The relationship of field dependence/independence to reading for the main idea in adult learners of English as a second language

Danik Inger Wold

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Bilingual, Multilingual, and Multicultural Education Commons, English Language and Literature Commons, and the First and Second Language Acquisition Commons

Recommended Citation


https://lib.dr.iastate.edu/rtd/16084

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
The relationship of field dependence/independence to reading for the main idea in adult learners of English as a second language

by

Danik Inger Wold

A Thesis Submitted to the Graduate Faculty in Partial Fulfillment of the Requirements for the Degree of

MASTER OF ARTS

Department: English
Major: English (Teaching English as a Second Language)

Approved:

Signatures have been redacted for privacy

Iowa State University
Ames, Iowa

1988
# Table of Contents

## Introduction

Page 1

## Review of Relevant Research

Page 5
- Field Dependence/Independence
- FD/I and Second Language Acquisition
- FD/I and Reading Comprehension
- Schema Utilization in Reading Comprehension
- Assessment of Background Knowledge
- Main Idea Comprehension
- Assessment of Main Idea Comprehension
- Summary

Page 7

## Method

Page 67
- Subjects
- Materials
  - Group Embedded Figures Test
  - Background Knowledge Survey
  - Main Idea Test
  - Test Development
  - Test Piloting
  - Test Review
  - Reading Proficiency
  - The Michigan Test
  - Test of English as a Foreign Language
- Testing Procedure

Page 71

## Analysis and Results

Page 94
- Characteristics of the Sample
- Descriptive Statistics for the Entire Sample
- Descriptive Statistics for Sample Subgroups
  - Graduate/Undergraduate Subgroups
  - First Language Subgroups
  - Major Academic Area Subgroups
- Summary of Sample Characteristics
- Reliability of Measures
- Accounting for Variance on the Main Idea Test
- Reanalysis of Variance on the MIT
  - Descriptive Statistics for the Reduced Sample
  - Summary of Reduced Sample Characteristics
- Accounting for Variance on the MIT: the Reduced Sample
- Summary of Reanalysis of Variance on the MIT

Page 95

## Discussion

Page 157
- Review of Statistical Analyses
- Interpretation of Results
LIST OF TABLES

TABLE 1. Distribution of subjects by language group, graduate/undergraduate status, and major academic area .......................... 69

TABLE 2. Distribution of subjects by language group and major academic area ............................ 70

TABLE 3. Frequency statistics on seven measures for the entire sample ................................. 96

TABLE 4. Comparison of means and standard deviations on overall TOEFL and reading section scores for females ........................................ 97

TABLE 5. Comparison of means and standard deviations on overall TOEFL and reading section scores for males ........................................ 98

TABLE 6. Comparison of means and standard deviations on the Michigan Test overall scores .......... 100

TABLE 7. Comparison of means and standard deviations on the Michigan Test reading scores .......... 101

TABLE 8. Comparison of means and standard deviations of females and males on the Group Embedded Figures Test ................................. 102

TABLE 9. Frequency statistics on seven measures for graduates ............................................ 104

TABLE 10. Frequency statistics on seven measures for undergraduates ..................................... 105

TABLE 11. ANOVA between graduate and undergraduate performance on seven measures ............... 107

TABLE 12. Frequency statistics on seven measures for Arabic speakers ..................................... 109
TABLE 13. Frequency statistics on seven measures for Japanese speakers
TABLE 14. Frequency statistics on seven measures for Spanish speakers
TABLE 15. Frequency statistics on seven measures for speakers of miscellaneous languages
TABLE 16. ANOVA among the performance of the language groups on seven measures
TABLE 17. Scheffé test for source of variance among language groups on overall TOEFL scores
TABLE 18. Scheffé test for source of variance among language groups on TOEFL reading scores
TABLE 19. Scheffé test for source of variance among language groups on Group Embedded Figures Test scores
TABLE 20. Linear ranking of mean scores of the language groups on seven measures
TABLE 21. Frequency statistics on seven measures for business/economics majors
TABLE 22. Frequency statistics on seven measures for engineering majors
TABLE 23. Frequency statistics on seven measures for science/mathematics majors
TABLE 24. Frequency statistics on seven measures for liberal arts majors
TABLE 25. ANOVA among the performance of academic major groups on seven measures
TABLE 26. Scheffé test for source of variance among academic major groups on Group Embedded Figures Test scores
TABLE 27. Reliability estimates for the TOEFL: overall and reading section
TABLE 28. Reliability estimates for dependent and independent variables
TABLE 29. Chi-square tests between background knowledge survey items and corresponding main idea test items 130

