INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in “sectioning” the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
THE EFFECTS OF SECOND-LANGUAGE STUDY ON THE COGNITIVE FLEXIBILITY OF FRESHMAN UNIVERSITY STUDENTS
THE EFFECTS OF SECOND-LANGUAGE STUDY
ON THE COGNITIVE FLEXIBILITY OF FRESHMAN UNIVERSITY STUDENTS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Kathleen Susan Cox, B.A., M.A.

* * * * *

The Ohio State University
1980

Reading Committee:
Gilbert A. Jarvis
James K. Duncan
Philip M. Clark

Approved By

Adviser
Department of Humanities
Education
To Fabián
who got me started
ACKNOWLEDGMENTS

Many people have had a part in bringing me to the position of being able to write this dissertation. My sincere thanks go out to each of them:

To Professors Edward D. Allen and Gilbert A. Jarvis for the fine preparation they have provided in the Foreign Language Education curriculum and to Professor John J. Kennedy for the the same in Research Design and Analysis.

To Professors Philip M. Clark and James K. Duncan for offering many excellent suggestions as members of my Reading Committee.

To Professor Jarvis, my dissertation director, for challenging me and inspiring me to do my best work and for the confidence he has always had in me.

To Jolaine Scholl for helping me set up and interpret the computer analyses; to Bette Stevens for scoring the Torrance Tests, for rounding up a validating group for the FRAT, and for helping me organize Chapter I; to Charlotte Leventhal for critiquing my drafts; and to all three of them for their friendship, without which I wouldn't have made it.

To Kent for his love and understanding.

And with love to my family for encouraging me to strive to do my best.
VITA

June 8, 1951.............. Born - San Francisco, California

1973...................... B.A. in Spanish, University of California, Davis, California

1973-1976................. Graduate Teaching Assistant, Department of Spanish and Classics, University of California, Davis, California

1975...................... M.A. in Spanish, University of California, Davis, California

1976-1978................. Graduate Research Associate, Department of Romance Languages and Literatures, The Ohio State University, Columbus, Ohio

1977-1980................. Graduate Teaching Associate, Department of Romance Languages and Literatures, The Ohio State University, Columbus, Ohio

PUBLICATIONS

The Ohio State University Spanish Individualized Program: Learning Packets for 101-103. Columbus: Department of Romance Languages and Literatures, The Ohio State University, 1978. With A. Cooper, A. Labarca, & M. Iglesias.


FIELDS OF STUDY

Major Field: Foreign Language Education

Studies in Foreign Language Education. Professors Gilbert A. Jarvis and Edward D. Allen

Studies in Research Design and Analysis. Professor John J. Kennedy
LIST OF TABLES

1. Reliability Coefficients for Torrance Tests of Creative Thinking, Form A with Form B .................................................. 33
2. Interrater Reliability Coefficients for Torrance Tests of Creative Thinking ................................................................. 34
5. Means and Standard Deviations of GPAs, ACT Scores and YRSLANG by YEAR, HSLANG, and CRSE101 ....................... 46
6. Values of $F$ for Manova of HSGPA and ACT Scores by YEAR, HSLANG, CRSE101 with Univariate $F$ Values .............. 48
7. Anova of YRSLANG by YEAR, HSLANG, and CRSE101 ................................................................................................. 52
8. Anova of GPAEND by YEAR, HSLANG, and CRSE101 ............................................................................................... 53
10. Values of $F$ for Manova of 1979 HSGPA and ACT Scores by HSLANG and CRSE101 with Univariate $F$ Values .......... 58
11. Means and Standard Deviations of Pre Flexibility Scores by HSLANG and CRSE101 ....................................................... 60
12. Values of $F$ for Manova of 1979 Pre Flexibility Scores by HSLANG and CRSE101 with Univariate $F$ Values ............... 61
13. Means and Standard Deviations of Post Flexibility Scores by HSLANG and CRSE101 ..................................................... 63
14. Values of $F$ for Manova of 1979 Post Flexibility Scores by HSLANG and CRSE101 with Univariate $F$ Values ............... 64
15. Values of $F$ for Ancovas Covarying Pre on Post Flexibility Measures ........................................................................ 67
LIST OF ABBREVIATIONS

Independent Variables

YEAR Year entered the university (1978 or 1979)
HSLANG Study of a second language in high school for longer than one year (yes or no)
CRSE101 Enrollment in a beginning second-language course (course number 101) Autumn quarter at OSU (yes or no)

Dependent Variables (A 1 or 2 after the abbreviation indicates a Pre or Post measure respectively.)

FIGFLU Figural fluency on Torrance Tests of Creative Thinking (TTCT)
FIGFLEX Figural flexibility on TTCT (number of categories)
FIGRAT Ratio of FIGFLEX to FIGFLU
FIGFLEX S Figural flexibility on TTCT (number of category shifts)
FIGRAT S Ratio of FIGFLEX S to (FIGFLU - 1)
FIGORIG Figural originality on TTCT
FIGELAB Figural elaboration on TTCT
VERFLU Verbal fluency on TTCT
VERFLEX Verbal flexibility on TTCT (number of categories)
VERRAT Ratio of VERFLEX to VERFLU
VERFLEX S Verbal flexibility on TTCT (number of category shifts)
VERRAT S Ratio of VERFLEX S to (VERFLU - 1)
VERORIG Verbal originality on TTCT
FRAT Functionally Remote Associates Test
I. INTRODUCTION TO THE PROBLEM

Statement of the Problem

Every moment of contact with the environment changes us. Within schools the educational environment is structured so as to effect change in an efficient and organized manner. Instruction is designed to increase students' knowledge of subject matter and to improve their proficiency in various skills. Growth in knowledge and proficiency is measured frequently, and learning is assumed to be taking place whenever changes in knowledge and skill are demonstrated.

Other types of learning or change also occur, however, as a result of the educational environment. These changes are less easily observable. They include changes in attitude, thinking skills, problem-solving abilities, and organizational skill. Changes in these areas are just as important for the overall development of the students as are changes in knowledge and skill. Yet we know little about them. In the field of foreign language education few research efforts have investigated these non-subject-matter outcomes of study. Educators have not determined exactly what is being learned in second-language classrooms in addition to the target language. There is a need to come to know these extra-linguistic outcomes of second-language study.

Early research efforts investigated the effects of second-language study on university academic achievement (Ross, 1931; Nelson, 1933; Fischer, 1945; Skelton, 1957). They found that students who had
studied a second language attained higher achievement levels in the university. These investigators admit, however, that it is often the good, hard-working student who chooses to study a second language. They were not able to make any causal statements in favor of second-language study.

Starch (1915) found no significant differences in the relative benefits of studying Latin or German. He also found that while second-language learners had a better knowledge of grammar, they were not able to use English any more correctly than those who had not studied a second language, and their English vocabularies were not greatly improved.

Thorndike (1924) investigated the mental disciplinary effect of various school subjects and found that students' intelligence and sex were much more important in predicting the score they received on the dependent measure than were any of the school subjects they studied.

These research efforts did not discover any important outcomes of second-language study. The research designs used in these investigations could, however, have lessened the likelihood of their being discovered if any were present. Studies conducted in the related fields of bilingualism and Foreign Language in the Elementary Schools (FLES), however, have shown that children who are bilingual or who have studied a second language in school for several years score significantly higher than monolingual children on various tests that seem to measure cognitive flexibility (Peal and Lambert, 1962; Balkan, 1970; Cummins and Gulutsan, 1974; Landry, 1974).
Cognitive flexibility is the term that many investigators have used to refer to a person's ability to change his or her patterns of thinking or his or her approaches to a task. Individuals who are said to be high in cognitive flexibility make these changes more easily and spontaneously than those who are low in cognitive flexibility. Those with greater cognitive flexibility are less rigid in their thinking. Many different cognitive tasks have been used to measure cognitive flexibility. Peal and Lambert (1962) used a test of general reasoning. Balkan (1970) used tests that he felt required subjects to change the mental sets they were using. Landry (1974), Carringer (1974), Cummins and Gulutsan (1974), and others have used divergent-thinking tests to approach this notion of cognitive flexibility. Despite the use of many different measures, in each of the studies the subjects who had had contact with a second language, either in the home or in school, showed evidence that they were less rigid in their thinking.

Two hypotheses have been suggested to account for the superior cognitive flexibility of bilinguals. Peal and Lambert (1962), Balkan (1970), Landry (1974), and Carringer (1974) suggest that students who study a second language develop a set for flexibility because they switch between languages often. This language switching, they feel, makes students more prone to switch to a different problem-solving approach whenever the original one loses its usefulness. Cummins and Gulutsan (1975) have cast doubt on the switching hypothesis. They found that bilinguals and monolinguals exhibited no differences in the rate at which they extinguished an haptic illusion. In order to extinguish the illusion, subjects must change an established response set.
If second-language study were to make one more prone to switching, bilinguals should extinguish the illusion more quickly. They did not.

Rather than being the actual language switching that contributes to heightened flexibility in second-language learners, Cummins (1976) suggests it is the recognition by students that language systems are arbitrary and that there are many ways to say something. It is this objectification of language, the separation of word and object, which, he feels, frees students to take diverse approaches to problem solving and which increases their cognitive flexibility.

Nespor (1970) investigated this phenomenon in third graders who had not yet separated word from object. This iconization of their language inhibited their ability to see relationships, to make inferences, and to synthesize when describing pictures. Nespor felt that foreign language learning might help them break their dependence on the concrete. In her study one group of students studied a foreign language in school while another group did not. Both groups were then asked to describe some pictures. The group that had studied a foreign language was significantly more expressive in its descriptions. The study of a second language had apparently helped them make the word/object separation and continue to grow in their communicative expressiveness.

Ianco-Worrall (1972) has also shown that second-language experience caused students to separate word and object earlier than students who were monolingual. A much greater number of bilingual than monolingual four- to six-year-olds grouped words semantically rather than acoustically. The bilingual children looked at language in a more analytic way than did the monolinguals.
None of these studies has investigated the effects of second-language study on college-aged learners. This study has undertaken such an investigation.

The use of children in the aforementioned studies somewhat limits the applicability of their findings as a theoretical base for research using college-aged learners. Children have not yet reached their full cognitive development. They are still learning concepts that adults can make full use of in learning. As Ausubel (1964) points out, adults learn a foreign language in a qualitatively different way than do children. Adults can learn from verbal propositions while children are still tied to experiences with concrete objects. Children cannot talk about language as a thing until they separate word and object. It is this objectification process that distinguishes adults from children in language learning.

This view is supported by Cook (1977) who feels that it is the use of general cognitive processes rather than language processes in language learning that characterizes the adult learner. Adults can see language as a thing.

In the Landry study (1974) it was only in the sixth grade that the FLES students showed significantly better divergent thinking ability though they began second-language study in the first grade. That a significant difference in ability did not appear until sixth grade is probably due, in part, to the cumulative effect of the amount of time spent on language study, as Landry suggests. It should also be noted, however, that it is in about sixth grade that children enter the Piagetian Formal Operations stage of learning in which language becomes a mode of learning new material directly and can be seen
objectively. Perhaps this ability is important in the attainment of
greater cognitive flexibility. If this is the case, there should be
even more of an effect shown with college-aged learners. They are all
capable of objectifying language. Moreover, the manner in which
second languages are taught at the university level requires that much
attention be given to the language as object, as a rule-governed,
arbitrary system.

The use of studies in bilingualism as a theoretical base for a
study using college-aged students as a population has the additional
limitation that very few college students reach the stage of complete
bilingualism, and certainly not after one quarter of study. Nevertheless,
whether one reaches a level of complete bilingualism or not, the study
of a second language does open one to the world of bilingualism. One
learns that all languages do not work alike, that different languages
express ideas differently, and, what may be most important in overall
mental growth, one learns that languages are arbitrary systems of
expression. It seems reasonable to expect that adult acquisition of
these insights would not depend on the attainment of high levels of
language proficiency. To whatever extent these insights produce changes
in cognitive flexibility, to that extent the studies in bilingualism and
FLES are applicable to studies of adult second-language learners.

An additional consideration in studies of cognitive flexibility
and second-language learning that use college students as subjects is
that many college students have studied a second language in high
school. In some cases high school second-language study is a college
entrance requirement. In other cases students have been advised that
the study of a second language in high school will help them in their college studies. Since some of the students used as subjects in the present investigation had studied a second language previously, two important questions arose, each of which had ramifications for the design of the study.

The first question involves the extent to which students who have studied a second language can be considered as belonging to the same population as those who have not. Are these two groups of students qualitatively different in ways that affect their cognitive flexibility? Do the more flexible students study second languages more often than the less flexible students? This question has been broached before with regard to intelligence, but not with regard to cognitive flexibility. It is an extremely important question because if the students who have studied a second language and those who have not are different, they cannot be compared in a meaningful manner.

The impossibility of absolute comparison of groups of individuals is the nemesis of behavioral research. Some attempt at comparison is, however, necessary. The grade-point averages and college entrance test scores of the groups can be compared to see if the students differ in scholastic aptitude, verbal skills, and school achievement. Their cognitive flexibility can be tested and compared at the outset. This will give some idea, albeit a limited one, of whether the groups can be considered as belonging to the same population.

The second question raised by the fact that some students had studied a second language in high school bears directly on the outcome of the study: How will this prior second-language study interact with college second-language study in affecting cognitive flexibility?
When does second-language study produce effects in flexibility? Is it on first exposure to second-language study? Does the type of second-language course make a difference? As was noted earlier, second-language study in college tends to be more intensive and grammar-oriented than high school second-language study. If Cummins is correct in suggesting that it is the objectification of language that increases flexibility, then perhaps the effect of college second-language study will overshadow any effect produced by high school study. This question is more easily studied than the first question. High school second-language study can be built into the design of the research as an independent variable, and its effects can be analyzed.

