validity of the TONI with the Wechsler Intelligence Scale for Children - Revised (WISC-R) and with the Children Word Finding Test (CWFT), was investigated and finally, the equivalence of the two forms of the TONI was verified.

CHAPTER III

METHODOLOGY

The purpose of the present study was to compare levels of verbal and non-verbal reasoning ability between Grade 3 bilingual children and unilingual children matched on socioeconomic background.

The Wechsler Intelligence Scale for Children - Revised (WISC-R), the Test of Nonverbal Intelligence (TONI) Form A and B and the Children's Word Finding Test (CWFT) were administered to two groups of students (French Immersion and Non-French Immersion) drawn from a total of six schools.

An analysis of variance tested for any differences between the two groups (French Immersion and Non-French Immersion) on socioeconomic status and on the number of books owned by the parents.

Another analysis of variance compared the two groups on the WISC-R, the TONI and the CWFT. In addition, correlations among the three tests were then calculated to assess the concurrent validity of the TONI with the WISC-R and with the CWFT and the equivalence of the two forms of the TONI.

Finally, an analysis of covariance was conducted, using the WISC-R test as a covariate, to analyse the difference between the two groups on the TONI and the Children's Word Finding Test when the level of intelligence was kept constant.

A complete description of the methodology is presented in this Chapter. The design and the sampling procedure are described. Similarly the tests, the questionnaire used and the data analyses are presented.

The Design

The study was designed to compare levels of verbal and non-verbal reasoning ability of bilingual children and unilingual children at the end of Grade 3 matched for socioeconomic level.

Bilingual children were defined as students registered in early French Immersion since Kindergarten and coming from anglophone families where the only language used in the home was English.

Unilingual children were students registered in a regular English program since Kindergarten and coming from anglophone families where the only language used in the home was English.

Verbal and nonverbal reasoning ability were measured according to results obtained by the students on a nonverbal reasoning ability test (the Test of Nonverbal Intelligence), and on a verbal reasoning test (the Children's Word Finding Test).

There was one independent variable, the condition bilingual or unilingual. There were two dependent variables, the level of verbal reasoning and the level of nonverbal reasoning.

Furthermore, the use of the TONI enabled an analysis of its concurrent validity with the WISC-R and the Children's Word Finding Test and, to verification of the equivalence between the two forms of the TONI, administrated in counter-balanced order.

The Sample

The sample for this study was selected from Grade 3 classrooms at six schools from the West and Centre areas of the Vancouver School Board. Three schools were French Immersion single track schools and three were regular English program schools.

All the schools were chosen in the West and Centre areas of the Vancouver School Board on the assumption that these were homogeneous areas regarding the general social and economic standing of the residents. This was important, as the literature on comparisons of socioeconomic backgrounds between unilinguals and bilinguals showed that French Immersion students came from families of higher socioeconomic status than Non-French Immersion students (Burns, 1983). Each French Immersion school was yoked with a regular English program school drawing students from approximately the same area (see map in Appendix A).

A total of 60 Grade 3 students, 10 in each of the six schools was chosen randomly. The French Immersion sample represented 14.85% of the total of the students registered in French Immersion at the Vancouver School Board in 1985-86. The Non-French Immersion sample represented

0.86% of the total number of students registered in Non-French Immersion schools at the Vancouver School Board in 1985-86.

Questionnaire

Letters informing parents of the purpose and nature of the study were sent home to every Grade 3 student enrolled in the schools selected. The letters were accompanied by forms requesting parental consent, and by a questionnaire to obtain background information. Copies of the letter, parental consent form and questionnaire are provided in Appendix A.

The questionnaire included questions on parents' education and occupation. The Blishen scale (1976) was used to determine the socioeconomic index of each family, from the information on occupation. The index scores of the Blishen scale are T scores with a mean of 50 and a standard deviation of 10. In the case of a family where both parents were working, the higher index was used. Data on parents' education were also used as a means of cross-checking the information on occupation. In some cases a more precise description of the profession of the parents was required in order to assure the correspondence with the Blishen scale.

Furthermore, the investigator phoned the parents to ask them how many books they owned. They could choose among 6 categories:

1. 0 to 9 books;,
2. 10 to 24 books;
3. 25 to 99 books;
4. 100 to 249 books;
5. 250 to 499 books;
6. 500 or more.

Information obtained on the questionnaire provided a basis for selecting the final sample for this study. The selection criteria were:

(1) Parental consent had been obtained for the child's participation in the study.

(2) Enrolment in the current language of instruction had been continuous from Kindergarten through to the time of testing;

(3) English was the only language spoken at home;

(4) Students were not attending Learning Assistance Centers; and

(5) Students had not been administered the tests used in the study during the last year.

Sampling

Of the 254 letters sent to the parents, 187 were returned. Table 3-1 shows how many letters were returned in each school. Slightly more French Immersion parents (76%) returned the letters than Non-French Immersion parents (71%).

Table 3-1

Number of Letters Returned in Each School

School	Number of Letters Sent	Number of Letters Returned	Percent Returned
Immersion			
1	59	41	69
2	24	21	88
3	48	38	79
Total Immersion	131	100	76
Non Immersion			
1	50	37	74
2	47	27	57
3	26	23	88
Total Non Immersion	123	87	71
Total	254	187	74

Among the 187 students whose parents returned the letters, 84 were eliminated on the basis of one or more of the criteria described in the preceding paragraph. Table 3-2 describes how many students were eliminated on the basis of each criterion in each school. 27% of the students used a language other than English at home, both in the French Immersion group and in the Non-French Immersion group. Of the French Immersion parents 6% did not give their consent compared to 18% in the Non-French Immersion group.

Finally, students were randomly chosen in each school from the group left after successive elimination. Table 3-3 recapitulates the different steps of elimination in arriving at the final sample of 10 in each school.

The original intent of the study was to select equal numbers of boys and girls, keeping in mind sampling requirements. Errors in interpreting the gender of two subjects arose from ambiguity in the interpretation in their Christian names. Therefore, there were 15 boys and 15 girls in the French Immersion sample and 13 boys and 17 girls in the Non-French Immersion sample.

Description of the Tests Used

Wechsler Intelligence Scale for Children-Revised (WISC-R)

The Wechsler Intelligence Scale for Children - Revised is a general intelligence test first used in 1949. The entire test can be administered in 75 minutes. It covers an age range from 6-0 to 16-11

Table 3-2

Elimination of Students by Criterion

School	Number of Letters Returned	Number of Students Eliminated on the basis of: (1) Consent not Given	(2) Lang. of Instrct.	(3) Lang. spoken at home	(4) Learning Assit.	(5) Prev. Tested
Immersion						
1	41	2	6	13	4	0
2	21	2	0	2	0	2
3	38	2	0	13	4	0
Non Immersion						
1	37	5	1	7	2	0
2	27	6	4	6	4	0
3	23	5	0	11	1	0
Total	187	22	11	52	15	2

Table 3-3

Recapitulative Table of the Formation of the Sample

School	Number of Letters Sent	Number of Letters Returned	Number of Students After Elimination Based on Criterion (1) to (5)	Final Sample Size
Immersion				
1	59	41	23	10
2	24	21	17	10
3	48	38	19	10
Non Immersion				
1	50	37	24	10
2	47	27	10	10
3	26	23	10	10
Total	254	187	103	60

years. It consists of 12 subtests. Six of these measure verbal skills, while the other six measure performance skills. The performance subtests provide a nonverbal measurement of cognitive ability, and refer to visual-motor perception.

A brief description of what each subtest measures follows (Sattler, 1982, pp. 188-189):

INFORMATION - The wealth of available information acquired as a result of native ability and early cultural experience.

SIMILARITIES - Verbal concept formation.

ARITHMETIC - Numerical reasoning ability.

VOCABULARY - Variety of functions, including language development, learning ability, and fund of information.

COMPREHENSION - Social judgement: the ability to use facts in a pertinent, meaningful, and emotionally appropriate manner.

DIGIT SPAN - Short-term memory and attention.