TABLE 30. Pearson product moment correlations between dependent and independent variables for the entire sample 133

TABLE 31. Multiple regression analysis using the main idea test scores as the dependent variable and the TOEFL reading and Group Embedded Figures Test scores as independent variables 135

TABLE 32. Multiple regression analysis with the main idea test scores as the dependent variable and the TOEFL reading scores, graduate/undergraduate status and the Group Embedded Figures Test scores as independent variables 137

TABLE 33. Frequency statistics on dependent and independent variables for the sample minus Spanish speakers 140

TABLE 34. Frequency statistics on dependent and independent variables for graduates minus Spanish speakers 142

TABLE 35. ANOVA between graduate and undergraduate performance on dependent and independent variables for the sample minus Spanish speakers 144

TABLE 36. ANOVA among the performance of language groups on dependent and independent variables for the sample minus Spanish speakers 145

TABLE 37. Frequency statistics on dependent and independent variables for science/mathematics majors from the sample minus Spanish speakers 146

TABLE 38. Frequency statistics on dependent and independent variables for liberal arts majors from the sample minus Spanish speakers 146
TABLE 39. ANOVA among the performance of major academic groups on dependent and independent variables for the sample minus Spanish speakers

149

TABLE 40. Pearson product moment correlations between dependent and independent variables for the sample minus Spanish speakers

152

TABLE 41. Multiple regression analysis with the main idea test as the dependent variable and the TOEFL reading and Group Embedded Figures Test scores as independent variables for the sample minus Spanish speakers

152

TABLE 42. Multiple regression analysis with the MIT scores as the dependent variable and the TOEFL reading scores, graduate/undergraduate status and GEFT scores as independent variables for the sample minus Spanish speakers

153
INTRODUCTION

Researchers in second language acquisition, including those who study English as a second language (ESL) learners, have long been aware that no one learner variable can account for performance on a given language task. In an effort to understand the factors that influence a student's ability to acquire a second language, researchers over the last ten to fifteen years have conducted numerous "good language learner" studies. In addition to extending the theoretical base, the results of these studies have been used in such practical endeavors as tailoring instructional strategies to individual students' needs and teaching students how to alter their learning styles to enhance language acquisition.

One student factor often incorporated into the research design of these studies is the cognitive style of field dependence/independence (FD/I). Simply put, field independence (FI) refers to an individual's tendency to perceive as discrete the relevant parts from a larger "field" of distracting items. Conversely, field dependence (FD) refers to an individual's tendency to be overwhelmed by the total field such that the discrete relevant parts are not easily perceived. A great deal of research has shown that the perceptual and intellectual style of processing
characteristic of FI individuals is a positive factor in second language acquisition; students who are relatively more FI tend to acquire a second language more easily and efficiently than students who are relatively more FD. The same observation has been made of the relationship of FD/I to reading comprehension; students who are more FI tend to score higher on tests of general reading proficiency than students who are more FD.

FD/I is by no means the only learner variable which affects a student's ability to read and comprehend a second language. Additional factors include intelligence and level of conceptual ability, other cognitive styles such as reflectivity/impulsivity, attitudes and level of motivation regarding reading and learning a second language, linguistic proficiency in the second language, and reading proficiency in the first language. In addition to these variables researchers, including many in ESL, have investigated the extent to which a student uses background knowledge to facilitate the comprehension of text. The research has shown that students who are able to actively incorporate background knowledge into their processing of text are better, more proficient readers in the second language.