Purpose of the Study

The purpose of this study was to evaluate the effects of the study of a 101-level beginning course in foreign language on measures of the cognitive flexibility of university freshman students who have a foreign language graduation requirement. These students were chosen in an attempt to standardize the motivation to study a second language. They all know that at some point they must study a second language in order to graduate.

The study addressed the following research questions:

1. What are the effects of 101-level study of a second language on measures of cognitive flexibility?

2. Are there interactive effects between prior high school second-language study and the 101-level study of a second language on cognitive flexibility?
3. Do students entering the university with high school foreign language credit differ from those who enter without it in their high school grade-point averages or their ACT English or Composite scores?

4. Do freshman students with a foreign language requirement who take a 101-level foreign language course during their first quarter at the university and students who do not take a 101 course that quarter differ from each other in their high school grade-point averages or their ACT English or Composite scores?

Operational Definitions

The following definitions have been used for this study:

1. **Cognitive flexibility** is defined as that ability measured by the Torrance Tests of Creative Thinking and the Functionally Remote Associates Test. Though the Torrance Tests are scored for fluency, flexibility, originality, and elaboration, flexibility will be the dimension of primary interest.

2. **Fluency** is the score on the Torrance Tests of Creative Thinking that consists of the number of different responses made to a given stimulus.

3. **Flexibility** is the score on the Torrance Tests of Creative Thinking that consists of the number of different functional categories into which responses to the stimulus fall. In the Functionally Remote Associates Test flexibility is defined as the number of correct associations given in response to the stimuli.

4. **Originality** is the score on the Torrance Tests of Creative Thinking that is derived by summing points given for unique and uncommon responses to the stimulus in comparison to the norm group of the test.

5. **Elaboration** is the score on the Torrance Tests of Creative Thinking that is derived by summing points given for extra details drawn in the figure completion task.

6. **High school grade-point average (HSGPA)** is the mean of all grades received in grades 9 through 12 as reported on the official transcript, weighted according to the credits awarded the course, where A=4 pts., B=3 pts., C=2 pts., D=1 pt., F(E)=0 pts. When the GPA was computed on some other scale, the researcher recomputed it on the four-point scale.
7. ACT English and Composite Scores are the scores received on the American College Testing college entrance examinations. These tests have a range of 1-36. The Composite score is the mean of the four subtests: English, Mathematics, Natural Science, and Social Science. Students enrolled at the Columbus campus of The Ohio State University Autumn quarter 1979 had a mean Composite score of 20.3 and a standard deviation of 5.5. In this study the English and Composite scores are considered indices of general verbal ability and general intellectual ability respectively.

8. GPA end of Autumn is the mean of all grades received Autumn quarter at The Ohio State University, weighted according to the number of credits awarded each course where A=4 pts..., E=0 pts. These data were compiled from official college records.

9. High school foreign language study is defined as having studied any second language in grades 9 through 12 for longer than one year as indicated on the official high school transcript.

10. 101-level foreign language study is defined as being enrolled in the first course (course number 101) in any second language at The Ohio State University for an entire quarter.

Limitations of the Study

The results of this study are generalizable to populations of similar composition to that of the one used here. That is, they are generalizable to freshman students at The Ohio State University who have a foreign language graduation requirement. Since students decide for themselves whether or not to take a second-language course in any given quarter, it was impossible to assign students randomly to the treatment groups. High school GPA and ACT scores were included in the analysis in an attempt to determine whether or not students who chose to study a second language and those who did not differed significantly in their academic and mental abilities. Since other differences among the students were not measured, the results of the study are biased to...
the extent that any of these other differences significantly discriminates among the groups. The use of volunteers also admits bias into the study. The GPAs and ACT scores of a random sample of students from the previous Autumn quarter were collected in an attempt to determine whether or not the volunteer group was representative of the larger population.

The question of whether tests of divergent thinking and remote associational ability are valid measures of cognitive flexibility is still largely unanswered. The use of these measures here was based upon precedents set in earlier studies. Whether or not the measures used in this study produce a true indication of cognitive flexibility, it is likely that they measure the same effect that has been observed in previous studies. The ultimate purpose of this study would still be achieved: the discovery of non-linguistic outcomes of second-language study.
II. RELATED RESEARCH

Introduction

Few research studies have attempted to isolate the non-linguistic outcomes of second-language study. The studies that have tried to observe the effects of study of a second language have used global dependent measures such as university achievement. This broad scope and the limitations imposed by the studies' research designs may have kept them from finding any reliably significant outcomes. For the most part, these studies were carried out during the first half of this century. In the 25 years since then, the question of how the college-aged second-language learner has been affected by his language experience has not been investigated.

There has been a great deal of research effort, however, in the areas of bilingualism and Foreign Language in the Elementary School (FLES). Researchers have been prompted by a desire to know how children's cognitive development is affected by having two language systems available to them. Many excellent studies have been conducted and certain outcomes seem to be found consistently. Because the majority of these studies are of children who have, for the most part, equal facility in both of their languages, they are not directly applicable to the present study. These studies do, however, point to an area of possible effects that should not be left uninvestigated. Indeed, certain of the hypotheses that the researchers offer to explain the greater
cognitive flexibility of the bilingual children seem to suggest that adults could experience the same effect without attaining levels of complete bilingualism. The present study attempted to ascertain whether or not those hypotheses are correct.

Early Research on Outcomes of Second-Language Study

Early research into the outcomes of second-language study centered on its effect on university academic achievement. There were several historical reasons for this emphasis. The notion of mental discipline was still somewhat popular during the first quarter of this century, and several studies sought to determine its credibility. Other studies stemmed from the desire on the part of university officials to find a good predictor of college success which could be used in admissions procedures. As times changed and more students entered the university without second-language credit, still other studies were designed to include students who had had no foreign language study at all.

In 1915 Starch undertook three studies that were among the first to provide empirical data on the effects of second-language study. His first study sought to determine the relative benefit gained from the high school study of Latin and German on college scholarship. He compared the median grade-point averages of groups of students who entered the University of Wisconsin with two to four years of high school Latin study to that of those who entered with two to four years of high school German study and found there to be a very slight advantage in
favor of the Latin group. His conclusion was that, rather than being attributable to any difference in disciplinary value between the languages, the difference was present in the students themselves, and that Latin attracted more of the better students than did German.

Still, the difference was very slight. In this study Starch also found that the average grade of the students increased with the number of years of foreign language study they had had.

Starch's second study investigated the effect of Latin study on the size of students' English vocabulary. He found that, especially at the high school level, students who had studied Latin had slightly larger vocabularies (4.5% larger). At the college level the difference was much smaller (2.7%), however, and Starch felt that the English meaning of Latin-root words had changed sufficiently from the original that the study of more English vocabulary would be a much better way to increase vocabulary size.

The third investigation sought to determine whether or not foreign language study increased knowledge of English grammar and correctness of English usage. He found that knowledge of grammar terminology was, indeed, improved through foreign language study, by as much as 34.9%. Correctness of English usage, however, was not very much affected. In this study it was also found that Latin had no advantage over other foreign languages in producing gains.

Starch concluded that claims of the mental disciplinary benefits of Latin study were exaggerated. Although students with more second-language study had better grades in college, he felt that this was more a function of the intelligence and determination of the students than of the value of second-language study.
Thorndike (1924) carried out a large-scale investigation to ascertain the mental discipline acquired in various high school subject-matter areas. He administered an IQ-type test at the end of two consecutive school years and compiled data on the school courses that each student took during the intervening year. An elaborate gain-score procedure was used to group courses that produced similar results, and then sets of courses were compared to determine the relative mental discipline acquired in each. French and Latin were grouped together and produced relatively large gains. Overall, however, original IQ score, sex, and race contributed more to the overall differences than did any school subject. What one was like as a person was more important in determining improved scores than were the courses one took in school. After Thorndike's study, little was made of the mental discipline idea.

In the 1930's, universities were eager to find some measure that would be a good predictor of college success. Ross (1931) and Nelson (1933) compared the number of entrance credits students had in various subject areas with their overall grade averages during their first year in college. Both found that the relationship was stronger for foreign language than for any other area including math.

Nelson found that the number of high school foreign language credits correlated .30±.05 with college scholarship. Math correlated only .11 and English rhetoric and composition -.13. Science correlated .07, social science -.10, and all subjects together correlated .32 with college success. Nelson also found that for students with foreign language entrance credit the average score on the American Council
Psychological Exam was in the 56th percentile whereas students without foreign language credit averaged a score in the 37th percentile.

Ross found that the more social science or natural science credits a student had when he entered college, the worse he did during his first semester. The relationship between number of math credits and achievement was positive, but it was not as strong as the positive relationship between foreign language and achievement. Ross totally rejected the mental discipline theory. Rather than foreign language study in high school being the cause of college achievement, he felt that foreign language study was chosen by those students who were willing to work hard. Consequently, these students worked hard in college as well. Social science and natural sciences were seen as easier subjects. Lazier students chose them and, likewise, did more poorly in college than those who chose foreign language and math.

The question of whether foreign language study improves overall achievement or whether better students choose to study a foreign language is present in all of the early investigations. Most researchers recognized this problem and tried to address it. Fischer (1945) compared the high school rank as expressed in percentiles of freshmen at the University of Illinois who chose to study a foreign language with that of those who did not. None of the students had a foreign language requirement. He found in two successive years that the high school rank of the students who chose to study a foreign language in college was dramatically higher than that of those who did not. In 1943-1944 there was a 15.3 percentile-point difference; in 1944-1945 there was a 19-point difference. He also found that
advocated strongly recommended foreign language study for good students, and that these students also recognized its value and, thus, chose it for themselves.

Skelton (1957) controlled for intelligence level when he compared the academic performance of college freshmen who had studied a foreign language in high school with that of those who had not. He found that the students who had studied a foreign language performed better. Skelton believed that foreign language study had improved the students' native language skills which, in turn, had allowed them to do better in their other courses. He did not, however, take into account the possibility that the students who chose to study a foreign language were those who worked harder as Ross (1931) had hypothesized. Many other variables could have influenced the better performance of the students who had studied a foreign language as well.

If the absolute effect of foreign language study on students is to be observed, the problem of self-selection needs to be overcome. Groups that begin at different levels on as important a variable as high school rank cannot easily be compared. In addition, if any non-linguistic outcomes of foreign language study are to be discovered, measures that are less global than academic achievement must be employed. If foreign language study does, in fact, improve academic achievement, what cognitive skills have been strengthened to produce those gains? As Jarvis (1978) points out, it is the micro-level thinking skills that are most likely to be the ones strengthened through second-language study. Which of these skills are affected cannot be determined without the use of more specific measurement instruments.
Research in FLES and Bilingualism

Recent research into the cognitive effects of being raised bilingually or of participating in FLES programs has tried to be more specific in its measurements. Most of this research has been conducted with children since it seems that true bilingualism is most easily attained in childhood. Researchers have been particularly interested in discovering how bilingual experiences affect the cognitive development of children. Studies have been conducted using children who attained bilingual status before coming to school as well as those engaged in bilingual education and foreign language immersion programs in school. The work of these researchers is applicable to the present study in that it offers some indication of possible effects of second-language study on mental processes.

Until the Peal and Lambert (1962) study made use of balanced bilingual children, that is, children who had relatively equal proficiency in both of their languages, studies in bilingualism had indicated that bilingual children were handicapped in their verbal intelligence. Peal and Lambert found that balanced bilingual ten-year-olds from middle-class homes in Montreal scored higher on both nonverbal and verbal measures of intelligence than did their monolingual counterparts. A factor analysis of their cognitive measures showed a more differentiated subtest profile for the bilingual group. This may indicate that bilingual experiences lead to a more complex and, therefore, more flexible cognitive structure.
Peal and Lambert's study served as a starting point for many other studies into the cognitive effects of balanced bilingualism. In Switzerland, Balkan (1970) felt that Peal and Lambert had confounded intelligence and flexibility. He designed an excellent study in which he tested bilingual and monolingual children on an Embedded Figures Test and on a test that required sensitivity to the different meanings of words. Balkan believed that both of these tests required that the subjects break their normal or established response sets in order to answer correctly. This, he believed, was a measure of their cognitive flexibility. He found that the bilingual children scored significantly higher in cognitive flexibility. The bilingual and monolingual children had been matched on nonverbal intelligence.

In Mexico, Carringer (1974) compared balanced bilingual 14- to 16-year-olds with a monolingual group on the Torrance Tests of Creative Thinking (Torrance, 1966). The bilinguals scored significantly higher on verbal flexibility and originality and figural fluency and originality. These results are somewhat questionable, however, because Carringer did not control for intelligence or socio-economic status (SES). His is the only study that has made use of subjects over the age of ten or eleven.

Cummins and Gulutsan (1974) repeated the Peal and Lambert study in Western Canada with balanced bilinguals matched with monolinguals for SES, age, and sex. In addition to the tests used in the earlier study, Cummins and Gulutsan gave the students a test of divergent thinking as a measure of cognitive flexibility. They found that the bilinguals had scored higher in nonverbal and verbal intelligence and
on the verbal originality measure of divergent thinking. When compared with non-balanced bilinguals, children whose native language was dominant, they found that the non-balanced children scored worse than the monolingual group on the flexibility and fluency dimensions of the divergent thinking test.

In another study (Cummins, 1975), balanced bilingual sixth graders scored significantly higher on fluency and flexibility and almost so in originality than had the non-balanced group. When these children were compared to a monolingual group matched on verbal and nonverbal intelligence and SES, Cummins (1977b) found that the monolinguals scored the same as the balanced group and higher than the non-balanced group on fluency and flexibility. In verbal originality they scored lower than the balanced group and the same as the non-balanced group.