PICTURE COMPLETION - Ability to differentiate essential from nonessential details.

PICTURE ARRANGEMENT - Nonverbal reasoning ability and planning ability.

BLOCK DESIGN - Visual-motor coordination and perceptual organization.

OBJECT ASSEMBLY - Perceptual organization ability.

CODING - Visual motor coordination, speed of mental operation, and short-term memory.

MAZES - Planning ability and perceptual organization.

Scaled scores from 10 subtests are combined to yield three summary IQ scores, each with a mean of 100 and a standard deviation of 15: the Verbal IQ, the Performance IQ, and the Full scale IQ. The remaining two subtests, Digit Span and Mazes, are supplementary tests used to replace invalidated results on regular subtests and to gain further information. Scaled scores for individual subtests have a mean of 10 and a standard deviation of 3.

The reliabilities of the Verbal, Performance and Full scales are reported in the manual as high (averages of .94, .90, and .96, respectively) with a standard error of measurement for the Full scale of about three IQ points. The WISC-R shows adequate concurrent validity with other measures of intelligence (Salvia & Ysseldyke, 1981; Sattler, 1982).

Test of Nonverbal Intelligence (TONI)

The Test of Nonverbal Intelligence is "a language-free measure of cognitive ability." It covers the age range from 5-0 to 85-11. It was created in 1972 to answer the need for a measure to evaluate the

intellectual potential of specific groups of individuals such as the aphasic, the retarded, the learning disabled, the deaf, non-English speakers and the culturally "different".

All items chosen require problem solving ability, defined as the ability to identify one or more rules for relationships among abstract figures. The rules include simple matching, analogies, classification, intersections, and progressions. The subject selects the correct figure among either four or six alternatives to complete a set of figures in which one or more of the figures are missing.

The test is a language-free measure. It may be administered entirely without the use of oral or written language, by using gestures. The subject has only to point out the figure he or she has chosen among the four or six possibilities. There are two equivalent forms of the test containing 50 items each. The entire test can be administered in 15 to 30 minutes.

The test has a mean of 100 and a standard deviation of 15. Its internal consistency and its stability reliability for special populations is .80 or higher. Observed correlations between the TONI and the WISC-R, ranges from .46 to .95 for the Learning Disabled and the Educable Mentally Retarded. The correlation for Normal groups has been established only with the Otis-Lennon Mental Ability test for age 8-1, and is .86 and .81 for Forms A and B respectively.

Equivalence of the two forms of the TONI has been calculated by the authors of the test and reported in the technical manual. The

correlation coefficients between the two forms all exceed r = .80, except for the age range 8-6 to 10-11 where it is .78.

Children's Word Finding Test

This test was formulated in 1976, on the model of Reitan's Word Finding Test (1972), designed for adults. In this test subjects must discern the meaning of a nonsense word through appreciation of its verbal context. The test serves as the verbal counterpart of nonverbal problem solving tasks most often used in neuropsychological batteries.

There are 13 items in this test. Each item is composed of five sentences in which "grobnick" is to be replaced by a word which is the same for all five sentences within the item. The sentences are tape-recorded with an interval of 5 seconds between each sentence. The entire test takes approximately 15 minutes to administer. The students receive one point for each correct answer.

For example, the first item is as follows:

1- Every classroom has at least one grobnick.

2- Grobnicks usually hang on the wall.

3- The teacher writes on the grobnick.

4- Grobnicks can be green or black.

5- You write on a grobnick with chalk.

Preliminary norms have been established for children from 6 to 10 years old. The test has been administered to 20 children in each age group (6-7-8-9-10). Means and standard deviations for number of

correct responses were calculated for normal children in 1982. For the age group 9, the mean was 36.80 and the standard deviation 7.61.

Testing

All tests were administered to each of the 60 children, according to the procedures described in the test manuals. All testing was done by 5 Level C testers (Cronbach, 1970) who had been trained and supervised in the administration and scoring of the WISC-R, TONI and Children's Word Finding Test. Testing took place in the children's schools during the regular instructional day in the period between May 26 and June 13, 1986. Each administration required approximately 2 hours. The order of the WISC-R and the TONI was counter-balanced. Half of the students were administered the TONI first and half were administered the WISC-R first. Furthermore, the two forms of the TONI were administered successively and in a randomized order. The CWFT was always administered last. Schools were contacted in advance to arrange time and appropriate space for test administration. A student code number, language of instruction, birthdate, gender and socioeconomic index of the parents was recorded on each test protocol. No names appeared on the protocols. This procedure allowed necessary information for the analysis while guaranteeing the anonymity of each subject.

Scoring and Data Preparation

Completed protocols were scored by the respective testers, following directions in the test manuals. All of the protocols were then checked for errors by a second party. Finally, scores were coded and entered into computer disk files for analysis.

Data Analysis

An analysis of variance was performed to test for any difference between the two groups (French Immersion and Non-French Immersion) on socioeconomic status and number of books owned by the parents.

A nested analysis of variance was then performed to detect possible differences among the types of school (French Immersion and Non-French Immersion) and within the schools of a type on the results of each of the tests administered. A mixed model was used: the type of school provided the fixed effects factor, and the schools within type the random-effects factor. The third level was the students themselves. The design is given in Table 3-4.

In this analysis, the independent variables were:

a) the type of school (fi or nfi)

b) the school within type (1 - 2 - 3)

and the dependent variables were the scores on each of the tests administered.

Table 3-4
Nested Analysis of Variance
Experimental Design Matrix

	School	Tests Administered					
IMMERSION	1	WVS.f1	WPS.f1	WFS.f1	TA.f1	TB.f1	F.f1
	2	WVS.f2	WPS.f2	WFS.f2	TA.f2	TB.f2	F.f2
	3	WVS.f3	WPS.f3	WFS.f3	TA.f3	TB.f3	F.f3
	Total	WVS.f	WPS.f	WFS.f	TA.f	TB.f	F.f
NON IMMERSION	1	WVS.nf1	WPS.nf1	WFS.nf1	TA.nf1	TB.nf1	F.nf1
	2	WVS.nf2	WPS.nf2	WFS.nf2	TA.nf2	TB.nf2	F.nf2
	3	WVS.nf3	WPS.nf3	WFS.nf3	TA.nf3	TB.nf3	F.nf3
	Total	WVS.nf	WPS.nf	WFS.nf	TA.nf	TB.nf	F.nf

Note:
WVS Wechsler Scale of Intelligence for Children - Revised, Verbal Scale
WPS Wechsler Scale of Intelligence for Children - Revised, Performance Scale
WFS Wechsler Scale of Intelligence for Children - Revised, Full Scale
TA Test of Non Verbal Intelligence, Form A (counter-balanced order with TONI B)
TB Test of Non Verbal Intelligence, Form B (counter-balanced order with TONI A)
F Children's Word Finding Test
f French Immersion
nf Non-French Immersion

The necessary computations were performed using the computer program BMDP P8V, General Mixed Model Analysis of Variance - Equal Cell Sizes (Jenrich and Sampson, 1985)

An analysis of covariance was conducted for the TONI Form A and Form B, and for the Children Word Finding Test (CWFT), with the WISC-R test Full Scale as a covariate. This analysis was performed to test for any differences between the two types of schools, when the level of cognitive ability was controlled.

The necessary computations were performed using the computer program SPSS-X (Nie & all, 1975), Analysis of Covariance.

Finally a multiple Pearson r correlation was calculated for all the tests, successively with each type and with the combined types. Special attention was paid to the correlation between the TONI and the other tests for the purpose of concurrent validity, and between the two equivalent forms of the TONI.

The necessary computations were performed using the computer program SPSS-X (Nie & all, 1975), Multiple Pearson r Correlation.

All computations were performed on a AMDAHL 470/V8 computer maintained by the Computing Centre, University of British Columbia.

CHAPTER IV
RESULTS

This Chapter presents the descriptive statistics of the study and contains the results of the statistical analysis outlined in the previous Chapter.