This study is concerned primarily with the effect of FD/I, and secondarily with the effect of background
knowledge utilization, on the reading comprehension of adult learners of academic English. As reading comprehension is a construct which subsumes a multitude of tasks and abilities, this study will concern itself with one particular aspect of the larger construct: main idea comprehension in expository prose. This subconstruct was chosen for study because it seems to have particular relevance to adult ESL learners in an academic setting. Most college students would agree that the assimilation of a great volume of complex information requires the comprehension and memory of main ideas as well as specific details; this is perhaps even more important for an ESL student who has a linguistic barrier to deal with as well.

It is hoped that the study of these three constructs in combination (FD/I, background knowledge utilization, and main idea comprehension) will extend what is already known about them in other domains. As mentioned before, numerous studies have been done on FD/I and second language acquisition, as well as FD/I and general reading comprehension. Far fewer studies have been done on FD/I and background knowledge utilization. Studies of the relationship between FD/I and main idea comprehension or background knowledge utilization and main idea comprehension are scant indeed. In fact, main idea comprehension is a
construct which appears to have escaped the direct attention of most ESL, FD/I, and background knowledge researchers to date.

Thus, the main purpose of this study is to discover the effects of two learner variables on the ability of adult ESL learners to comprehend the main idea in expository writing. Because previous research has indicated that general second language reading proficiency is a significant predictor of performance on specific reading skills tests, reading proficiency will necessarily be considered when analyzing the effects of the other two variables. Therefore, the main research question being asked here is as follows: does FD/I account for variance in the performance of adult ESL learners on a test of finding the main idea beyond that which can be accounted for by reading proficiency and background knowledge?
REVIEW OF RELEVANT RESEARCH

This chapter reviews literature from several interrelated strands of research germane to the inquiry at hand: field dependence/independence (FD/I), second language acquisition, and reading comprehension. Because it is a central construct in this investigation, the first section of the review is devoted to the definition of FD/I. The next two sections of the chapter focus on two areas of FD/I research. The second section reviews research on the role of FD/I in second language acquisition. The third section discusses research on the role of FD/I in reading comprehension.

The last four sections of the chapter focus on aspects of reading comprehension which are pertinent to this study. The fourth section of the chapter opens with an overview of current models of reading comprehension. Over the last ten years, great strides have been made in the understanding of reading comprehension in English as a second language (ESL) through research into the role of background knowledge, or schema utilization. Thus, this section continues with a discussion of schema theory and its particular application in ESL research. Because schema utilization is being examined here as a potential predictor variable, the fifth section of the chapter reviews ways in which background
knowledge has been assessed in empirical investigations. The sixth section of the chapter narrows in focus to the specific subconstruct of reading for the main idea. This section reviews previous studies of the subconstruct in an attempt to define how FD/I might affect main idea comprehension. Finally, because the main idea test in this study was constructed by the investigator, the seventh section of the chapter reviews some ways in which main idea comprehension has been assessed in research situations.

The main reason for conducting this study is to extend what is known about the acquisition of English as a second language by adult learners. Thus, it should be noted that wherever possible the research selected for review comprised studies done using adult ESL learners. However, the reach of ESL research has apparently not yet extended fully into all the domains of inquiry described above. For example, much of the early research on the relationship of FD/I to second language acquisition was conducted using adolescent native speakers of English in foreign language learning situations (e.g., Bialystok & Frohlich, 1978; Naiman, Frohlich, Stern, & Todesco, 1978; Tucker, Hamayan, & Genesee, 1976). Likewise, most of the research concerning the role of FD/I specifically in reading comprehension was performed on children and adolescents who were native
speakers of English (e.g., Pitts & Thompson, 1982; Roberge & Flexer, 1981, 1984; Spiro & Tirre, 1980). Further, while a great deal of research has been done on the reading comprehension of adult ESL learners (the extensive work of Patricia Carrell, for example), no ESL studies were found which focused specifically on reading for the main idea.

The review of the above areas in which ESL studies were scant was made possible by research which used native English speakers as subjects. This study acknowledges the wisdom of proceeding cautiously when formulating hypotheses based on such work since the learners and learning situations in these studies may not be completely analogous to those in ESL. However, out of necessity and in many instances with sound results, ESL investigators have found an empirical foundation for their research in work from outside but related fields. Such is the case with this study, which has drawn theoretical bases from research in foreign language acquisition and reading research in elementary and secondary education, as well as research in ESL.