These results suggested to Cummins that perhaps there was a threshold level of second-language competence that children had to reach before they could be affected positively by bilingual experiences (Cummins, 1977a). He believes that unless students attain a level of second-language proficiency that lets them learn effectively in school, they will become confused and experience negative effects in their cognitive development, especially in verbal flexibility and fluency. Once a level of balance is achieved, students seem to be positively affected in their cognitive development by bilingual experiences. Since the students in the present study did not use the second language to learn new material, the threshold-level theory, even if valid, should not apply.
Bruck, Lambert and Tucker (1973) reported that the pilot class in the St. Lambert Home School Language Switch program scored significantly better than the controls on divergent thinking tasks in grades 3, 5, and 6. Moreover, Scott (cited in Cummins, 1977b) analyzed the St. Lambert experiment for the relationship between divergent thinking and bilingualism. With IQ and SES as covariates, he found that the level of French speaking skill in late elementary school was significantly predicted by third grade divergent thinking ability. This study opens the possibility that bilingualism can both influence and be influenced by divergent thinking.

Landry (1974) moved the research on FLES in a new direction by investigating the positive effects of a FLES program on students' divergent thinking ability. Previously, FLES research had been directed toward showing that there were no detrimental effects on FLES students. Landry administered several subtests of both the figural and verbal sections of the Torrance Tests of Creative Thinking (Torrance, 1966) to first, fourth, and sixth grade students in FLES and non-FLES schools. The tests were scored for fluency, flexibility and originality. In the first and fourth grade samples no significant differences were found, although in the fourth grade the directionality of the means favored the FLES group in all variables but verbal fluency. At the sixth grade level, however, the multivariate $F$ test was significant at the .05 level in favor of the FLES group. The sixth grade FLES students showed significantly greater divergent thinking ability than the non-FLES sixth graders. The students had begun second-language study in first grade and had received from 20 to 45 minutes of instruction per day in the second language.
Landry concluded that a certain amount of exposure to and experience with the target language must be achieved before the relationship between second-language study and divergent-thinking ability becomes significant. This somewhat parallels Cummins' threshold theory. In the program Landry studied, reading and writing of the second language are not introduced until grade four. It is only then, according to Landry, that the students become consciously aware of the differences between the two languages. At that point students begin to compare and contrast the languages. By sixth grade, it seems, the effect of this contrastive linguistic process becomes significant.

Landry believes that experience with two languages leads students to develop a set to switch—a set to be flexible. This, Landry feels, is then put to use in the rest of the student's life, making him more flexible in his thinking in general. This is the view suggested by Peal and Lambert (1962), Balkan (1970), and Carringer (1974) as well.

Cummins and Gulutsan (1975) have brought this hypothesis into question, however, in a test of the rate of extinction of the Uznadze haptic illusion. In this test the illusion is acquired by developing a response set. Subjects sit with their eyes closed and are handed two spheres. They feel them, and then the spheres are taken away. Subjects must say whether the spheres were the same size or not. Unequal spheres are used several times and then equal spheres are presented. If the subject says they are different, the response has been fixated. To extinguish the set the equal spheres are presented consistently until the subject says they are equal five times in a row. Using grade six balanced bilinguals and monolinguals, Cummins and
Gulutsan found no differences in the rate at which the illusion was extinguished. This result is inconsistent with the switching hypothesis because if language switching leads to a more flexible response set, the bilinguals should have extinguished the set more quickly. They did not do so. Of course, the Uznadze illusion is a test of motor set extinction. It is possible that verbal response set is acquired apart from motor response set and that a test of verbal set extinction would, indeed, have shown the bilinguals to extinguish the set more rapidly than the monolinguals.

Cummins (1976) suggests that perhaps it is the process of objectification of language rather than language switching that influences the cognitive flexibility of bilinguals. He feels that by having two linguistic perspectives the child can see more relevant aspects of problem situations and act in a more diversified way. When an object is seen as separate from the word used to describe it, it can become the focus of special attention, which can yield new knowledge. Likewise, the words themselves can be seen objectively.

That early bilingualism accelerates this separation of meaning and sound was first suggested by Leopold (1949). It has been empirically shown by Ianco-Worrall (1972). She asked Africaans-English four- to six-year-old bilinguals and English monolinguals to tell her which two of three words were similar. The triplets included two words that meant the same thing and two that sounded similar. Fifty-four percent of the bilingual children chose the semantically similar words, whereas practically none of the monolingual group did. The bilinguals were more sensitive to the semantic aspects of the words and seemed to look at language more analytically.
Nespor (1970) has shown that even small amounts of second-language study can help children make the word/object separation. In her study one group of third graders who were experiencing iconicity in their oral expression was exposed to foreign language study in school. Another group was not. Later they were asked to describe several pictures. The students who had not studied a foreign language continued to describe only the concrete, factual information in the pictures. The students who had studied a foreign language, however, described relationships and synthesized information. Second-language study apparently helped the students break through the object/word barrier.

Summary

Early studies of the outcomes of second-language study centered on global measures of academic achievement and did not reveal many promising areas of investigation. Studies in bilingualism and FLES have shown that balanced bilingual children score higher than non-balanced bilingual and monolingual children in cognitive flexibility. This advantage may be the result of an awareness of the arbitrariness of language labels and the separation of word and meaning. The objectification of language may allow children to focus their attention on more relevant aspects of a problem and free them to be more flexible in their approaches to it. The bilingual studies are relevant to this research in that they point the way to an area of investigation using college-aged second-language learners.
III. DESIGN AND ANALYSIS

Population and Sample

The population for this study consisted of freshman students enrolled in the University Survey course that enrolls students who plan to enter the Colleges of Arts and Sciences (ASC) of The Ohio State University. In ASC, students have a graduation requirement of four quarters of study of a foreign language.

The study proceeded in two phases. In the first phase a sample of 137 students was randomly selected from the ASC sections of the University Survey course of Autumn quarter 1978. These students were assigned to one of four groups according to whether or not they had studied a second language in high school and whether or not they were enrolled in a 101-level second-language course that quarter.

In the second phase the sample consisted of 51 volunteers from the ASC sections of the University Survey course of Autumn quarter 1979. These students were likewise assigned to one of the same four groups on the basis of their foreign language learning background.

Design

The first phase of the study employed a 2 x 2 factorial design. The independent variables were (a) second-language study in high school (two levels: yes, no) and (b) 101-level second-language study during
their first quarter of study at OSU (two levels: yes, no). Of the 137 subjects, 31 studied a second language in high school and were in a 101-level course, 78 studied a second language in high school but did not take a 101-level course Autumn 1978, 6 had not studied a second language in high school but did take a 101-level course, 22 had not studied a second language in high school and did not take a 101-level course Autumn 1978.

The dependent variables were (a) high school grade-point average (HSGPA), (b) ACT English score (ACTENG), (c) ACT Composite score (ACTCOMP), and (d) end of Autumn 1978 grade-point average (GPAEND). These measures were used to determine whether or not the four groups could be said to have come from the same population on the basis of their scores.

The second phase of the study also employed a 2 x 2 factorial design and used the same independent variables as the first phase: second-language study in high school and 101-level second-language study during their first quarter of study at OSU. Of the 51 subjects, 15 had studied a second language in high school and took a 101-level second-language course Autumn 1979, 25 had studied a second language in high school but did not take a 101-level course, 4 had not studied a second language in high school but did take a 101-level course, 7 had not studied a second language in high school and did not take a 101-level course or any other second-language course Autumn 1979.

The dependent variables of the second phase were the following:

(a) high school grade-point average (HSGPA)
(b) ACT English score (ACTENG)
(c) ACT Composite score (ACTCOMP)
(d) end of Autumn 1979 grade-point average (GPAEND)
(e) verbal fluency from the Torrance Tests of Creative Thinking (TTCT): Unusual Uses subtest (VERFLU)

(f) verbal flexibility from TTCT: Unusual Uses—number of categories (VERFLEX)

(g) ratio of (f) to (e) (VERRAT)

(h) verbal flexibility from TTCT: Unusual Uses—number of category shifts (VERFLEX S)

(i) ratio of (h) to ((e) - 1) (VERRAT S)

(j) verbal originality from TTCT: Unusual Uses (VERORIG)

(k) figural fluency from TTCT: Incomplete Figures subtest (FIGFLU)

(m) figural flexibility from TTCT: Incomplete Figures—number of categories (FIGFLEX)

(n) ratio of (m) to (k) (FIGRAT)

(o) figural flexibility from TTCT: Incomplete Figures—number of category shifts (FIGFLEX S)

(p) ratio of (n) to ((k) - 1) (FIGRAT S)

(q) figural elaboration from TTCT: Incomplete Figures (FIGELAB)

(r) figural originality from TTCT: Incomplete Figures (FIGORIG)

(s) score on the Functionally Remote Associates Test (FRAT)

The flexibility measures were taken at the beginning of the quarter and again at the end. Equivalent forms of the measures were used. The scores were used to determine whether or not the four groups could be said to have come from the same population at the beginning and at the end of the quarter. GPAs and ACT scores were collected from the subjects' University College files with permission from Dean Mount. These scores were used to determine whether or not the four groups could be said to have come from the same population and whether or not that population was the same one as the one from which the 1978 random sample was drawn.
Instrumentation

Three instruments were used to measure the cognitive flexibility of the subjects: the figural Incomplete Figures subtest of the Torrance Tests of Creative Thinking (TTCT), the verbal Unusual Uses subtest of the TTCT, and the Functionally Remote Associates Test (FRAT). Equivalent forms of each of these instruments were given at the beginning and end of the quarter.

The Torrance Tests of Creative Thinking were developed to measure figural and verbal divergent thinking ability. Both the figural and verbal tests are composed of several subtests each of which is intended to measure a different aspect of creative thinking behavior. The figural subtests are:

1. Picture Construction: Subjects must make a picture out of an abstract form. The subtest is scored for originality and elaboration. The subject is required to find a purpose for a purposeless form.

2. Incomplete Figures: Subjects must complete 10 incomplete line drawings in an original manner and give each a title. The incomplete drawings set up a tension in the subject to complete the drawing in the easiest way. To be original, the subject must resist the impulse to closure and make a jump away from the obvious drawing.

3. Repeated Figures: Subjects are asked to make different drawings out of a repeated pattern of lines (a circle or parallel lines). Here the task is to make multiple and varied associations to a single stimulus.

The verbal subtests are as follows:

1, 2, 3. Asking, Guessing Causes, Guessing Consequences: Subjects must show their curiosity and ability to develop hypotheses.
4. Product Improvement: Subjects must give suggestions for the improvement of an object. This allows the subjects to "brainstorm" and play with ideas.

5. Unusual Uses: Subjects must list as many unusual uses as they can think of for tin cans or cardboard boxes. The task is to free the mind from an established set (i.e. box, can as containers) and, thus, create unique responses.

6. Unusual Questions: Subjects must invent questions about boxes or cans that could have many possible answers.

7. Just Suppose: Subjects must imagine the consequences of a supposed event. This task requires the subjects to carry thoughts out to logical conclusions.

Because all testing of subjects had to be completed in one 48-minute period, only one figural and one verbal subtest could be selected for use in the study. A conscious effort was made to choose the subtest that seemed most likely to be affected by second-language study. That is, the subtest most similar to the cognitive requirements of second-language study was chosen from each test. For the figural test, the Incomplete Figures subtest was chosen. To succeed in second-language study, the learner must avoid the impulse of assigning meaning too quickly according to the common language code: English. The learner must tolerate ambiguity for a period of time. This process would seem to be similar to the avoidance of easy closure in order to draw unusual pictures.

The verbal subtest chosen was the Unusual Uses. In language learning, students must free the mind of the well-established mindset of the English language. They must begin to think in new, unusual ways and sometimes assign new meanings to well known things. Likewise, to score highly on this subtest, students must see familiar objects in new ways, to break the response set of box or can as container.

Each of the subtests is scored for several different constructs: fluency, flexibility, originality, and in the figural test, elaboration.
The scoring procedures are outlined very carefully in the scoring guide to the tests. For the Incomplete Figures test, fluency is the number of figures out of 10 that a subject completes. Each of the figures is then given a flexibility category number and an originality weight ranging from 0 (common) to 2 (unusual). There are 68 flexibility categories to choose from. For each figure, a list of the most common responses is given with category number and originality weight. For these examples the originality weights were based upon the responses of a norm group of 500 subjects ranging in age from kindergarten to college-aged. A weight of 0 was awarded if the response was given by 5% or more of the subjects. One point was awarded for responses given by 2% - 4.99% of the subjects. Two points were awarded for responses given by fewer than 2% of the subjects. For responses not given in the examples, the scorer must determine the category and originality weight him or herself. The scorer also has to determine the elaboration points awarded— one point for each extra detail in the figure drawn.

Once each response is categorized, the scoring takes place. Fluency is the number of figures completed out of the 10. Flexibility is the number of different categories of response given. Originality is the sum of the originality points given. Elaboration is the sum of the elaboration points given. For this study several other measures were taken as well. Since Torrance states (Torrance, 1974, p. 57) that flexibility is the ability to produce a variety of kinds of ideas and to shift from one approach to another, it was decided to compute the number of shifts in category to see if any different results would occur. In addition, a ratio was taken of the flexibility to fluency
scores and of the flexibility (shift) to (fluency - 1, the number of possible shifts). This was done to try to get a flexibility score more free of the influence of fluency: The number of categories or category shifts being determined by the number of responses, as Clark and Mirels (1970) point out.