Descriptive Statistics

Care was taken to ensure that the analyses relected a comparison between the levels of language learning: French Immersion students and Non French Immersion. In particular, the socioeconomic status of the student's families and the number of books owned at home were identified as two potential confounding variables that could influence the results of the tests administered.

Indices of the socioeconomic status and of the number of books owned by the student's families were obtained from questionnaire responses (Table 4-1). Analysis of variance on these two indexes indicated no significant difference between the two types of schools (Table 4-2). The families of both French Immersion students and Non-French Immersion had on average between 250 to 499 books (index level 5). The number of books owned by the families of the students in the sample of this study is higher than the provincial average for Grade 4 reported in the British Columbia Reading Assessment (1984). None of the students in this study were at index level 2 or 1 (less than 24 books) compared to 9% in the provincial average. Only 7% of

Table 4-1

Indeces of Home Book Numbers and SES
by Type of School

	French Immersion		Non French Immersion		Total Sample	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
number of books index	5.03	0.93	5.10	0.89	5.07	0.90
socioeconomic status index	64.50	11.81	65.17	10.06	64.83	10.88

Note: All information was obtained from questionnaires.
The book index used was (1) 0-9 books, (2) 10-24, (3) 25-99, (4) 100-249, (5) 250-499, (6) 500 books or more.

Table 4-2

Analysis of Variance for Number of Books and SES
in French Immersion and Non-French Immersion Groups

	df	SS	MS	F	p
Books/Between Type of School	1	0.07	0.07	0.08	0.78
Books/Within Type of School	58	47.67	0.82	----	----
SES/Between Type of School	1	6.67	6.67	0.06	0.82
SES/Within Type of School	58	6973.67	120.24	----	----

the students in the entire sample were at the index 3 compared to 20% in the provincial average. A high percentage of the sample in this study (41.67%) came from families who owned more than 500 books (index level 6) compared to a provincial average of 25.4%. Both French Immersion and Non-French Immersion students' families had a socioeconomic status (index 65) that was higher than average (index 50).

The results of analyses of the WISC-R, the TONI, Form A and B and the CWFT are displayed in Tables 4-3, 4-4, 4-5, and 4-6.

Analysis of Variance

One of the assumptions of the analysis of variance is that the samples were taken from normal populations. This was tested by applying the Kolmogorov-Smirnov Goodness of Fit to the scores of all the WISC-R subtests, subscales and full scale, the TONI A and B, and the CWFT (Table 4-7). Expected frequencies from a normally distributed population were calculated for each test and compared to the scores obtained. All the distributions, except for the TONI.B, proved to be normal. The TONI.B distribution was leptokurtic and did not appear to seriously threaten the validity of analysies of variance.

French Immersion students scored higher than Non-French Immersion Students in all three sets of tests administered. In the WISC-R, TONI A, TONI B, and CWFT the French Immersion students scored 114, 102, 106, and 40 respectively; whereas the Non-French Immersion

Table 4-3

WISC-R Scale Means by Type of School

	French Immersion		Non French Immersion		Total Sample	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Information	10.80	3.10	9.77	2.37	10.28	2.79
Similarities	12.87	3.29	11.67	3.28	12.27	3.31
Arithmetic	11.17	2.45	12.03	2.68	11.60	2.59
Vocabulary	12.97	2.59	12.60	2.50	12.78	2.53
Comprehension	11.73	3.06	11.37	2.86	11.55	2.94
Digit Span	10.77	2.28	10.13	2.11	10.45	2.21
Picture Completion	12.27	2.24	10.43	2.81	11.35	2.69
Picture Arrangement	12.10	2.55	11.40	2.90	11.75	2.73
Block Design	13.13	3.08	12.00	2.41	12.57	2.80
Object Assembly	12.30	2.84	11.07	2.12	11.68	2.56
Coding	10.37	2.36	8.90	2.51	9.63	2.52
Verbal IQ	111.57	13.90	108.77	10.31	110.17	12.22
Perform. IQ	114.00	11.18	105.97	11.99	109.48	12.36
Full Scale IQ	114.20	12.21	107.67	9.87	110.93	11.49

Table 4-4

TONI Results by Type of School

	French Immersion		Non French Immersion		Total Sample	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
TONI A	102.87	14.00	96.07	10.10	99.47	12.58
TONI B	106.33	14.37	99.83	11.81	103.08	13.45

Table 4-5

Children's Word Finding Test Results by Type of School

	French Immersion		Non French Immersion		Total Sample	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Children's Word Finding Test	40.17	5.72	36.60	5.16	38.38	5.69

Note: The above results are given in raw scores. The maximum raw score on this test is 65. For the age group 9, the mean is 36.80 and the standard deviation is 7.61

Table 4-6

Range of Test Scores for the WISC-R, the TONI, and the CWFT by Type of School

	WISC-R Verbal I.Q.	WISC-R Perfor. I.Q.	WISC-R Full Scale	TONI.A	TONI.B	CWFT
Immersion	92-147 55	82-139 57	71-141 50	82-140 58	69-145 76	26-53 27
Non Immersion	85-128 43	80-128 48	89-124 35	82-126 44	79-120 41	27-46 19
All	85-147 62	80-139 59	89-141 52	82-140 58	69-145 76	26-53 27

Table 4-7

Kolmogorov-Smirnov Goodness of Fit Test

	Kolmogorov-Smirnov Z Value	2 Tailed P Value
WISC-R		
Information	0.92	0.36
Similarities	0.88	0.42
Arithmetic	0.96	0.31
Vocabulary	1.03	0.24
Comprehension	0.83	0.50
Digit Span	1.05	0.23
Picture Completion	1.15	0.14
Picture Arrangement	0.71	0.70
Block Design	1.06	0.21
Object Assembly	0.65	0.79
Coding	1.10	0.18
Verbal IQ	0.72	0.67
Performance IQ	0.66	0.77
Full Scale	0.55	0.93
TONI A	1.02	0.25
TONI B	1.56	0.02
Children's Word Finding Test	0.72	0.68

students scored 107, 96, 100, and 37 respectively. The differences were statistically significant on the Performance Scale ($p = .01$) and on the Full Scale ($p = .05$). The greatest difference was in the Performance Scale of the WISC-R where French Immersion scored 114 and the Non-French Immersion scored 105. French Immersion students scored slightly lower than Non-French Immersion students on only one of the WISC-R subtests, Arithmetic. The significance of these results was determined by an analysis of variance.

The results of the analysis of variance on the WISC-R are shown in Table 4-8. They indicate a difference between the French Immersion and Non-French Immersion group on the Performance Scale at $p < .01$ level of significance and on the Full Scale ($p < .05$). However, no significant difference ($p > .05$) was found between the two groups on the Verbal Scale of the WISC-R test.

The results of the TONI are contained in Table 4-9. A difference between French Immersion and Non-French Immersion ($p = .07$) was found on the TONI Form A, and no significant difference was found on the TONI Form B.

An analysis of variance of the Children's Word Finding Test found a difference ($p = .10$) between French Immersion and Non-French Immersion groups and is shown in Table 4-10.

In summary, the analysis of variance showed a significant difference between the French Immersion scores and the Non-French Immersion scores on the WISC-R Full Scale ($p < .05$) and WISC-R Performance Scale ($p < .01$). A difference was found on the TONI.A

Table 4-8

Analysis of Variance on the WISC-R

Source of Variance	Degrees Freedom	Sum of Squares	Mean Square	F	p
WISC-R FULL SCALE					
Between Type of School	1	640.27	640.27	16.39	0.02
Within Type of School	4	156.27	39.07	0.30	0.88
Within School	54	6993.20	129.50	----	----
WISC-R PERFORMANCE SCALE					
Between Type of School	1	1224.02	1224.02	28.29	0.01
Within Type of School	4	173.07	43.27	0.31	0.87
Within School	54	7619.90	141.11	----	----
WISC-R VERBAL SCALE					
Between Type of School	1	117.60	117.60	1.44	0.30
Within Type of School	5	327.73	81.93	0.53	0.72
Within Type of School	54	8359.00	154.70	----	----

Table 4-9

Analysis of Variance on the TONI

Source of Variance	Degrees Freedom	Sum of Squares	Mean Square	F	p
TONI A					
Between Type of School	1	693.60	693.60	5.87	0.07
Within Type of School	4	472.73	118.18	0.78	0.54
Within School	54	8168.60	151.27	----	----
TONI B					
Between Type of School	1	633.75	633.75	2.20	0.21
Within Type of School	4	1150.73	287.68	1.75	0.15
Within School	54	8886.10	164.56	----	----

($p = .07$), and on the CWFT ($p = .10$). However, there was no significant difference on the WISC-R Verbal Scale and on the TONI.B.