A similar procedure was followed for the verbal subtest. Each response was given a flexibility category (of 28) and an originality weight. Many example responses are given with categories and weights. Here again, fluency is the number of different responses. Flexibility is the number of different categories of response. Originality is the sum of the originality points. As with the figural test, the flexibility (shifts) and the two ratios were also computed. A potential problem arose during the scoring of originality. The scorers had the impression that they were giving higher originality weights for Form B than for Form A. As a result, the example responses given in the scoring guide were scrutinized. It was found that for Form A (cardboard boxes) there were 113 examples with a mean originality weight of .93 (range, 0 - 2) and a standard deviation of .66. For Form B, however, there were 180 examples with a mean of 1.25 and a standard deviation of .69. The mean originality weight of Form B is 34% higher than that of Form A. An increase in score on Form B of 34% could actually mean no increase at all. Therefore, caution will need to be exercised in the interpretation of verbal originality scores during certain of the analyses.

The Torrance Tests of Creative Thinking were chosen for use in this study because they had been used in many of the previous studies with bilingual children. The choice to use the Incomplete Figures and
Unusual Uses subtests was made in an effort to maximize the content and construct validity.

The Norms-Technical Manual (Torrance, 1974) reports many studies validating the use of the TTCT for measuring creativity. The manual also reports reliability data for the test. These test-retest reliability studies used subjects of various ages, various batteries of subtests, sometimes using both Form A and Form B, sometimes not, and with various amounts of time between tests. The reliability coefficients reported ranged from .34 to .97. When 118 sixth graders took alternate forms of the test (AB, BA) one to two weeks apart, the reliability coefficients ranged from .71 to .93. The only single subtest reliability study applicable to the present study had 22 college seniors take the Unusual Uses of Tin Cans test twice with 10 weeks between testings. The reliabilities were .75, .60, and .64 for fluency, flexibility, and originality respectively. Torrance recommends caution in interpreting reliability coefficients since creativity is influenced by emotions, motivation, and learning experiences.

Reliability coefficients obtained in the present research are given below. They are based upon the 51 subjects taking Form A followed 8 to 9 weeks later by Form B of the two subtests.
Table 1

Reliability Coefficients for Torrance Tests of Creative Thinking, Form A with Form B

<table>
<thead>
<tr>
<th>Measures</th>
<th>Incomplete Figures</th>
<th>Unusual Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>.53</td>
<td>.79</td>
</tr>
<tr>
<td>Flexibility (categories)</td>
<td>.48</td>
<td>.64</td>
</tr>
<tr>
<td>Flexibility (shift)</td>
<td>.51</td>
<td>.77</td>
</tr>
<tr>
<td>Originality</td>
<td>.46</td>
<td>.70</td>
</tr>
<tr>
<td>Elaboration</td>
<td>.62</td>
<td>---</td>
</tr>
</tbody>
</table>

Interrater reliabilities are generally quite high for the TTCT. Torrance (1974) reports reliability coefficients ranging from .80 to .99 for both experienced and inexperienced scorers. In the present study interrater reliability coefficients were computed for a random sample of 20% of the subjects. The second scorer was a doctoral candidate in Foreign Language Education at The Ohio State University. The reliability coefficients are presented below in Table 2.
Table 2
Interrater Reliability Coefficients for Torrance Tests of Creative Thinking

<table>
<thead>
<tr>
<th>Measures</th>
<th>Incomplete Figures</th>
<th>Unusual Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Form A</td>
<td>Form B</td>
</tr>
<tr>
<td>Fluency</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Flexibility (categories)</td>
<td>.90</td>
<td>.90</td>
</tr>
<tr>
<td>Flexibility (shifts)</td>
<td>.99</td>
<td>.96</td>
</tr>
<tr>
<td>Originality</td>
<td>.98</td>
<td>.92</td>
</tr>
<tr>
<td>Elaboration</td>
<td>.98</td>
<td>.96</td>
</tr>
</tbody>
</table>

A third measure of cognitive flexibility was obtained by use of the Functionally Remote Associates Test (FRAT) (Worthen and Clark, 1971), an adaptation of Mednick's Remote Associates Test (RAT) (Mednick, 1962). The FRAT presents subjects with items consisting of three words. The subject must produce a fourth word that is related to each of the other words. In order to answer the item correctly, the subject must go through many words associated with each of the given words until he/she finds the word that is associated with all three words. Mednick reports that some people cannot think of many words beyond their first strong association. These people are less likely to be able to find the word associated with all three words. The person more flexible in his or her thinking can produce many words associated with each of the given words and, thus, is likely to find the missing word.
The FRAT consists of 38 items arranged in order of increasing difficulty, each with one correct response. Forty minutes is allowed to complete the test. Since two versions of the test would be necessary for the present study, it was decided to divide the test into two 19-item, 20-minute subtests. The odd-numbered items became the new Short Form A, and the even-numbered items became Short Form B.

To check the validity of such a division of the test, the entire 38-item test was taken by 13 students of beginning Spanish at Ohio Dominican College. A split-half, odd-even reliability coefficient was computed and found to be .94. This coefficient was considered to be high enough to merit division of the test in this manner. Short Form A was given at the beginning of the quarter, Short Form B at the end.

KR-20 reliability coefficients obtained for the FRAT in the main study were as follows: Short Form A, r=.64; Short Form B, r=.77. When the Spearman-Brown correction is used to estimate the reliability of 38-item tests, the coefficients are: Form A, r=.78; Form B, r=.87. These results compare favorably with those reported by Worthen and Clark (1971) for the FRAT as a whole. Their KR-21 estimates ranged from .68 to .76.

After preliminary data analysis, it was decided not to use the FRAT in the subsequent analyses. The FRAT proved to be significantly correlated with ACTENG: PREFRAT, r=.4755, p < .001; POSTFRAT, r=.3846, p < .005. It was also nearly significantly correlated with ACTCOMP: PREFRAT, r=.2574, p < .068; POSTFRAT, r=.2682, p < .057. None of the 13 TTCT measures correlated with the ACT scores except the post verbal ratio shift (VERRAT2S) with ACTENG, r=.3057, p < .029. Furthermore,
the PREFRAT was not significantly correlated with any of the other variables except POSTFRAT, \( r = .4892, p < .001 \). The POSTFRAT was correlated significantly only with PREFRAT, pre verbal originality (VERORIG1) \( (r = .3696, p < .008) \), and pre figural elaboration (FIGELAB1) \( (r = .2803, p < .046) \). Thus, in this study, the FRAT appears more likely to be a measure of scholastic aptitude (more precisely, of verbal aptitude) than a measure of flexibility.

This finding is in keeping with the results of a factor analysis study by Cropley (1966). He administered two kinds of tests to seventh grade students in Canada: convergent tests such as IQ tests and vocabulary tests, and divergent or creativity tests such as the Torrance tin can Uses and the Mednick RAT. With both orthogonal and oblique rotation, the RAT consistently loaded on the convergent factor, whereas tin can Uses loaded strongly on the divergent factor. This finding prompted Cropley to cite Ketcham and Kheiralla (1962) who also found that the RAT correlated significantly with measures of achievement and to cast doubt on the usefulness of the RAT as a measure of creativity.

The high correlation of the RAT with intelligence was one of the factors that led Worthen and Clark (1971) to develop the FRAT. They reported generally lower correlations between the FRAT and a quick word test intelligence measure \( (r = .26 - .36) \) than between the RAT and the quick word test \( (r = .33 - .59) \). Nevertheless, in the present study, the correlations between the FRAT and the ACT English scores were significant. In addition, several subjects had difficulty with the vocabulary of the FRAT. The word "mezzanine" was especially difficult for them. "Petite" was also difficult.
Procedures

First Phase

A random sample of 306 students was chosen from the class rosters of the Colleges of Arts and Sciences sections of the University Survey Course (UVC 100.11) for Autumn quarter 1978. The following data were collected for each of the students from the University College files: high school GPA, ACT English and Composite scores, the number of years of study of a second language in high school, the language studied, whether or not the student completed a 101-level second-language course Autumn quarter 1978, GPA at end of Autumn 1978. Students were eliminated from the study if: Autumn 1978 was not their first quarter in a college or university, they were non-native speakers of English, they had credit by examination for any second-language course, they were enrolled in any second-language course beyond the 101-level, or were missing any of the data collected. After elimination of ineligible subjects, 137 remained.

The subjects were then divided into groups according to their high school and Autumn quarter second-language learning history. Specifically, they were divided into four groups defined by the independent variables: high school language learning and course 101 Autumn. Each of these variables had two levels: enrollment and non-enrollment.

Second Phase

During the first week of Autumn quarter 1979, the researcher visited five ASC sections of the University Survey course lecture to
explain the research project to the students and to ask for volunteers to participate in it. Volunteers had a choice of participating on either of two days during the following week at the same time as their lecture class. It was explained that to qualify as a participant students had to be native speakers of English who were not fluent in another language and who had not lived an extended period of time in a country where a language other than English was spoken. They had to be beginning their first quarter of university study, and they either had to be enrolled that quarter in a 101-level second-language course or in no second-language course at all. They also had to have taken the ACT college entrance tests. Furthermore, they were told that, while it didn't matter whether or not they had studied a second language in high school, students who had not studied another language before were especially encouraged to volunteer, particularly if they were enrolled in a 101-level language course that quarter. It was explained that when the volunteers met with the researcher they would be administered some tests of cognitive flexibility and be asked to fill out a brief questionnaire stating whether or not they were enrolled in a 101-level second-language course that quarter and whether or not they had studied a second language in high school. During the last week of the quarter the volunteers would be asked to return to take another version of the flexibility tests and to fill out another questionnaire confirming their enrollment status in 101-level second-language courses that quarter. Students were assured complete anonymity in any presentation or discussion of the study, and all volunteers would be asked to sign a consent form agreeing to participate in the study.
In response to this request, only 12 subjects volunteered to participate in the study. The researcher, therefore, returned the following week to the recitation sections of the course to solicit more help. As a result of the second request, 49 more subjects volunteered, making a total of 61 subjects. Of these subjects, 10 did not return at the end of the quarter to be retested. Thus, a total of 51 subjects qualified for inclusion in the study.

The subjects met with the researcher to be tested for the first time either the second or third week of an 11-week quarter. Participants were handed a booklet of materials with a consent form on top, which each subject read, signed, and dated. The instructions for the FRAT were read, and 20 minutes were allowed to complete the 19 items. Subjects were instructed not to go on in the booklet if they finished early. When time was called, the instructions for the Incomplete Figures test of the TTCT were read and 10 minutes were allowed to finish the 10 items. Next, the instructions for the Unusual Uses test of the TTCT were read and 10 minutes were allowed to work on the activity. When the tests were completed, subjects were asked to complete the questionnaire as fully as they could. The entire testing period lasted 48 minutes: one university class period.

During the tenth week of the quarter, all subjects who could be reached were called by the researcher to remind them of the retesting the following week. The retesting session followed the same procedure as the original testing session. That is, subjects completed the FRAT (20 minutes), Incomplete Figures (10 minutes), and Unusual Uses (10 minutes) with instructions read before each test. Form B of the tests
was given. Lastly, subjects filled out a second questionnaire confirming their enrollment status in 101-level courses and giving their opinions on the effects of second-language study on the mind.

**Analysis of the Data**

The analysis of the data attempted to answer three major questions. They are as follows:

1. Are the 1979 volunteer subjects from the same population as the 1978 randomly selected subjects?

2. Are the subjects in each of the four groups of each year's sample all from the same population?

3. After treatment, are the subjects in each of the four 1979 groups from the same population?

Each of the major questions subsumes several other questions, the answers to which help answer the major questions. These subquestions and the analyses used to answer them are listed below, grouped by major question.

1. A. Does the same proportion of subjects fall into each of the four groups each year?

   A Chi Square goodness of fit test with the 1978 proportions as the expected proportions and the 1979 proportions as observed proportions was employed.

   B. Do the subjects in the 1979 groups differ significantly from the subjects in the 1978 groups in (a) their HSGPAs and ACT scores, (b) their number of years of high school second-language study, (c) their GPAs at the end of their first quarter in the university?

      (a) A 3-way Manova with YEAR (1978, 1979), HSLANG (yes, no) and Course 101 (CRSE101) (yes, no) as independent variables and HSGPA and ACT scores as dependent variables. YEAR is the variable under study. Univariate F tests and Discriminant Analysis as follow ups.
(b) A 3-way Anova as above with years of high school language study (YRSLANG) as the dependent variable, followed up as in (a).

(c) A 3-way Anova with GPAEND as the dependent variable as above. This is a post-treatment measure.

2. A. Do the subjects in each of the four groups of 1978 differ significantly from each other in their HSGPAs and ACT scores?

   A 2-way Manova with HSLANG (yes, no) and CRSE101 (yes, no) as independent variables and HSGPA, ACTENG, and ACTCOMP as dependent variables. Univariate F tests and Discriminant Analysis as follow ups.

   B. Same as above for 1979 data.

   C. Same as B except Pre Flexibility measures as dependent variables.

3. A. Do the subjects in each of the four groups of 1979 differ significantly from each other after treatment in their Post Flexibility measure scores?

   A 2-way Manova with HSLANG and CRSE101 as independent variables and Post Flexibility measures as dependent variables. Univariate F tests and Discriminant Analysis as follow ups.

   B. Do the subjects in each of the four groups of 1979 differ significantly from each other after treatment in their Post Flexibility measure scores when the Pre Flexibility score for each measure is used as a covariate?

      Univariate 2-way Ancovas with HSLANG and CRSE101 as independent variables on each of the Flexibility measures covarying the Pre measure on the Post measure.