Analysis of Covariance

The central problem of this study was to test the difference in verbal and non-verbal reasoning ability between the students of the two types of schools. This was accomplished by using the TONI as a non-verbal reasoning test and the CWFT as a verbal reasoning test. However, it is possible that the differences between French Immersion students and Non-French Immersion students were due in part to differences in cognitive abilities. Using the WISC-R as a covariate to measure cognitive abilities, an analysis of covariance was done to test whether there was a difference between French Immersion Schools and Non-French Immersion schools on the TONI and the CWFT scores. The results of the analysis of covariance (Table 4-11) confirmed no significant difference on the TONI.A ($p = .09$) and on the CWFT ($p = .09$). There was no significant difference on the TONI.B as in the previous analysis after controlling for differences in cognitive abilities. The adjusted means were calculated for each test and were found to be higher for the French Immersion group than for the Non-French Immersion groups, although not significantly different (Table 4-12).

Table 4-10

Analysis of Variance on the Children's Word Finding Test

Source of Variance	Degrees Freedom	Sum of Squares	Mean Square	F	p
CWFT					
Between Type of School	1	190.82	190.82	4.45	0.10
Within Type of School	4	171.67	42.92	1.50	0.22
Within School	54	1547.70	28.66	----	----

Table 4-11

Analysis of Covariance of the TONI Form A and B, and of the CWFT with the WISC-R as covariate

Source of Variance	Degrees Freedom	Sum of Squares	Mean Square	F	p
TONI A					
Covariate	1	191.22	191.22	1.27	ns.
Between Type of School	1	449.56	449.56	2.99	0.09
Within Type of School	4	400.10	100.03	0.66	ns.
Within School	53	7977.38	150.52	----	----
TONI B					
Covariate	1	993.38	993.38	6.67	0.01
Between Type of School	1	224.03	224.03	1.50	ns.
Within Type of School	4	1204.88	301.22	2.02	ns.
Within School	53	7892.72	148.93	----	----
CWFT					
Covariate	1	251.16	251.16	10.27	0.01
Between Type of School	1	74.51	74.51	3.05	0.09
Within Type of School	5	154.52	38.63	1.58	ns.
Within Type of School	53	1296.54	24.46	----	----

Table 4-12

French Immersion Versus Non-French Immersion
Unadjusted and Adjusted Means for Level of Cognitive Ability
on TONI.A, TONI.B, and Children's Word Finding Test

Test	French Immersion Unadjusted Means	French Immersion Adjusted Means	Non-French Immersion Unadjusted Means	Non-French Immersion Adjusted Means
TONI A	102.87	102.33	96.07	96.61
TONI B	106.33	105.10	99.83	101.06
CWFT	40.17	39.55	36.60	37.22

Table 4-13

Pearson Correlation Coefficients
for the Entire Sample

	WISC-R Verbal I.Q.	WISC-R Perfrm. I.Q.	WISC-R Full Scale	TONI.A	TONI.B	CWFT
WISC-R Verbal IQ	----					
WISC-R Perfrm. IQ	0.34	----				
WISC-R Full Scale	0.84	0.79	----			
TONI.A	0.16	0.25	0.24	----		
TONI.B	0.28	0.31	0.35	0.52	----	
CWFT	0.40	0.31	0.45	0.36	0.38	----

TONI, Concurrent Validity and Equivalence of the Two Forms

The second purpose of this study was to verify the concurrent validity of the TONI with the WISC-R and CWFT. This was accomplished by calculating the correlation coefficients between each test for the entire sample and for each subgroup (Tables 4-13, 4-14, and 4-15). Attention should be drawn to two results in particular: (1) The correlations of the TONI Form A and B, with the WISC-R and with the CWFT are extremely low for the Non-French Immersion group. (2) The correlations of the TONI Form A and B appear to be higher with the Children's Word Finding Test than with the WISC-R; although, a t-test showed no statistical differences between the correlated correlation coefficients ($o = .05$, 2 tails, 57 df).

The third question dealt with the equivalence of the two forms of the TONI. A Pearson correlation coefficient ($r = .52$) was calculated between the TONI.A and the TONI.B for the entire sample of French Immersion and Non-French Immersion students combined (Table 4-12). This value is much lower than that reported in the technical manual ($r = .78$) for the same age group (8 years 6 months to 10 years 11 months). In fact, these two correlation coefficients are significantly different according to the Fisher's Z-transformation test ($Z = 0.57$, $df = 57$, $p < .01$).

The low correlation coefficient found between the two forms of the TONI test prompted a further analysis on the effect of treatment order. A procedure outlined by Gart (1969) was used to test for the difference in response between the two sequentially applied

Table 4-14

Pearson Correlation Coefficients
for the French Immersion Group

	WISC-R Verbal I.Q.	WISC-R Perfrm. I.Q.	WISC-R Full Scale	TONI.A	TONI.B	CWFT
WISC-R Verbal IQ	----					
WISC-R Perfrm. IQ	0.42	----				
WISC-R Full Scale	0.90	0.78	----			
TONI.A	0.17	0.26	0.24	----		
TONI.B	0.41	0.46	0.50	0.60	----	
CWFT	0.47	0.40	0.53	0.46	0.53	----

Table 4-15

Pearson Correlation Coefficients
for the Non Immersion Group

	WISC-R Verbal I.Q.	WISC-R Perfrm. I.Q.	WISC-R Full Scale	TONI.A	TONI.B	CWFT
WISC-R Verbal IQ	----					
WISC-R Perfrm. IQ	0.20	----				
WISC-R Full Scale	0.78	0.77	----			
TONI.A	0.06	0.05	0.07	----		
TONI.B	0.02	0.02	0.02	0.29	----	
CWFT	0.28	0.06	0.22	0.07	0.05	----

Table 4-16

Pearson Correlation Coefficients
for Each Type of School Between TONI A and TONI B
According to the Order of Administration of the TONI

	TONI A First	TONI A Second	FI / NFI Combined
French Immersion	0.78	0.43	0.60
Non-French Immersion	0.02	0.53	0.29
FI / NFI Combined	0.58	0.46	0.52

treatments, TONI.A and TONI.B, and to test whether the order of administration had an effect on the scores. A contingency table of the observed and expected number of students that scored higher and lower on the second administered TONI was tested by the chi-square test. The results showed that the scores on the TONI.A and B were affected by the order in which they were administered (X = 5.91, $p < .025$) and suggests a learning effect.

When the two TONI tests are given one after another, the score of the second should be better than the first because the first test is a form of training. This was indeed the case when the TONI.A was given before the TONI.B. The average increase in score was 5.60 and was significantly different from zero ($t = 2.46$, $p = .02$). However, there was an average decrease in score of -1.63, although not significant ($t = -0.69$, $p = .50$), when the TONI.B was given before the TONI.A. The learning effect was also compared between French Immersion and Non-French Immersion and support the results of the combined samples. There was an increase in the TONI.B score when the TONI.A test was given first ($t = 1.78$, $p = .09$ FI; $t = 1.73$, $p = .11$ NFI) but no detectable increase in the TONI.A score was found when the TONI.B was given first ($t = -.61$ $p = .55$ FI; $t = -0.31$ $p = .76$ NFI).

As previously noted, the correlations of the TONI, Form A and B, with the WISC-R and with the CWFT were extremely low for the Non-French Immersion group, ranging from .02 to .29 (Table 4-15). Correlation coefficients were recalculated to see if these results might have been due to the order effect. TONI.A and TONI.B scores were segregated by the order the tests were administered and by the

type of school, French Immersion and Non-French Immersion. After this breakdown, there were 15 subjects in each of the subgroups. The Pearson correlation coefficient between Form A and B was high for all categories (Table 4-16) except for the Non French Immersion group when the TONI.A was administered first (r = .018). This is in sharp contrast to the French Immersion group (r = .78).

CHAPTER V

DISCUSSION AND CONCLUSIONS

The results of the WISC-R, the TONI, Form A and Form B, and the Children's Word Finding Test (CWFT), as administered to 30 Grade 3 French Immersion students, were compared to the results of the same tests administered to 30 Grade 3, Non French Immersion students.

All students came from homes in which English was the only language commonly used, and all had been enrolled continuously since Kindergarten in the current language of instruction.

The purpose of the study was to answer three questions:

1. Is there a difference in the level of verbal and non-verbal reasoning ability of Grade 3 French Immersion students and that of Grade 3 Non French Immersion students?

2. What is the concurrent validity of the TONI with the WISC-R and with the CWFT?

3. Are the two forms of the TONI equivalent?

As a first step, the socioeconomic status and the size of personal libraries of parents from the two groups (French Immersion and Non-French Immersion), were compared. These variables have been found to be correlated with level of achievement and cognitive ability of the children (Jeroski, 1984). No significant difference was found between the two groups on the number of books (p ($F = .0811$) $= .78$), and on socioeconomic status (p ($F = .0554$) $= .81$). Therefore, we

assumed that the socioeconomic status and library sizes of families in the two groups were equivalent. As a result, there was no need to use these variables as covariates in further analysis of the data.

The two groups on the other hand did come from families whose socioeconomic status was higher than average (1.4 SD above the mean). On the average, their home libraries contained between 250 and 499 books.

In regard to the three questions, various analyses led to the following results:

1. Difference in the levels of verbal and non-verbal reasoning ability between Grade 3 French Immersion group and Grade 3 Non-French Immersion group:

Data analysis shows that the French Immersion group scored higher on all the tests than did the Non-French Immersion group. There was a significant difference between the two groups on the WISC-R Performance Scale ($p = .01$), and on the WISC-R Full Scale ($p = .02$). There was a difference on the TONI A ($p = .07$), and on the CWFT ($p = .10$). There was no significant difference on the WISC-R Verbal Scale, nor on the TONI B.

The results on the WISC-R test differ from those of Nielsen (1983) who found no significant difference on the Full Scale and subscales of the WISC-R for Grade 2 students. The fact that there is a significant difference between the two groups on the Performance Scale, and no significant difference on the Verbal Scale may be due to the fact that French Immersion students had only one hour every day of

formal English in Grade 3 and none before. Although they speak English at home, they may be at a disadvantage as compared to Non-French Immersion students on subtests such as Vocabulary, Comprehension and Arithmetic. In fact, it is in these subtests that the smallest mean differences are found between the two groups. Moreover, in the Arithmetic subtest the difference favours the Non-French Immersion group.

On this analysis, the level of verbal and non-verbal reasoning ability, seems higher for French Immersion students than for Non-French Immersion students, although it is not significantly higher ($p = .07$ for TONI A and $p = .10$ for CWFT). One might suppose that this was due to the higher cognitive ability, of French Immersion students as compared to Non-French Immersion students, as measured by the WISC-R; A further analysis of data controlling for the level of cognitive ability was therefore undertaken to compare the results of the two groups in the TONI Form A and B and the CWFT. The adjusted means were still higher for French Immersion than for Non-French Immersion students on all three tests. The results showed a difference between the two groups on the TONI Form A ($p = .09$), on the TONI Form B ($p = .21$) and on the CWFT ($p = .10$). Those differences were not significant.

One might speculate that these results are explained by the learning process through which French Immersion students must go in order to learn a second language. In this process, they must attend more closely to the general properties of things in order to understand new words that are explained to them. They have to analyse

what the teacher says and reason constantly to verify whether what they understand make sense to them or not. Finally, they tend to think at a more abstract level since they have two codes for each word. This specific language learning activity constitutes an intense training in inductive and deductive reasoning, the abilities required for the TONI and for the CWFT.

Other variables could explain the difference between the two groups on reasoning ability. Any variable that would explain why parents chose French Immersion for their children, could be responsible for the difference found in the two groups on the three tests. Motivation is certainly an important variable in need of further investigation. Also, learning activity in French Immersion classroom should be compared with that in the Non-French Immersion classroom.

Research is needed to determine whether either treatment - French Immersion or Non-French Immersion - makes the difference, or whether the subjects (those who chose French Immersion and those who don't) make the difference, or whether the explanation is a combination of the treatment and of subjects' characteristics.

The results of the CWFT were compared to the norms calculated in 1982 (Table 4-3). The mean was higher for the entire sample (38.80) than the mean of the norms (36.80) for the same age group. However, it is important to note that in this study, the students had no limitation of time in which to give their answers. It is possible that such a limitation would have given results lower or equal to the 1982

norms. The standard deviation was smaller in the present study because of the restriction of range

The CWFT seems to involve memory. The testers noticed, while administering the test, that some children did not use their inductive or deductive reasoning abilities because they did not remember the clues given in previous sentences. It would be valuable to conduct a study on the correlation between results on a verbal memory test and on the CWFT to find out what part memory plays in the score obtained on CWFT

Finally, the CWFT appears to be a useful tool to test verbal reasoning ability. The fact that CWFT measures something different than general intelligence was demonstrated in this study by the difference between the two groups after controlling for differences in cognitive ability, as measured by the WISC-R. More research is needed on this promising test for purposes of standardization, construct validity and concurrent validity.

2. Concurrent validity of the TONI with the WISC-R and the CWFT:

Correlation coefficients between the TONI and the WISC-R were not very high ($r = .45$ for both forms). This was also the case for correlations of the TONI with the CWFT ($r = .36$) for Form A, and ($r = .39$) for Form B. The TONI manual indicates moderate to high correlations between the TONI and the WISC-R (.62 to .95). However, those correlations were established with Learning Disabled and Educable Mentally Retarded groups; therefore, probably with a restricted range of scores mainly at the lower end of the scale. It is

possible that subjects in those categories would score low on any test of cognitive ability. Thus, correlations between various tests related to cognitive ability would be fairly high. In the present study, the subjects scored at the higher end of the scale and the range of scores was wide. The low correlations obtained in this study constitute a warning against using the TONI as a substitute for the WISC-R with normal populations.

Although it might be tempting to use the TONI instead of the WISC-R in French Immersion situations as a language-fair test, it would be dangerous to do so for the results of this study show that the two tests are not equivalent.

The difference in age range norms in the TONI and the WISC-R could offer an explanation for the low correlations between the two tests. In this study, the age interval was 8-5 to 9-11, with only three children age 8-5. This age interval corresponds to only one age interval in the norms of the TONI (8-6 to 10-11), except for the three children aged 8 years 5 months, while it corresponds to five age intervals in the norms of the WISC-R (8-4 to 8-7, 8-8 to 8-11, 9-0 to 9-3, 9-4 to 9-7 and 9-8 to 9-11). In the age range 8-5 to 9-11, cognitive development is important and differences between the results of a child at 8-5 and at age 9-11 are expected to be large. However, the TONI gives only one set of norms for the huge interval of 29 months between 8-6 and 10-11. If the WISC-R and the TONI were equivalent tests, and assuming that norms of the TONI were based on the middle of the age interval (8-5 to 10-11), we should expect large differences at the two ends of the age interval (8-6 to 10-11) and

small differences at the middle of the age interval (9-9). In this case a graph of the differences between the WISC-R scores and the TONI scores should be positively skewed, since we only deal with the lower half of the TONI age interval. This was verified by plotting the difference between the WISC-R and the TONI A scores against the age of each subject, then fitting a smooth curve to the data (Figure 5-1). The scatterplot was smoothed using the LOWESS procedure (LOcally WEighted Smoothing of Scatterplots, Cleveland (1979)). The figure lends support to the hypothesis; it shows an increase between age 8-4 to 8-8 followed by a decrease through the age 9-9.