   C. On which Flexibility measures are the four groups most differentiated?

      A Discriminant Analysis on the Pre and Post Flexibility measures.
IV. ANALYSIS AND RESULTS

Introduction

As described at the end of Chapter III, the analysis of data attempted to answer the following questions:

1. Are the 1979 volunteer subjects from the same population as the 1978 randomly selected subjects?

2. For each year's sample, are the subjects in each of the four groups defined by the independent variables all from the same population?

3. After treatment are the subjects in each of the four 1979 groups from the same population?

This chapter will present the results of the analyses performed to answer each of the questions.

Question 1.

Two different analyses were used to help answer the question: Are the 1979 volunteer subjects from the same population as the 1978 randomly selected subjects? The first analysis attempted to determine whether or not the same proportion of students fell into each of the four groups defined by the independent variables each year. These four groups are (1) subjects who studied a second language in high school and a 101-level course Autumn quarter (yes, yes) (2) subjects who studied a second language in high school but did not take a 101-level course Autumn quarter (yes, no) (3) students who did not study a second
language in high school but did take a 101-level course Autumn quarter
(no, yes) (4) subjects who did not study a second language in high
school and did not take a 101-level course Autumn quarter (no, no).

To determine whether or not the proportions in each group were
the same from year to year, a Chi Square goodness of fit test was
performed on the data with the 1978 group proportions as the expected
proportions and the 1979 group proportions as the observed proportions.
These proportions are presented in Table 3.

Table 3
Proportions of Subjects in Each Group: 1978 and 1979
With Sample Size in Parentheses

<table>
<thead>
<tr>
<th>Group</th>
<th>1979</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (yes, yes)</td>
<td>.29</td>
<td>.23</td>
</tr>
<tr>
<td>2 (yes, no)</td>
<td>.49</td>
<td>.57</td>
</tr>
<tr>
<td>3 (no, yes)</td>
<td>.08</td>
<td>.04</td>
</tr>
<tr>
<td>4 (no, no)</td>
<td>.14</td>
<td>.16</td>
</tr>
</tbody>
</table>

Computation yielded a Chi Square value of 6.9380 which, when
compared to the tabled critical value at the .05 level of significance
on 3 degrees of freedom of 7.815 proved to be not significant. When
viewed as four separate groups, the proportion of students falling into
each of them in 1979 was not significantly different from the proportion
in each of the four groups of 1978.

For further comparison, the groups were collapsed over each of
the variables to produce two other sets of proportions each with two
groups of subjects. First, the groups were collapsed over the CRSE101
variable so that there were two groups: subjects who studied a second language in high school and subjects who did not. Secondly, the four groups were collapsed over the HSLANG variable so that there were two groups: subjects who took a 101-level course Autumn quarter and subjects who did not. These two sets of group proportions are presented in Table 4.

Table 4
Proportions of Subjects in Collapsed Groups: 1978 and 1979
With Sample Size in Parentheses

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 2</td>
<td>.78 (40)</td>
<td>.80 (109)</td>
<td>1 &amp; 3</td>
<td>.37 (19)</td>
<td>.27 (37)</td>
</tr>
<tr>
<td>(HS yes)</td>
<td></td>
<td></td>
<td>(101 yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td>.22 (11)</td>
<td>.20 (28)</td>
<td>2 &amp; 4</td>
<td>.63 (32)</td>
<td>.73 (100)</td>
</tr>
<tr>
<td>(HS no)</td>
<td></td>
<td></td>
<td>(101 no)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When the groups are collapsed over the CRSE101 variable so that there are two groups: those who studied a second language in high school and those who did not, the computed Chi Square value is .25. This value is non-significant when compared to the tabled critical value at the .05 level of significance on 1 degree of freedom of 3.841. Thus, each year approximately 80% of the subjects had studied a second language in high school.

The computed Chi Square value for the groups collapsed over the HSLANG variable was 5.07 which proved to be significant at the .05 level on 1 degree of freedom (critical value: 3.841). A significantly larger proportion of the subjects studied a 101-level second-language course in 1979 than did in 1978. This result is important because it
is the first indication that the volunteer group was not from exactly the same population as the 1978 randomly selected group. The volunteer group had a higher proportion of students enrolled in a 101-level course than did the random group. Since the study centered on the effect of second-language study, perhaps students who were enrolled in language courses were more interested in participating. Likewise, the same motivation that students had for not enrolling in a 101-level course may also have kept them from volunteering to participate in the study.

The second part of the analysis to answer Question 1 aimed to determine whether or not the 1979 groups differed significantly from the 1978 groups in their HSGPAs, their ACT scores, the number of years of high school language study they had had, and their GPAs at the end of Autumn quarter. As outlined in Chapter III, three separate Analyses of Variance were undertaken for this purpose. All three were 3-way analyses with YEAR, HSLANG, and CRSE101 as independent variables. The first was a multivariate analysis with HSGPA, ACTENG, and ACTCOMP as dependent variables. The second analysis had YRSLANG as the dependent variable, and the third analysis had GPAEND as the dependent variable. In these three analyses, YEAR was the independent variable being scrutinized. If none of the effects including YEAR was significant, the 1978 and 1979 groups could not be shown to have come from different populations, and an analysis of the 1979 data could be justified, bearing in mind, of course, that there were significantly more subjects in 101-level courses in 1979 than there had been in 1978. Means and standard deviations relevant to the analyses are presented in Table 5.
Table 5

Means and Standard Deviations of GPAs, ACT Scores and YRSLANG by YEAR, HSLANG, and CRSE101

<table>
<thead>
<tr>
<th>YEAR:</th>
<th>1978</th>
<th>1979</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSLANG:</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Variable</td>
<td>CRSE101:</td>
<td>YES</td>
</tr>
<tr>
<td>HSGPA</td>
<td>M</td>
<td>3.03</td>
</tr>
<tr>
<td>SD</td>
<td>.54</td>
<td>.65</td>
</tr>
<tr>
<td>ACTENG</td>
<td>M</td>
<td>20.06</td>
</tr>
<tr>
<td>SD</td>
<td>3.60</td>
<td>4.88</td>
</tr>
<tr>
<td>SD</td>
<td>4.14</td>
<td>5.35</td>
</tr>
<tr>
<td>YRSLANG</td>
<td>M</td>
<td>2.48</td>
</tr>
<tr>
<td>SD</td>
<td>.72</td>
<td>.82</td>
</tr>
<tr>
<td>GPAEND</td>
<td>M</td>
<td>2.56</td>
</tr>
<tr>
<td>SD</td>
<td>.59</td>
<td>.72</td>
</tr>
</tbody>
</table>

\( ^a \)HSGPA = High school grade-point average
ACTENG = ACT English score
ACTCOMP = ACT Composite score
YRSLANG = Number of years of high school second-language study
GPAEND = Grade-point average at end of Autumn quarter
The 3-way Manova on HSGPA and ACT scores was undertaken to determine whether or not the subjects from 1978 and 1979 differed significantly in their academic ability upon entering the university. A summary table of the Manova with univariate F's included is presented in Table 6.

As seen in the table, none of the effects with the YEAR variable was significant. Thus, the 1979 subjects cannot be shown to be significantly different from the 1978 subjects in their academic ability upon entering the university. There are, however, significant multivariate effects in the analysis. These effects apply to the subjects of both years taken as one pooled group. The discussion of these effects does not apply directly to the purpose of the present study which includes the 1978 subjects in order to validate the analysis of the data collected in 1979 on volunteer subjects. Nevertheless, the analysis of these pooled effects provides an overview of the academic ability of freshman students at OSU on a larger scale.

As seen in Table 6, there was a significant first-order interaction between HSLANG and CRSE101. The univariate F's showed the same effect for the ACT scores. A discriminant analysis of the data showed group 4 (no, no) to be significantly different and lower in its scores than the other groups. One significant discriminant function was produced. It consisted of only the ACTENG variable and accounted for 84% of the overall variance.

There was also a significant main effect for HSLANG multivariately and in all three univariate tests. This effect was due largely to group four's low scores, but group three's scores were low enough,
Table 6

Values of F for Manova of HSGPA and ACT Score by YEAR, HSLANG, CRSE101 with Univariate F Values

<table>
<thead>
<tr>
<th>Source</th>
<th>Univariate Fs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Manova</td>
</tr>
<tr>
<td>ABC(^\text{a})</td>
<td>.478</td>
</tr>
<tr>
<td>BC</td>
<td>5.029^{***}</td>
</tr>
<tr>
<td>AC</td>
<td>1.623</td>
</tr>
<tr>
<td>AB</td>
<td>.614</td>
</tr>
<tr>
<td>C</td>
<td>1.449</td>
</tr>
<tr>
<td>B</td>
<td>8.069^{***}</td>
</tr>
<tr>
<td>A</td>
<td>1.316</td>
</tr>
</tbody>
</table>

\(^\text{a}\)A: YEAR, B: HSLANG, C: CRSE101

\(^*\) \(p < .05\)
\(^{**}\) \(p < .01\)
\(^{***}\) \(p < .001\)
especially on HSGPA, to contribute to the main effect. Thus, overall, students who had studied a second language in high school were higher in academic ability than those who had not.

The discriminant functions correctly categorized only 40% of the subjects, but the success rate varied greatly by group. Only 6% of those in group one were correctly placed into group one. Fifty percent of group two, 30% of group three, and 62% of group four were correctly categorized. Thus, group four seems to be a more distinct group than the others, distinguished by its lower scores.

When discriminant analyses of the GPA and ACT data were performed for each of the years separately, the groups could be compared more closely. In the 1979 analysis, group four (no, no) was significantly lower than the others in scores, and group three (no, yes) was higher, but not significantly so. Forty-five percent of the subjects were correctly classified into their groups: 20% of group one, 52% of group two, 50% of group three, and 71% of group four. Thus, group four was more easily discriminated because of its low scores. Group one was most often miscategorized into group two. Forty-two percent of the 1978 subjects were correctly classified by the 1979 functions. The percentages by group matched those of 1979 except for that of group three. Twenty-six percent of group one was correctly classed, 44% of group two, only 17% of group three, and 64% of group four. This suggests that although the subjects did not prove to be significantly different on the YEAR variable in the Manova, the subjects of group three in 1979 were not the same as those in 1978.
The 1978 discriminant analysis seemed to bear this out. Again, group four was lower than the rest in its scores, but not lower than group three in HSGPA or ACTCOMP. Forty-three percent of the subjects were correctly classified: 13% of group one, 56% of group two, 17% of group three, and 50% of group four. Subjects in group one were often placed in group two; those in group three were placed in group four. Of the 1979 subjects, 41% were correctly classified by the 1978 functions: 0% of group one, 60% of group two, 50% of group three, and 57% of group four. The 50% for group three is deceiving since the total group consisted of only four subjects. The other two were placed into group two whereas the 1978 group three subjects were often placed into group four.

Overall, then, the 1979 groups seem to be slightly more differentiated than the 1978 groups. In both years, group four (no, no) had lower scores than the other groups, and in 1979, group three (no, yes) had higher scores than it did in 1978. Except for group four, none of the groups was highly differentiated and, therefore, they could all be from the same population.

The second analysis undertaken to answer part two of Question 1 was a 3-way Anova on YRSLANG. The purpose of this analysis was to determine whether or not the subjects in the 1979 groups had had the same number of years of high school language study as had those in the 1978 groups. Here it needs to be noted that in order to maximize the size of the groups with no high school study of second languages, any student with one year or less of study was included in the no high school second-language group. In general, high school second-language
classes present the structures of the language at a very slow pace. Thus, only a small portion of the second language is learned in one year's time. Many students begin high school second-language study in grade nine and very often complete two years of study. Of course, even a small amount of second-language learning should have some effect on the individual, but the effect of one year of study four years earlier would likely be minimal. Students with one year or less of second-language study were felt to be more like those without second-language study than like those with two years or more of study.

The summary table of the 3-way Anova on YRSLANG is presented in Table 7. There were no significant effects involving the YEAR variable. Thus, the groups vary the same way each year with regard to the number of years of high school language study that subjects had had. The main effect for HSLANG was obviously significant since that variable was defined by the subjects' number of years study of second languages in high school. There were no other significant differences.

The third analysis undertaken to help answer the second part of Question 1 was a 3-way Anova on subjects' GPAs at the end of Autumn quarter. This was a post-treatment measure, included to determine whether or not the testing experience of the 1979 groups influenced them with regard to their academic achievement in the university. If it did, there should have been an effect in the YEAR variable. A summary table of the Anova is presented in Table 8. There were no significant effects involving the YEAR variable, hence, it cannot be shown that the 1979 groups behaved differently during their first quarter in the university than did the 1978 groups.
Table 7
Anova of YRSLANG by YEAR, HSLANG, and CRSE101

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC(^a)</td>
<td>1</td>
<td>.093</td>
<td>.170</td>
</tr>
<tr>
<td>BC</td>
<td>1</td>
<td>.054</td>
<td>.099</td>
</tr>
<tr>
<td>AC</td>
<td>1</td>
<td>.532</td>
<td>.970</td>
</tr>
<tr>
<td>AB</td>
<td>1</td>
<td>.038</td>
<td>.068</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>.224</td>
<td>.408</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>98.452</td>
<td>179.509***</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>.092</td>
<td>.168</td>
</tr>
<tr>
<td>Residual</td>
<td>180</td>
<td>.548</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>187</td>
<td>1.279</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) A: YEAR, B: HSLANG, C: CRSE101

*** \(p < .001\)
Table 8
Anova of GPAEND by YEAR, HSLANG, and CRSE101

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCa</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>BC</td>
<td>1</td>
<td>.144</td>
<td>.253</td>
</tr>
<tr>
<td>AC</td>
<td>1</td>
<td>1.366</td>
<td>2.394</td>
</tr>
<tr>
<td>AB</td>
<td>1</td>
<td>.476</td>
<td>.833</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>.052</td>
<td>.092</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>6.926</td>
<td>12.137***</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1.042</td>
<td>1.826</td>
</tr>
<tr>
<td>Residual</td>
<td>180</td>
<td>.571</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>187</td>
<td>.641</td>
<td></td>
</tr>
</tbody>
</table>

*a: YEAR, B: HSLANG, C: CRSE101
*** P < .001
The only significant result was the main effect for HSLANG. Examination of the means showed that subjects who had studied a second language in high school achieved higher GPAs at the end of one quarter in the university than did those who had not studied languages. This could indicate that the academically better students chose to study a second language in high school (these students also had higher HSGPAs and ACT scores), or that the study of a second language helps improve academic ability in some way, perhaps in study skills. This question could not be answered by this study.