This explains in part the low correlations found in the age group with which this study is concerned (8-5 to 9-11), between the TONI and the WISC-R scores. Because of the difference in age range norms between the TONI and the WISC-R, it is recommended not to use the TONI test in place of the WISC-R test especially when the age of the students corresponds to a wide age interval of the TONI.

Although the way in which the TONI test is normed could be responsible for the low correlation coefficients between the TONI and the WISC-R, this still does not explain why those correlation coefficients are lower (.07 and .02) for the Non-French Immersion group than for the French Immersion group (.24 and .50). The intercorrelation between the Performance scale and the Verbal Scale of the WISC-R is lower for the Non-French Immersion group (.20) than for the French Immersion group (.42), both being compared to the standardization group (.64). An unexplained characteristic of the Non-French Immersion group distinguishes it from the standardization

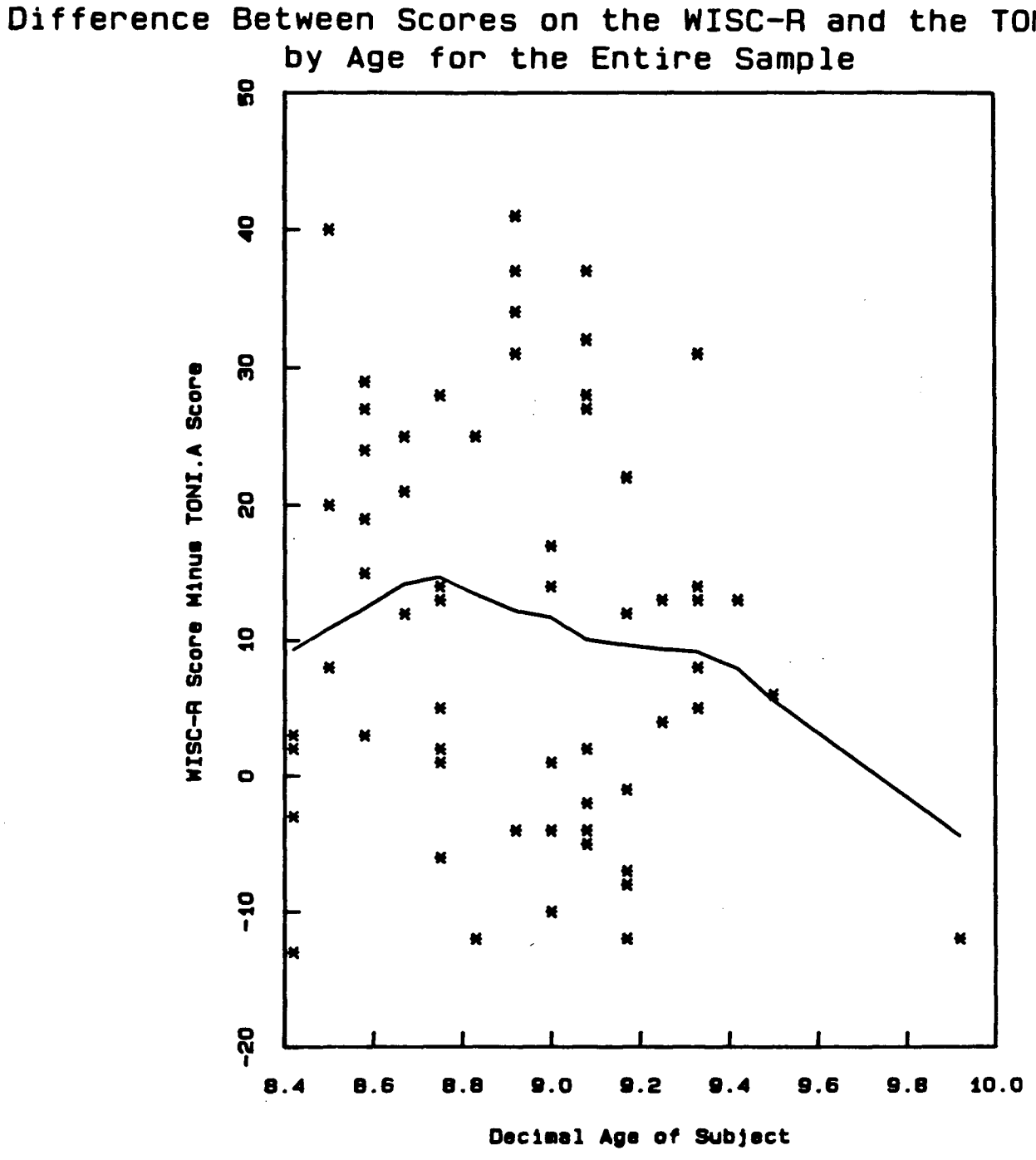

Figure 5-1 Differences between scores on the WISC-R and the TONI A by Age for the entire sample. The scatterplot was smoothed using the LOWESS procedure (LOcally WEighted Smoothing of Scatterplots, Cleveland (1979)). The figure confirms the hypothesis that the scores are different because of the age interval; it shows an increase between age 8-4 to 8-8 followed by a decrease through the age 9-9.

group, on the abilities required to complete visual and perceptual tasks. This characteristic could partly explain the low correlation coefficients between the TONI and the WISC-R found in the Non-French Immersion group.

Before we can use the TONI reliably, it would be appropriate to establish norms for smaller age intervals especially where cognitive development is important.

In comparing correlation coefficients between the TONI Forms A and B and the WISC-R on one hand, and between the TONI Forms A and B and the CWFT on the other hand, no significant difference was found for TONI A ($.20 < p\ (t > -.94) < .50$), nor for TONI B ($p\ (t > -.20) > .50$). From this analysis it appears that the TONI is no more correlated with the WISC-R than it is with the CWFT. Although it is called an "Intelligence test", the TONI does not measure general intelligence as the WISC-R does. On the other hand, the moderate correlation between the TONI and the CWFT could be explained by the fact that the kind of ability necessary to solve problems in the TONI is different from that needed in the CWFT. The TONI as well as the CWFT measure reasoning ability. However, the TONI requires visual and perceptual abilities while the CWFT requires verbal and linguistic capacities. Those two reasoning tests might discriminate between learning styles of students, some being more visual and other being more verbal.

3. Equivalence of the two forms of the TONI:

This study found a moderate correlation between Form A and Form B of the TONI test ($r = .52$). This result is significantly different ($p < .01$) from results obtained by the authors ($r = .78$) for age range 8-6 to 10-11. Among the equivalence coefficients by age reported in the technical manual for the test, the equivalence coefficient for age range 8-6 to 10-11 was the lowest of age ranges.

This finding suggests the wisdom of not using one form of the TONI as equivalent to the other in age range of 8.6 to 10-11.

Furthermore, the analysis showed that the order in which the two forms were administered had an effect on the results of the students. When Form A was administered first there was a significant learning effect on the results, whereas when Form B was administered first, there was no significant learning effect. These results show that the two forms are not equivalent.

The calculation of the correlation coefficients between the two forms of the TONI, separating the order effect shows interesting results. When TONI A is given first, the correlation between the two forms of the TONI is high ($r = .78$) for the French Immersion group and low ($r = .02$) for the Non-French Immersion group. The order effect seems to favour the French Immersion group and not to favour the Non-French Immersion group. It appears that the French Immersion group generalizes and consistently applies in Toni B the rules they have learned in TONI A whereas the Non-French Immersion population does not. Bain's study (1975), on the difference in discovery time and

transfer time between unilinguals and bilinguals when presented with problem solving, found no difference between the two groups at the concrete operation stage. This study is different in the sense that it investigates the consistency, as opposed to the time, with which unilinguals and bilinguals generalize the rules discovered in one task to other tasks. The ability to generalize and transfer is higher for French Immersion than for Non-French Immersion.

The TONI's ceiling rule suggests another explanation of the low correlation coefficients between the two forms of the test, and also between the TONI and the WISC-R. According to the rule, the tester must "discontinue testing when the subject has missed three out of five items" (Technical manual, p. 20). The ceiling rule for most tests of this kind is, by way of contrast, five items missed in a row. As a result, the students end up meeting the ceiling when they might have been able to get more items correct under the rule of five items missed out of five, This would be especially so if the order of difficulty of the items were not strictly controlled. A rule of five out of five items missed would give more stable results.

Differences in motivation between the two groups could also offer another explanation as to why in French Immersion the correlations between TONI Form A, TONI Form B, WISC-R, and CWFT were moderately strong, while they were extremely low for Non-French Immersion. In this explanation, one assumes that French Immersion students are more motivated than are Non-French Immersion students. As a consequence, French Immersion students would do their best on all tests while one would expect Non-French Immersion students to obtain inconsistant

results on the three tests. However, this assumption has not been tested here or elsewhere and is not a strong argument in explaining the results.

Before drawing any conclusion from this result, it would therefore be necessary to replicate the study using another sample, and measuring the level of motivation in each group.

Limitations of the Study

While the implications of this study are encouraging -- they show that French Immersion students scored higher on cognitive abilty tests, such as the WISC-R , they also indicate that French Immersion students have a tendency to obtain higher results, although not significantly higher, on reasoning ability tests -- there are some reservations. First, the study was limited to a small sample (N = 30) of French Immersion children at only one grade level (Grade 3). Second, the choice of schools was arbitrary and not random; all the students were from families with high socioeconomic status. Theoretically, generalization of results to different grade levels or even to all Grade 3 French Immersion students from any or all socioeconomic status, is not yet justified.

Conclusions and Implications for Practice

The results of this study indicate that for the age group 8-4 to 9-11 the TONI correlates only moderately with the WISC-R as well as

with the CWFT. Therefore, the TONI is an unreliable substitute for the WISC-R especially for the age group mentioned above.

The two forms or the TONI have a very low coefficient of equivalence for the age range 8-6 to 10-11. They are not comparable enough to be considered equivalent. Furthermore, there is a learning effect when Form A is administered first and no learning effect when Form B is administered first.

The French Immersion group scored higher than the Non-French Immersion group on the TONI A, and on the CWFT and significantly higher on the Performance Scale and Full Scale of the WISC-R. After controlling for variability on cognitive ability levels, the French Immersion group still scored higher, although not significantly higher, than the Non-French Immersion group on the TONI A and CWFT.

The French Immersion group has been shown to be better able consistently to transfer rules learned in the TONI A to the TONI B than the Non-French Immersion group.

Despite the fact that confounding differences between the two groups - such as motivation or the reasons that make parents chose French Immersion or not - we can conclude that in this study French Immersion students demonstrated higher general cognitive ability, as measured by the WISC-R, but no significantly different verbal and non-verbal reasoning ability than Non-French Immersion students.

The results on the TONI test suggest that more research is needed with regard to normalization, concurrent validity and equivalence of

the forms before it can be considered as a reliable substitute for the WISC-R test.

Implications for Further Research

The conclusions drawn from this study are limited to Grade 3 students from high socioeconomic status. It would be wothwhile to conduct further research on children from other grade levels, and from a wider range of socioeconomic status to verify the consistency of the results. More research is also needed to investigate the importance that motivation may have had on the results of the French Immersion group.

Moreover, there is a need to further study and compare the ability of French Immersion groups and Non-French Immersion groups to consistently generalize and transfer rules. This would provide further insights into how learning a second language might affect learning development with regard to cognitive processes.

This study questioned the equivalence of the two forms of the TONI and the difference in learning effect according to the order of administration of the two forms. Therefore, additional studies should be conducted to compare the equivalence of the two forms of the TONI and the order effect with other groups, at various age, grade and from a wide range of socieconomic status.

This study found also a need for standardization of the TONI using smaller age intervals, especially between age 8 and 12, when cognitive development is significant.

Finally, because the CWFT shows differences other than cognitive ability between groups such as French Immersion and Non-French Immersion students, it would be worthwhile conducting research on this test with regard to its standardization and the establishment of its concurrent and construct validity.

REFERENCES

Ahmed, M.A.S. (1954), Mental manipulation. Egyptian Yearbook of Psychology,1,23-88.

Bain, B. (1975) Toward an integration of Piaget and Vygotsky: bilingual considerations. Linguistics, 160, 5-20.

Balkan, L. (1970), Les effets du bilinguisme francais-anglais sur les aptitudes intellectuelles. Aimav, Brussels.

Barik, H.C. & Swain, M. (1976), A longitudinal study of bilingual and cognitive development. International Journal of Psychology. 11(4), 251-263.

Ben-Zeev, S. (1977), The influence of bilingualism on cognitive development and cognitive strategy. Child Development, 48, 1009-1018.

Binet, A. (1916), The Development of Intelligence in Children (The Binet-Simon Scale). Baltimore: Williams & Wilkins.

Blau, P. & Duncan, O.D. (1967), with Andrew Truce. The American Occupational Structure. New York: Wiley.

Blishen, Bernard, R. and Mc Roberts, Hugh, A. (1976) A revised socioeconomic index for occupations in Canada. Revue Canadienne de Sociologie et d'Anthropologie/Canadian Review of Sociology and Anthropology, 13(1), 71-79.

Brown, L., Sherbenou, R. & Dollar, S. (1982), The Test of Nonverbal Intelligence Manual. pro-ed, Texas.

Burns, G.E. (1982), Charges of Elitism in Immersion Education: The Case for Improving Program Implementation. Toronto: The Ontario Institute for Studies in Education.

Buros, O.K. (1985), the Ninth Mental Measurements Yearbook. New Jersey: Gryphon Press.

Carringer, D.C. (1974), Creative thinking abilities of Mexican youth: The relationship of binlingualism. Journal of Cross-Cultural Psychology,5, 492-504.

Cleveland, W. S. (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots", JASA, 74, 829-836.

Coleman, J. (ed) (1966), Equality and Educational Opportunity. Washington, D. C.: U.S. Governement Printing Office.

Cronbach, L. J. (1970), Essentials of Psychological Testing (3rd ed.). N.Y.: Harper & Row.

Cummins, J. (1975), Cognitive factors associated with intermediate levels of bilingual skills. Unpublished manuscript, Educational Research Centre, St Patrick's College, Dublin.

Cummins, J, (1976), The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypotheses. Working Papers on Bilingualism, 9, 1-43.

Cummins, J. (1978),Bilingualism and the development of metalinguistic awareness. Journal of Cross-Cultural Psychology,9, 131-149.

Cummins, J, (1979), Linguistic interdependance and the educational development of bilingual children. Review of Educational Research, 49(2), 222-251.

Cummins, J., & Gulutsan, M. (1974), Some effect of bilingualism on cognitive functioning. In S.T. Carey (ed), Bilingualism, Biculturalism and Education, Edmonton: the University of Alberta Press.

Cummins J.& Mulcahy, R. (1978), Orientation to language in Ukrainian-English bilingual children, Child Development, 49, 1239-1242.

D'Anglejan, A. Gagnon, N. Hafez, M. Tucker, G.R. Winsberg, S. (1978), Solving problems in deductive reasoning: three experimental studies of adult second language learners. Working Papers on Bilingualism, 17, 1-23.

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

Gart, J. J. (1969), An exact test for comparing matched proportions in crossover designs. Biometrika, 56, 75 - 80.

Harvey, F.B. (1974), Educational Systems and the Labour Market, Don Mills, Ontario: Longman Canada Ltd.

Heath, S.B. (1982), What no bedtime story means: narrative skills at home and school. Language in Society, 11, 49-76.

Holmes, B. J. (1981), Individually-administered intelligence tests: An application of anchor test norming and equating procedures in British Columbia. Doctor of Education thesis, University of British Columbia.