In summary, there were very few differences between the subjects in year 1978 and those in year 1979. The same proportion of the total group of each year fell into each of the four groups. Only when they were divided as to whether or not they were enrolled in a 101-level course was there a significant difference. More students in 1979 were in a 101 course than there were in 1978. This may have been due to the nature of the study and its effect on motivation to volunteer.

There were no significant differences between the 1978 and 1979 groups on their HSGPAs, ACT scores, years of high school language study, or GPAs at the end of Autumn quarter. The four groups varied on these measures in the same manner in both years. The only indication that this statement may not be true came from the comparison of the group classifications of the discriminant analyses which suggested that group three in 1979 differed from group three in 1978. This difference was not found to be significant in the Manova, although the small sample sizes may have prevented its detection. Nevertheless, from the analyses performed here, the 1979 volunteers cannot be said to be a
sample from a different population than the 1978 random sample. They may, of course, differ in ways not measured or not detected here.

**Question 2.**

The second major question sought to determine whether or not the four groups could be said to all be from the same population within each year. The discussion of Question 1 indicated that the groups varied in the same ways from year to year. This question sought to determine in what ways the groups differed from each other. For the 1979 groups this question was important in establishing equality of the groups before treatment in their academic ability and in their flexibility scores. The 1978 groups were also analysed to provide further confirmation that the two years varied in the same ways. There were three analyses performed to answer Question 2: A 2-way Manova on HSGPAs and ACT scores of the 1978 subjects; a 2-way Manova on HSGPAs and ACT scores of the 1979 subjects; and a 2-way Manova on the Pre Flexibility measures of the 1979 subjects.

A summary table of the multivariate and univariate $F$ values for the 2-way Manova on 1978 HSGPAs and ACT scores is presented in Table 9. Although the multivariate first-order interaction was not significant, the univariate interaction for ACTENG was. Also significant were the multivariate and all univariate main effects for HSLANG.

The discriminant analysis showed that ACTENG was the best discriminator with group four (no, no) having significantly lower scores than all other groups. This accounted for the significant interaction
Table 9

Values of $F$ for Manova of 1978 HSGPA and ACT Scores by HSLANG and CRSE101 with Univariate $F$ Values

<table>
<thead>
<tr>
<th>Source</th>
<th>Manova</th>
<th>HSGPA</th>
<th>ACTENG</th>
<th>ACTCOMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB$^a$</td>
<td>1.949</td>
<td>.962</td>
<td>5.435*</td>
<td>2.033</td>
</tr>
<tr>
<td>B</td>
<td>.387</td>
<td>.380</td>
<td>.056</td>
<td>.711</td>
</tr>
</tbody>
</table>

$^a$A: HSLANG, B: CRSE101

* $p < .05$

*** $p < .001$
on ACTENG. When the other variables were included in the analysis, group three (no, yes) ceased to be differentiated from group four. Thus, the interaction was not significant for those variables. Since groups three and four were both no high school language groups, their lack of differentiation on HSGPA and ACTCOMP was enough to produce a significant HSLANG main effect. Although group four had ACTENG scores significantly lower than those of group three, their combined mean was low enough to produce the highly significant main effect. In 1978, then, students who had not studied a second language in high school had significantly lower HSGPAs and ACT scores than students who had studied a second language.

A summary table of the multivariate and univariate F values for the 2-way Manova on 1979 HSGPAs and ACT scores is presented in Table 10. The multivariate first-order interaction proved to be significant. Inspection of the univariate F's showed that the interaction for ACTENG was significant. Discriminant analysis indicated that ACTENG was the best discriminator showing group four (no, no) to have significantly lower scores than all other groups. No other results were significant. Group three had higher scores than groups one and two, but not significantly so. These higher scores precluded the HSLANG main effect that was seen in 1978. They also suggest that the 1979 group three was composed of a different kind of student than the 1978 group three, though this could not be directly shown statistically. In 1979, the only observed significant difference among the four groups in their academic ability upon entering the university was that students who did not study a second language in high school and, likewise, chose not
Table 10

Values of $F$ for Manova of 1979 HSGPA and ACT Scores by HSLANG and CRSE101 with Univariate $F$ Values

<table>
<thead>
<tr>
<th>Source</th>
<th>Manova</th>
<th>HSGPA</th>
<th>ACTENG</th>
<th>ACTCOMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>3.891*</td>
<td>.935</td>
<td>10.093</td>
<td>3.213</td>
</tr>
<tr>
<td>B</td>
<td>2.486</td>
<td>3.510</td>
<td>2.187</td>
<td>.392</td>
</tr>
<tr>
<td>A</td>
<td>1.724</td>
<td>2.995</td>
<td>2.445</td>
<td>.633</td>
</tr>
</tbody>
</table>

$^a$A: HSLANG, B: CRSE101

* $P < .05$

*** $P < .001$
to enroll in a 101-level second-language course Autumn quarter had lower scores than all other students on the ACTENG test.

The third analysis performed in answer to Question 2 was a 2-way Manova on the Pre Flexibility scores of the 1979 subjects. The purpose of the analysis was to determine whether or not the subjects in each of the four groups could be considered to have come from the same population with regard to their cognitive flexibility upon entering the university. Relevant means and standard deviations for the Pre Flexibility data are presented in Table 11.

A summary table of the multivariate and univariate F values for the 2-way Manova on Pre Flexibility scores is presented in Table 12. The analysis did not produce any significant multivariate effects. Thus, the groups cannot be shown to be from different populations on the basis of their cognitive flexibility upon entering the university. All univariate interactions except the ratios were, however, significant. Discriminant analysis showed that FIGORIG was the best discriminator, distinguishing among the groups better by itself than in combination with any other variables. It discriminated among all the groups except group two from group three and group one from group four. Groups two and three had significantly higher means than groups one and four. FIGFLEX also discriminated among the groups significantly by itself.

Thus, while multivariately there were no significant differences among the four groups on the Pre Flexibility measures, univariate tests indicate that groups two and three scored significantly higher than groups one and four on all measures but the ratios.
### Table 11

Means and Standard Deviations of Pre Flexibility Scores by HSLANG and CRSE101

<table>
<thead>
<tr>
<th>Variable</th>
<th>HSLANG:</th>
<th></th>
<th>CRSE101:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>FIGFLU</td>
<td>M</td>
<td>5.933</td>
<td>7.080</td>
<td>8.500</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.154</td>
<td>2.597</td>
<td>2.380</td>
</tr>
<tr>
<td>FIGFLEX</td>
<td>M</td>
<td>5.200</td>
<td>6.280</td>
<td>8.000</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.781</td>
<td>2.112</td>
<td>2.000</td>
</tr>
<tr>
<td>FIGRAT</td>
<td>M</td>
<td>.890</td>
<td>.906</td>
<td>.950</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.121</td>
<td>.109</td>
<td>.058</td>
</tr>
<tr>
<td>FIGFLEX S</td>
<td>M</td>
<td>4.733</td>
<td>5.840</td>
<td>7.500</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.120</td>
<td>2.511</td>
<td>2.380</td>
</tr>
<tr>
<td>FIGRAT S</td>
<td>M</td>
<td>.962</td>
<td>.962</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.082</td>
<td>.079</td>
<td>.000</td>
</tr>
<tr>
<td>FIGORIG</td>
<td>M</td>
<td>5.867</td>
<td>8.840</td>
<td>10.500</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.295</td>
<td>3.325</td>
<td>3.873</td>
</tr>
<tr>
<td>FIGELAB</td>
<td>M</td>
<td>37.200</td>
<td>46.520</td>
<td>54.500</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>17.436</td>
<td>18.617</td>
<td>47.177</td>
</tr>
<tr>
<td>VERFLU</td>
<td>M</td>
<td>17.867</td>
<td>23.280</td>
<td>28.750</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>7.110</td>
<td>9.529</td>
<td>20.516</td>
</tr>
<tr>
<td>VERFLEX</td>
<td>M</td>
<td>10.533</td>
<td>12.120</td>
<td>13.500</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.997</td>
<td>2.934</td>
<td>7.550</td>
</tr>
<tr>
<td>VERRAT</td>
<td>M</td>
<td>.636</td>
<td>.570</td>
<td>.610</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.165</td>
<td>.147</td>
<td>.283</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>5.551</td>
<td>7.084</td>
<td>16.132</td>
</tr>
<tr>
<td>VERRAT S</td>
<td>M</td>
<td>.889</td>
<td>.854</td>
<td>.830</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.100</td>
<td>.149</td>
<td>.147</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>6.486</td>
<td>10.126</td>
<td>17.369</td>
</tr>
</tbody>
</table>
Table 12

Values of $F$ for Manova of 1979 Pre Flexibility Scores by HSLANG and CRSE101 with Univariate $F$ Values

**Manova**

<table>
<thead>
<tr>
<th>Source</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB$^a$</td>
<td>1.664</td>
</tr>
<tr>
<td>B</td>
<td>.531</td>
</tr>
<tr>
<td>A</td>
<td>.625</td>
</tr>
</tbody>
</table>

**Univariate Fs**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source:</th>
<th>AB</th>
<th>B</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGFLU</td>
<td></td>
<td>5.184*</td>
<td>.245</td>
<td>.044</td>
</tr>
<tr>
<td>FIGFLEX</td>
<td></td>
<td>8.729*</td>
<td>.136</td>
<td>.105</td>
</tr>
<tr>
<td>FIGRAT</td>
<td></td>
<td>1.870</td>
<td>.046</td>
<td>.047</td>
</tr>
<tr>
<td>FIGFLEX S</td>
<td></td>
<td>6.596*</td>
<td>.104</td>
<td>.024</td>
</tr>
<tr>
<td>FIGRAT S</td>
<td></td>
<td>2.322</td>
<td>.604</td>
<td>.438</td>
</tr>
<tr>
<td>FIGORIG</td>
<td></td>
<td>12.548***</td>
<td>2.337</td>
<td>.030</td>
</tr>
<tr>
<td>FIGELAB</td>
<td></td>
<td>4.129*</td>
<td>.256</td>
<td>.039</td>
</tr>
<tr>
<td>VERFLU</td>
<td></td>
<td>5.661*</td>
<td>.428</td>
<td>.013</td>
</tr>
<tr>
<td>VERFLEX</td>
<td></td>
<td>5.389*</td>
<td>.176</td>
<td>.210</td>
</tr>
<tr>
<td>VERRAT</td>
<td></td>
<td>.357</td>
<td>.957</td>
<td>.117</td>
</tr>
<tr>
<td>VERFLEX S</td>
<td></td>
<td>4.822*</td>
<td>.276</td>
<td>.005</td>
</tr>
<tr>
<td>VERRAT S</td>
<td></td>
<td>.929</td>
<td>.160</td>
<td>.001</td>
</tr>
<tr>
<td>VERORIG</td>
<td></td>
<td>4.614*</td>
<td>.075</td>
<td>.000</td>
</tr>
</tbody>
</table>

$^a$ A: HSLANG, B: CRSE101

* $p < .05$

** $p < .01$

*** $p < .001$
In answer to Question 2, then, 1979 subjects who did not study a second language in high school and did not enroll in a 101-level course Autumn quarter had significantly lower ACTENG scores than those in the other groups. In addition, those same students and those who did study a second language in high school and did enroll in a 101-level course had significantly lower scores on all of the Pre Flexibility measures except the ratios, although multivariately the difference was not significant.

**Question 3.**

Three analyses were conducted to help answer Question 3 which asked whether or not the four groups of 1979 could be considered to belong to the same population after the treatment. The first analysis was a 2-way Manova on the Post Flexibility measures. The second was a series of Ancovas on each of the Flexibility measures covarying the Pre on the Post measures. The final analysis was an overall discriminant analysis with the Pre and Post Flexibility measures included. Relevant means and standard deviations for the Post Flexibility measures are presented in Table 13.

A summary table of multivariate and univariate F values for the Manova on Post Flexibility measures is presented in Table 14. There were no multivariate significant differences among groups on the Post Flexibility measures. Thus, it cannot be stated that the four groups do not come from the same population.