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

Jenrich, R. & Sampson P. (1985), General mixed model analysis of variance - equal cell size. In Dixon, W. J. (chief ed.), BMDP Statistical Software, 1985, Printing. Berkeley, Los Angeles, London.

Jeroski, S., C. Tolsma, G. Labercane. (1984), British Columbia Reading Assessment: Summary Report.

Lambert, W.E. (1974), Culture and language as factors in learning and education. In F. Aboud and F. D. Meade (eds), Cultural Factors in Learning. Bellingham: Western Washington State College.

Lambert, W.E. (1977), Effects of bilingualism on the individual. In Hornby, P.A. (ed), Bilingualism: Psychological, Social and Educational Implications. New york, San Francisco, London: Academic Press Inc.

Lambert, W.E. & Tucker, G.R. (1972), Bilingual education of children. The St Lambert experiment.Rowley: Newbury House.

Landry, R.G.(1974), A comparison of second language learners and monolinguals on divergent thinking tasks at the elementary school level. Modern Language Journal, 58, 10-15.

Laxer, G. et al (1974), Student Social and Achievement Patterns. Toronto: OISE, H.S. 1 Studies.

Liedke W.W. and Nielson L.D. (1968), Concept formation and bilingualism. Alberta Journal of Educational Research,14, 225-232.

MacNab, G.L. (1979), Cognition and bilingualism: a reanalysis of studies. Linguistics, 17, 231-255.

Macnamara, J.(1966), Bilingualism and primary education. Edinburgh: Edinburgh University Press.

Mehrotra, K.K. (1968), A Comparative study of WISC and Raven's Progressive Matrices, Psychological Studies, 13, 47-50.

Nie, N. H., C. H Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent. (1975) SPSS. Statistical Package for the Social Sciences. 2nd ed. McGraw Hill, New York.

Nielsen, B. (1983), Validation of the WISC-R for grade two French immersion students. Unpublished Master's thesis, University of Bristish Columbia, Vancouver.

Ninio, A. (1980), Picture-book reading in mother-infant dyads belonging to two sub-groups in Israel. Child Development, 51, 587-590.

Peal, E. & Lambert, W.E.,(1962), The relation of bilingualism to intelligence, Psychological Monographs, 76, 1-23.

Piaget, J. (1923), Le language et la pensee chez l'enfant, Neuchatel et Paris: Delacheux et Niestle.

Pintner, R. (1932), The influence of language background on intelligence tests. Social Psychology, 3, 235-240.

Raven, J.C. Court, J.H. & Raven, J. (1976), Manual for Raven's Progressive Matrices and Vocabulary Scales. H. K. Lewis & Co Ltd.: London.

Reitan, R. M., (1972) Verbal problem solving as related to cerebral damage. Perceptual and Motor Skills, 34, 515-524.

Rourke, Byron, P. and Fisk John, L. (1976), Children's Word Finding Test (Revised). University of Windsor and Windsor Western Hospital Centre.

Saer, D.J. (1923), Effect of bilingualism on intelligence. British Journal of Psychology, 14, 25-38.

Salvia, Y., & Ysseldyke, J.E. (1978), Assessment in Special Remedial Education. Boston, N. Y.: Houghton Mifflin.

Sattler, J. M. (1982) Assessment of Children Intelligence (2nd ed.) Boston, N.Y.: Allyn & Bacon.

Skutnabb-Kangas, T. & Toukomaa, P.(1976), Teaching Migrant Children Mother Tongue and Learning the Language of the Host Country in

the Context of the Socio-Cultural Situation of the Migrant Family. Helsinki: The Finnish National Commission for UNESCO.

Smith, F.(1978), Understanding Reading. New York: Holt, Rinehart and Winston, 2nd ed.

Smith, F. (1923), Bilingualism and mental development. Brit. Journal of Psychology, 13, 270-282.

Teale, W.H. (1982), Reading to Young Children: Its Significance for Literacy Development. In H. Goelman, A. Oberg, & F. Smith (eds), Awakening to Literacy, London: Heinemann Educational Books.

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

Trites, R. & Price, M.A. (1980), Assessment of Readiness for Primary French immersion: Grade one Follow-up Assessment. Toronto: Ministry of Education, Ontario.

Tsushima, W.T. & Hogan, T.P. (1975), Verbal ability and school achievement in bilingual and monolingual children of different ages. Journal of Education Research, 68, 349-353.

Vigotsky, L.S. (1962), Thought and Language. Cambridge, Mass.: MIT Press.

Wechsler, D. (1974), Manual for the Wechsler Intelligence Scale for Children - Revised. The Psychological Corporation, New York.

Whorf, B. (1956), Language, Thought and Reality. Cambridge, Mass.: MIT Press.

APPENDIX A

LETTER TO PARENTS

CONSENT FORM

QUESTIONNAIRE

VANCOUVER SCHOOL BOARD MAP

CHILDREN'S WORD FINDING TEST

Dear Parents:

_____'s School has agreed to participate in a research project: "Bilingualism and Reasoning Ability". The project involves the use of individual intelligence tests (verbal and nonverbal) and verbal reasoning tests in Early French Immersion schools and regular anglophone schools. The project requires the cooperation of 60 children in the Vancouver School District to take tests which are widely used. In addition, the parents of participating students will be asked to complete a short questionnaire which will be used in establishing home background information.

The researcher seeks to discover a possible relationship between bilingualism and the level of reasoning ability of Grade 3 French Immersion children. Also, the project will compare these children with their Grade 3 English classroom counterparts. A further purpose of the study is to statistically check the validity of a nonverbal test of intelligence when used with normal children.

The research project is being undertaken as a master's thesis in the department of Educational Psychology at the University of British Columbia. It has been approved by the Vancouver School Board's Student Assessment and Research office and by the principal of your school.

_____'s name was chosen as a possible participant in this research. If you and your child agree to participate, _____ will be asked to take part in

BILINGUALISM AND REASONING ABILITY

PARENT CONSENT FORM

I consent to _____'s participation in the testing research study at _____ school. I am aware that this will involve testing sessions totalling approximately two hours duration. I understand that confidentiality of test results will be maintained and that no individual scores will be released. I also understand that participation in this project is voluntary and may be terminated at any time.

_____

Signature

I would like a copy of the group results to be mailed to:.....
.....
.....

I am unwilling to have _____ involved in the testing research study.

_____

Signature

BILINGUALISM AND REASONING ABILITY

QUESTIONNAIRE

Your assistance in providing the following information would be very helpful in making this a meaningful study:

1 - What is the main language spoken in your home?

_____

2 - Do you sometimes use another language in the home?

Yes___ No___

If Yes, which one? _____

3 - Has your child been enrolled in the French Immersion program continuously from Kindergarten through to the present time?

Yes___ No___

4 - Has your child been enrolled in the regular English program continuously from Kindergarten through to the present time?

Yes___ No___

5 - Do you read to your child?

Yes___ No___

If Yes,

How often? _____

Do you discuss with him/her what you have read?

Yes___ No___ Sometimes___

How old was your child when you started reading to him/her?

_____

Did you read to him/her, more often when he/she was younger?

Yes___ No___

6 - How many hours a day on the average does your child watch T.V.?

7 - Approximately how many hours a week does your child read?

8 - QUESTIONS ADDRESSED TO THE MOTHER

8a) Do you work outside the home?

Yes, part-time___ No___

Yes, full-time___

8b) What is your occupation?_____

8c) Please circle the number in front of the category below, which best describes your completed level of education.

I Less than High School completion

II High School completion

III Post-Secondary, no degree

IV University or College degree

9 - QUESTIONS ADDRESSED TO THE FATHER

9a) Do you work outside the home?

Yes, part-time___ No___

Yes, full-time___

9b) What is your occupation?_____

9c) Please circle the number in front of the category below, which best describes your completed level of education.

I Less than High School completion

II High School completion

III Post-Secondary, no degree

IV University or College degree

Location of the Schools

Participating in the Study

in the Vancouver School Board District