Although the multivariate first-order interaction was not significant, one of the univariate interactions was significant:
Table 13
Means and Standard Deviations of Post Flexibility Scores by HSLANG and CRSE101

<table>
<thead>
<tr>
<th>Variable</th>
<th>HSLANG: Yes</th>
<th>HSLANG: No</th>
<th>CRSE101: Yes</th>
<th>CRSE101: No</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGFLU</td>
<td>M 7.400</td>
<td>7.960</td>
<td>8.000</td>
<td>6.857</td>
</tr>
<tr>
<td></td>
<td>SD 1.920</td>
<td>1.968</td>
<td>2.828</td>
<td>2.734</td>
</tr>
<tr>
<td>FIGFLEX</td>
<td>M 6.533</td>
<td>7.120</td>
<td>7.000</td>
<td>5.571</td>
</tr>
<tr>
<td></td>
<td>SD 1.506</td>
<td>1.716</td>
<td>2.160</td>
<td>2.070</td>
</tr>
<tr>
<td>FIGRAT</td>
<td>M .898</td>
<td>.907</td>
<td>.900</td>
<td>.846</td>
</tr>
<tr>
<td></td>
<td>SD .104</td>
<td>.109</td>
<td>.141</td>
<td>.171</td>
</tr>
<tr>
<td>FIGFLEX S</td>
<td>M 6.200</td>
<td>6.840</td>
<td>6.750</td>
<td>5.714</td>
</tr>
<tr>
<td></td>
<td>SD 1.740</td>
<td>1.930</td>
<td>2.630</td>
<td>2.563</td>
</tr>
<tr>
<td>FIGRAT S</td>
<td>M .976</td>
<td>.984</td>
<td>.973</td>
<td>.984</td>
</tr>
<tr>
<td></td>
<td>SD .050</td>
<td>.044</td>
<td>.055</td>
<td>.042</td>
</tr>
<tr>
<td>FIGORIG</td>
<td>M 10.800</td>
<td>10.200</td>
<td>11.250</td>
<td>8.571</td>
</tr>
<tr>
<td></td>
<td>SD 3.028</td>
<td>3.329</td>
<td>5.909</td>
<td>4.353</td>
</tr>
<tr>
<td>FIGELAB</td>
<td>M 53.000</td>
<td>48.960</td>
<td>45.250</td>
<td>38.000</td>
</tr>
<tr>
<td></td>
<td>SD 21.341</td>
<td>18.390</td>
<td>30.511</td>
<td>19.706</td>
</tr>
<tr>
<td>VERFLU</td>
<td>M 17.467</td>
<td>20.800</td>
<td>25.500</td>
<td>17.000</td>
</tr>
<tr>
<td></td>
<td>SD 7.596</td>
<td>10.271</td>
<td>20.158</td>
<td>7.703</td>
</tr>
<tr>
<td>VERFLEX</td>
<td>M 10.333</td>
<td>11.960</td>
<td>12.000</td>
<td>8.857</td>
</tr>
<tr>
<td></td>
<td>SD 3.754</td>
<td>3.494</td>
<td>6.782</td>
<td>3.388</td>
</tr>
<tr>
<td>VERRAT</td>
<td>M .641</td>
<td>.622</td>
<td>.615</td>
<td>.581</td>
</tr>
<tr>
<td></td>
<td>SD .189</td>
<td>.140</td>
<td>.274</td>
<td>.195</td>
</tr>
<tr>
<td>VERFLEX S</td>
<td>M 14.333</td>
<td>17.800</td>
<td>21.250</td>
<td>13.000</td>
</tr>
<tr>
<td></td>
<td>SD 6.662</td>
<td>8.544</td>
<td>16.860</td>
<td>6.028</td>
</tr>
<tr>
<td>VERRAT S</td>
<td>M .882</td>
<td>.895</td>
<td>.905</td>
<td>.839</td>
</tr>
<tr>
<td></td>
<td>SD .117</td>
<td>.085</td>
<td>.071</td>
<td>.082</td>
</tr>
<tr>
<td>VERORIG</td>
<td>M 18.800</td>
<td>23.320</td>
<td>33.750</td>
<td>18.857</td>
</tr>
<tr>
<td></td>
<td>SD 10.489</td>
<td>9.357</td>
<td>28.987</td>
<td>9.045</td>
</tr>
</tbody>
</table>
Table 14

Values of $F$ for Manova of 1979 Post Flexibility Scores by HSLANG and CRSE101 with Univariate $F$ Values

### Manova

<table>
<thead>
<tr>
<th>Source</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB$^a$</td>
<td>1.226</td>
</tr>
<tr>
<td>B</td>
<td>.808</td>
</tr>
<tr>
<td>A</td>
<td>1.107</td>
</tr>
</tbody>
</table>

### Univariate Fs

<table>
<thead>
<tr>
<th>Source:</th>
<th>AB</th>
<th>B</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGFLU</td>
<td>1.278</td>
<td>.101</td>
<td>.433</td>
</tr>
<tr>
<td>FIGFLEX</td>
<td>2.689</td>
<td>.096</td>
<td>1.868</td>
</tr>
<tr>
<td>FIGRAT</td>
<td>.567</td>
<td>.016</td>
<td>.889</td>
</tr>
<tr>
<td>FIGFLEX S</td>
<td>1.377</td>
<td>.232</td>
<td>.548</td>
</tr>
<tr>
<td>FIGRAT S</td>
<td>.011</td>
<td>.459</td>
<td>.006</td>
</tr>
<tr>
<td>FIGORIG</td>
<td>.663</td>
<td>.996</td>
<td>.512</td>
</tr>
<tr>
<td>FIGELAB</td>
<td>.049</td>
<td>.638</td>
<td>2.001</td>
</tr>
<tr>
<td>VERFLU</td>
<td>2.681</td>
<td>.074</td>
<td>.024</td>
</tr>
<tr>
<td>VERFLEX</td>
<td>3.070</td>
<td>.297</td>
<td>1.060</td>
</tr>
<tr>
<td>VERRAT</td>
<td>.014</td>
<td>.192</td>
<td>.367</td>
</tr>
<tr>
<td>VERFLEX S</td>
<td>3.750</td>
<td>.151</td>
<td>.029</td>
</tr>
<tr>
<td>VERRAT S</td>
<td>1.413</td>
<td>.023</td>
<td>.721</td>
</tr>
<tr>
<td>VERORIG</td>
<td>5.329$^*$</td>
<td>.012</td>
<td>.427</td>
</tr>
</tbody>
</table>

$^a$: HSLANG, B: CRSE101

$^*$ $p < .05$
VERORIG, $p < .025$. Approaching significance were VERFLEX S, $p < .059$ and VERFLEX, $p < .086$. The discriminant analysis showed that none of the variables by itself distinguished among the groups significantly. VERORIG was the best discriminator by itself with $p < .1388$. It was the first included in the overall analysis. The discrimination was not significant until VERFLEX was also included in the analysis. These two variables together discriminated among the groups significantly with a $p < .0154$. The best discrimination occurred when VERRAT was also added to the other two. Then group three was shown to be significantly higher in scores than the other groups. When FIGELAB was included in the analysis the same was true.

None of the three generated functions was significant, the first accounting for 43% of the variance. It included VERORIG and VERFLU with Originality more heavily weighted than Fluency. Thus, since Originality is somewhat determined by Fluency, this function seems to be a measure of the ability to list many, unusual ideas. This is the dimension in which group three excelled.

The second function accounted for 32% of the variance and included all of the Flexibility measures, and, less strongly, FIGFLU and FIGRAT. This is obviously a measure of Flexibility. Group two scored somewhat higher than the others on this dimension while group four was lower.

The third function included FIGORIG and FIGELAB, and, less strongly, FIGRAT S and VERRAT. The function seems to be the figural counterpart of the first function: The ability to draw unusual pictures with many details. Since there were only 10 pictures to draw, many
subjects completed all 10. Thus, the Fluency aspect manifested itself more in Elaboration than in Fluency per se. This function accounted for 23% of the variance, and group four scored lower than the others while groups one and three scored higher on it.

It is interesting to note that whereas on the Pre tests the groups differed on both figural and verbal measures (with groups two and three high and groups one and four low), on the Post tests the groups differed only on the verbal Originality measure (again with group three high and groups one and four low).

It is also of note that on all the figural tests except Originality, groups one, two and four scored higher on the Post than on the Pre tests. Group three scored lower. On Originality, all four groups scored higher on the Post test. Since group three was much higher on the Pre tests than the other groups, this could be a result of regression toward the mean.

On the verbal tests the pattern is more difficult to see. On Originality all groups went up (perhaps a function of the scoring procedure). On Fluency and Flexibility, all went down--group three more than the others. On FLEX S all went down very slightly except group four which went up. On both ratios, all groups went up except group four which went down.

In order to understand better these changes in score and to determine whether or not other differences would result in the Post tests when Pre test differences were controlled for, a series of univariate Analyses of Covariance (Ancova) were performed covarying each of the Pre Flexibility measures on its Post measure. A summary table of the $F$ values for the Ancovas is presented in Table 15.
Table 15

Values of F for Ancovas Covarying Pre on Post Flexibility Measures

<table>
<thead>
<tr>
<th>Source</th>
<th>AB</th>
<th>B</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGFLU</td>
<td>.002</td>
<td>.007</td>
<td>.737</td>
</tr>
<tr>
<td>FIGFLEX</td>
<td>.103</td>
<td>.005</td>
<td>2.253</td>
</tr>
<tr>
<td>FIGRAT</td>
<td>.705</td>
<td>.362</td>
<td>.472</td>
</tr>
<tr>
<td>FIGFLEX S</td>
<td>.007</td>
<td>.121</td>
<td>.828</td>
</tr>
<tr>
<td>FIGRAT S</td>
<td>.013</td>
<td>.150</td>
<td>.020</td>
</tr>
<tr>
<td>FIGORIG</td>
<td>.997</td>
<td>1.033</td>
<td>.940</td>
</tr>
<tr>
<td>FIGELAB</td>
<td>2.021</td>
<td>.140</td>
<td>4.051*</td>
</tr>
<tr>
<td>VERFLU</td>
<td>.091</td>
<td>.014</td>
<td>.000</td>
</tr>
<tr>
<td>VERFLEX</td>
<td>.168</td>
<td>.002</td>
<td>.620</td>
</tr>
<tr>
<td>VERRAT</td>
<td>.418</td>
<td>.029</td>
<td>.730</td>
</tr>
<tr>
<td>VERFLEX S</td>
<td>.179</td>
<td>.066</td>
<td>.003</td>
</tr>
<tr>
<td>VERRAT S</td>
<td>2.113</td>
<td>.777</td>
<td>.189</td>
</tr>
<tr>
<td>VERORIG</td>
<td>1.201</td>
<td>.540</td>
<td>1.292</td>
</tr>
</tbody>
</table>

*a: HSLANG, B: CRSE101
*p < .05
The only significant effect in the set of Ancovas was the main effect for HSLANG on FIGELAB, \( p < .050 \). When Pre test results were controlled for, subjects with second-language study in high school elaborated significantly more in their figure completions than did those who had not studied a second language in high school. This result was due, in part, to the 16-point increase in mean Elaboration score for group one and the 9-point decrease in score for group three. Groups two and four increased slightly in their mean scores on the Post test from the Pre test. The verbal Originality interaction effect from the Post test analysis did not prove to be significant when Pre test scores were controlled for.

Finally, a discriminant analysis was performed with both the Pre and Post test scores as dependent measures. Recall, that individually, FIGORIG1 and FIGFLEX1 significantly discriminated among the groups. In this overall analysis, FIGORIG1 was the best discriminator, significantly distinguishing all of the groups except group two from group three and group one from group four with groups two and three having higher scores.

The order of inclusion in the analysis was interesting. Of the figural measures, 70% were included in the first half of the analysis. Thirty percent of the verbal measures were in the first half. It would seem, therefore, that the groups varied more on the figural measures. Of the first five measures included in the analysis, four were figural and one was verbal.

Forty percent of the first half of the analysis was made up of Pre measures while 60% were Post measures. Thus, the groups differed
more on the Post measures than on the Pre measures. This would seem to indicate that the subjects were influenced in some way in the course of the quarter, although the change was not detected by the previous analysis.

The breakdown by subtest provides more insight. Of the first nine variables included (the first third of the analysis), three were Originality measures (two figural, one verbal), four were Flexibility measures (three figural, one verbal), and two were ratios (one figural, one verbal). That means that 75% of the Originality measures, 50% of the Flexibility measures (75% of the FLEX, 25% of the FLEX S), and 25% of the ratios were in the top one third in discriminating ability. The groups seem to differ most in their figural Originality and figural Flexibility.

None of the three functions produced was significant. The first accounted for 59% of the variance and included FIGORIG1, FIGFLEX2S, and VERRAT1 with FIGORIG most highly weighted. It would seem to be a figural Originality function. Group two placed high on this function, and groups one and four placed low.

The second function accounted for 29% of the variance and included three Fluency measures (two verbal, one figural), FIGFLEX1, VERORIG2 and VERRATLS. It seems to be a Fluency function with two more Pre than Post measures and two more verbal than figural measures. Group three placed high and group one low on the function.

The third function was composed of all the other variables, most notably six Flexibility measures and six ratios (which are also measures of flexibility). It is obviously a Flexibility function.
There were more Post than Pre measures and more figural than verbal measures included. It accounted for only 12% of the variance. On this function, group three placed high and group four low.

It would seem, then, that group two had a facility for figural Originality. Group three could be characterized by Fluency and Flexibility. Groups one and four seem more characterized by a lack of figural Originality. In addition, group one lacks Fluency and group four lacks Flexibility.

It is of note that the two groups that seem to lack flexibility (groups one and four), are the two groups that chose to maintain the same status of enrollment in second-language courses that they had in high school. Group one continued to study second languages, and group four continued not to study them. Groups two and three, on the other hand, changed their enrollment status with regard to languages. Group two stopped studying, and group three began to study a second language. The results of this research project may say less about the effects of second-language study than they do about personalities open to change and those fixed in habit. The results must always be seen as tentative, however, because of the small sample sizes and the possibility that group three may be atypical of the students of that group in the population.

Summary

The analyses performed to answer the three major questions yielded the following results:
**Question 1:** The 1979 volunteer subjects could not be shown to have come from a different population than the 1978 randomly selected subjects except in that a significantly larger proportion of the volunteer subjects were enrolled in a 101 course than were the randomly selected subjects. There were several indications, however, that group three of the 1979 sample was composed of a different type of student than was the 1978 group three.

**Question 2:** In the 1979 sample, group four had significantly lower ACTENG scores than did the other groups. Multivariately, there were no differences among the groups in their cognitive flexibility upon entering the university, but univariately, groups two and three had significantly higher scores on all measures except the ratios than did groups one and four.

**Question 3:** Multivariately, there were no differences among the groups in their cognitive flexibility at the end of the quarter, but univariately, group three had significantly higher scores than the other groups on verbal Originality. When the Pre Flexibility measures were covaried on the Post measures, however, this effect was no longer significant. This analysis did show that subjects who had studied a second language in high school had higher scores in figural Elaboration. Discriminant analysis of all the flexibility measures showed that, overall, groups two and three were most highly differentiated in their Post figural Originality and Flexibility. Since the more flexible groups were the two that changed their enrollment status in second-language courses, it was suggested that perhaps a personality trait of
willingness to change was being reflected in the flexibility scores rather than an effect of second-language learning.

All results of the study must be interpreted with great caution since they are based on the scores of a small and, perhaps, atypical sample.
V. SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Summary

This study attempted to discover whether or not the study of a beginning-level second-language course during the first quarter of enrollment in the university would affect the cognitive flexibility of a student. A group of 51 volunteer subjects was tested at the beginning and end of Autumn quarter 1979 on verbal and figural measures of cognitive flexibility. The subjects were placed into one of four treatment groups according to their second-language learning history; that is, according to whether or not they had studied a second language in high school and whether or not they were enrolled in a 101-level second-language course that quarter.

Before the end-of-the-quarter flexibility scores could be analyzed and the principal research question answered, two preliminary questions had to be answered. First, since the subjects were volunteers, a determination of their representativeness of the population as a whole had to be made. To answer this question, the HSGPAs, ACT scores, and GPAs at the end of Autumn quarter of the volunteers were compared with those of a randomly selected sample of subjects from Autumn 1978. The proportion of subjects in each group from year to year was also compared. The analyses performed to answer Question 1 indicated that there were no significant differences between the volunteer and randomly
selected subjects on these measures except that a larger proportion of volunteers was enrolled in a 101 course than was true of the randomly selected subjects. This was thought to be due to the nature of the study. It seemed reasonable that students who were enrolled in a second-language course would be interested in participating in a study on second-language learning.

Although there were no other significant differences between the 1978 and 1979 subjects, there were several indications that the 1979 subjects in group three (no, yes) were of a different type than those in the 1978 group three. This suggestion and the very small sample size make necessary the use of great caution in the acceptance and interpretation of the results of the study.

The second question that needed to be answered before the post-treatment data could be analyzed was the following: Could subjects in each of the four groups be considered to have come from the same population before treatment with regard to their academic ability and cognitive flexibility? The analyses performed indicated that group four had significantly lower ACTENG scores than the other groups. Multivariately there were no significant differences among the groups in their flexibility. Univariately, however, groups two and three had significantly higher scores than groups one and four on all measures except the four ratios.

The third question to be answered was the principal one—the analysis of the post-treatment flexibility data. Because of the small sample size and the suggestion that group three was an atypical group, the interpretation of the data must be tentative. Multivariately, there
were no significant differences among the groups, but univariately, group three had higher scores than the other groups on verbal originality. This result did not prove to be significant when the pre-treatment flexibility measures were covaried on the post measures. Then the only significant result was that subjects with high school second-language study had higher scores on figural elaboration. Discriminant analysis including all of the flexibility measures indicated that groups two and three had higher scores than groups one and four and that, in general, the groups were most highly discriminated by their post figural originality and flexibility. This result suggested that it could be a personality trait of openness to change that was being reflected in the flexibility measures rather than an effect of second-language learning.

Discussion

The principal research question of this study as stated in Chapter I—What are the effects of 101-level study of a second language on measures of cognitive flexibility?—could not be answered with confidence. In that sense, the research effort failed. There were two principal and interrelated reasons for the failure: the small sample size and the use of volunteers.

Although the sample size is taken into consideration by the statistical analyses, and all results that proved to be significant were so, the small sample size did not provide the necessary power to the statistical tests for detecting other significant differences that may have existed in the data. Moreover, extreme scores have more of an
effect on means in small samples than in large ones. Group three of 1979 is a case in point. If the four subjects of group three are typical of that group in 1979, then had there been more of them, a significant difference between their scores and those of the 1978 group-three subjects probably would have resulted in the analysis. If they are not typical of the subjects of the 1979 group three, then a larger sample including more of the typical subjects would have lessened the effect of the four subjects' scores on the overall group means. With such a small sample, it is impossible to tell which is the case. A comparison of the scores of a random sample of 1979 students in addition to those of the 1978 students would probably have made the interpretation easier. It is especially unfortunate that group three was such a small group because it was the group of primary interest in the study: students who were studying a second language for the first time.

The use of volunteers determined the small sample size. It also produced the lack of representativeness of the subjects. The greater proportion of students in 101-level courses in 1979 than in 1978 was a function of the use of volunteers as was the seeming atypicalness of the 1979 group three. In fact, the first may explain the unusually high scores of the group-three subjects. Perhaps only good, interested, hard-working students volunteered from that group.

The success of the study was limited by several other factors not directly related to sample size or the use of volunteers, although a larger, random sample would have improved the chances for success. The decision to include subjects with up to one year of high school second-language study in the no-language groups is one such factor.
While defensible, and, for practical reasons, necessary, it precluded the absolute determination of the effect of a first experience in second-language learning. Had the sample been larger or not composed of volunteers, perhaps it would not have been necessary to include subjects with language study in the no-language groups.

Another problem arose from the decision to use the 101-level second-language course as the treatment. It was the only choice, since it is the course that students who have never studied a second language take. Because it is the first course, however, students who have studied a second language in high school long enough to enter a more advanced course in the university do not take it. Thus, the students who learned the most and in some ways benefited the most from their language-learning experience in high school were eliminated from the study altogether. Group one, therefore, was composed of students who learned so little language in high school that they had to start all over again in the university and students who purposely did not take the Placement Exam so that they could take the first course for an easy grade. This situation would also have been true for the 1978 group one. The results of the study would, therefore, not be generalizable to all students in the ASC sections of the University Survey course, but rather only to those who would have taken a 101-level course.

For group two the situation is different. Students in the 1978 population who did not study a second language their first quarter in the university were not included in the sample if, in a subsequent quarter, they began second-language study at a level more advanced than the 101 course. Thus, as nearly as possible, all subjects in the 1978
sample were beginning-level students. Volunteers for the 1979 sample could not be screened in this way. Hence, some of the subjects in the group that had studied a second language in high school but did not enroll in the 101 course Autumn quarter 1979 could have been more advanced students of the second language who, for one reason or another, did not choose to begin their university language study during their first quarter in the university and, therefore, did not take the Placement Exam at that time. The composition of group two may, therefore, be different from that of group one.

The decision to use only the ASC sections of the University Survey course was made in an attempt to standardize students' motivation for studying a second language: If they graduated from ASC, they would have to complete four quarters of study of a second language. Although this is true, there are still those who study a language because they like it and those who study a language because they have to. Every person has his or her own reasons for the decision. By only including students in the ASC sections, the number of students who had never studied a second language was severely limited. In the study, 20% had not studied a language, but the percentage in the population would have been smaller since the students in advanced second-language courses were eliminated from the study. Moreover, being such a small part of the ASC group, the students who had never studied a second language were atypical of the group from the outset. This can be seen in group four's lower scores in everything. In 1978, group three also had lower scores. To have tried to build these students into the study in numbers equal to those of the students who had studied a second language may have been
biasing the study. Perhaps it would have been wiser to have used subjects from all of the University Survey sections. There would have been more variability in motivations for studying a second language, and, probably, more variability in academic ability. Perhaps, though, the pool of students who had never studied a second language would have been larger, and the results would have been generalizable to the entire population of freshman students.

Another confounding variable in the study may be the nature of the second-language courses in which the subjects were enrolled. These courses will most likely vary in the language learning activities they employ along the continuum of mechanical to meaningful activities. If cognitive flexibility can, indeed, be altered by second-language study, it would seem probable that the deeper processing of meaningful language learning activities would produce greater, more rapid changes. This variable was not controlled at all in the study, but should be in any future investigation.

Bearing in mind that the study is plagued by problems that greatly limit the generalizability of the results, there are several comments that can be made about the cognitive flexibility of these subjects. It is quite clear that the fluency variable pervades the others with the exception of the ratios. The pattern of the group means for each of the variables is very similar to that of its fluency variable. This observation is reasonable since all of the variables are computed on the basis of the same responses and their number directly influences the total possible score. It was for this reason that the ratio scores were included in the analysis. By dividing the flexibility
score by its fluency score, it was hoped that a more fluency-free index of flexibility could be achieved. These ratios did not follow the patterns of the fluency-influenced flexibility scores. In the pre flexibility analysis the ratios were the only variables that did not prove to contain significant differences among the group means. This outcome may mean that although groups two and three could produce more responses to the stimuli, they were not any more flexible in their responses than groups one and four. The same was true for the post flexibility tests. The ratio scores never showed any significant differences, though the sample size may have lessened the chance of their being detected.

It is interesting that it was verbal originality that showed a significant interaction in the post measures with group three being higher than the others. Cummins (1975), Cummins and Gulutsan (1974), and Carringer (1974) all found that their bilingual groups excelled especially in verbal originality. Group three was experiencing the study of a second language for the first time, and although they scored higher than the other groups in verbal originality at the outset, they also had the greatest increase in score. When the pre measures were controlled for, the difference was not significant, but this result may be an indication that second-language study increases verbal originality. The final discriminant analysis showed that the groups were differentiated well by the originality variables as well.

The significant main effect in the Ancova for HSLANG on the figural elaboration variable is more difficult to interpret. Those who studied a second language in high school drew more details in their
figure completions. This result reflects the performance on the post test when pre test performance was controlled for. It reflects the influence of some aspect of first quarter university life since both a course-101 and a no-course-101 group are represented. Perhaps the study of a second language in high school allows the student to deal with complexities better—to be able to branch out and embellish his life more.

The earlier suggestion that the flexibility measures reflect an openness on the part of the subject to change remains tenable. Groups two and three—those that changed their enrollment status in second languages—consistently scored higher than the groups that continued as they had. Further research would be needed to determine whether or not an openness trait was at work here.

Recommendations for Further Research

This study was a first attempt at discovering non-linguistic outcomes of second-language study at the college level. No previous research had been completed in the area. As a first study in the area, the study was in many ways a shot in the dark. Studies using bilingual children had found that cognitive flexibility seemed to be a dimension that was influenced by the ability to use two languages. The intuitions of foreign language educators also indicated that second-language learning should influence the cognitive flexibility of the student (Jarvis, 1978). This study, then, attempted to determine whether or not beginning college-level second-language study could influence the
students' flexibility. Even had it not been handicapped from the outset by the necessity of using volunteers, its results would have been tentative. Many more research efforts will be needed before it can be said with some certainty that cognitive flexibility is or is not influenced by second-language study. Even more studies will be needed to determine what other non-linguistic outcomes there are from language study.

The researcher was gratified to see that the subjects in the study also believed that second-language study can change people in non-linguistic ways. In response to the question "Do you think that studying a foreign language can change a person in any way (thinking ability, verbal skills, openness, flexibility, etc.)? How and why?" on the questionnaire completed after the post flexibility tests many subjects indicated that language study can make people open to new ideas and more flexible. They thought it taught people to think more concisely, more analytically, to reason better, to be more aware of details. It could give people more ways of approaching things, more options so they could make better decisions. One subject even mentioned that it made him see that words are arbitrary—a book is a book because we call it that. The subjects also mentioned the cultural and linguistic outcomes of language study: that it helps enlarge vocabulary, to understand English better, to increase verbal skills; that it helps understand other peoples and their countries and cultures. Several subjects also mentioned that learning a second language improved their self images. Clearly, second-language learners believe that they can be changed by their experience.
The problem is one of research methodology. How does one reduce the number of variables at play? How does one isolate the variables in question? How does one decide which variables may be worthy of investigation? The cognitive flexibility dimension is still promising. Perhaps if the study were done with younger subjects none of whom had studied a second language before, the task would be less complex. If a junior high school could be located in which all seventh graders studied a language for one semester and the group that studied the first semester were chosen at random, the research project would have good possibilities of success. With college-aged subjects a random sample more representative of the entire population would increase the chances of success over those of the present study. A larger, more diverse battery of tests would also improve the chances of discovering a consistent effect.

The time element is also important. Perhaps nine weeks between testings is not sufficient to detect changes. In the university, however, if the study lasts longer than one quarter, the number of extraneous variables introduced becomes enormous. Attrition would also increase. University students are a convenient group to use for studies of this kind, but they may not be the population most suited to finding an answer to the question.

There are many other possible non-linguistic outcomes of second-language study. The subjects of this study mentioned several. They should all be investigated. The more that is known about what happens to people when they learn a second language, the better able we will be to refine those learnings and the more precise we can be when we
talk to prospective students. Vague notions that second-language learners are somehow able to cope better in an ever shrinking world are not enough.

If the non-linguistic outcomes of second-language learning are to become known, there must be an organized plan of investigation. A group of researchers interested in these outcomes must meet and discuss ways of delimiting the problem. They must use their powers of reasoning and their problem-solving strategies to sort out the variables and to design a series of creative studies that can narrow the scope of the problem. Only then will there be a chance of specifying the non-linguistic outcomes of the study of a second language.
BIBLIOGRAPHY


Cummins, J. Cognitive factors associated with the attainment of intermediate levels of bilingual skills. *Modern Language Journal*, 1977, 61, 3-12. (b)


Thorndike, E. L. Mental discipline in high school studies. *Journal of Educational Psychology*, 1924, 15, 1-22; 83-98.