Field Dependence/Independence

FD/I is one dimension of cognitive style, the unique way each individual has of perceiving, organizing,
analyzing, or recalling information and experience. Cognitive styles, as conceived by Witkin, Moore, Goodenough, and Cox (1977), are the characteristic ways in which an individual processes external stimuli; they are a composite of individual tendencies in perceptual and intellectual functioning. These characteristic tendencies are seen to be pervasive within the individual. They influence not only perceptual and intellectual processing, but also aspects of personality, including the way in which an individual relates to others. Cognitive styles are seen to be stable over time. Finally, cognitive styles are conceived as bipolar to distinguish them from abilities. When one speaks of an ability, the implication is that it is better to have more of the ability than less of it; when one refers instead to a cognitive style, the implication is that "each pole has adaptive value under specified circumstances, and so may be judged positively in relation to those circumstances" (Witkin et al., 1977, p. 16).

Field independence (FI) is defined by Witkin et al. (1977) as "the extent to which the person perceives part of a field as discrete from the surrounding field as a whole, rather than embedded in the field" (p. 6). Field dependence (FD), on the other end of the continuum, is "the extent to which the organization of the prevailing field determines
perception of its components" (Witkin et al., 1977, pp. 6-7). In other words, an individual who is more FI has a tendency to perceive environmental stimuli analytically. This individual is able not only to overcome the organization of the field, but can restructure a field which has a dominant organization, and impose structure on a field with no inherent organization. On the other hand, an individual who is more FD tends to perceive external stimuli holistically; that is, perception is dominated by an existing field such that the individual can not see the tree for the forest. This individual is also less likely to use mediators such as analysis and restructuring to facilitate cognitive processing.

This dimension of cognitive style was originally explored using visual, tactile, and kinesthetic stimuli. The instrument used in this study, the Group Embedded Figures Test (GEFT), is an example of a visual test of FD/I in which the individual must locate a simple geometric figure in a complex line drawing. In addition to affecting visual perception, FD/I also influences the perception of symbolic representations, such as an individual encounters in thinking and problem solving. The more FI person will be at advantage with that class of problems "where the solution depends on taking some critical element out of the context
in which it is presented and restructuring the problem material so that the item is now used in a different context" (Witkin et al., 1977, p. 8). Conversely, the more FD person is likely to have difficulty with this sort of task, though Witkin et al. (1977) are careful to point out that this is the only class of problems they have identified in which a more FD person would likely be disadvantaged.

In addition to the perceptual and intellectual aspects of processing described above, the influence of FD/I also extends into characteristics of personality. It is in this domain that the attributes of the more FD individual are described in the literature as strengths. Someone who is more FD is seen to be generally more gregarious, attentive to social cues, interested in what others say and do, and is perceived by others as warm, tactful, considerate, outgoing, and affectionate. Individuals who are more FI, on the other hand, are seen to have a more impersonal social orientation. They tend to show little sensitivity to social cues, may seem distant or aloof with others, and generally prefer individualistic, solitary pursuits. Through this study of FD/I and personality, Witkin et al. (1977) have found that individuals often make educational and vocational choices that are in keeping with their relative levels of FI or FD.
FD/I and Second Language Acquisition

Much work has been done in the last ten to fifteen years to seek evidence for a relationship between FD/I and second language acquisition. One important question in this area has been whether students who are more FI make better second language learners than students who are more FD. The answers vary depending on whose review of the literature one subscribes to. Some researchers have found support for their hypotheses that FI students are better second language learners in those studies, or parts of studies, which point to FI as a predictor of success on a number of second language tasks (e.g., Genesee & Hamayan, 1980; Naiman, Frohlich, Stern, & Todesco, 1978; Tucker, Hamayan, & Genesee, 1976). Other researchers have been circumspect, noting the inconclusive results of studies, or parts of studies, which find little to no direct relationship between FI and proficiency in second language tasks (e.g., Bialystok & Frohlich, 1977; Tucker, Hamayan, & Genesee, 1976).

These differences in interpretation of the literature are sustained by a number of empirical considerations. For example, most of the second language tasks in research studies were characteristically academic in form and function. For this reason, the finding that FI predicts success in the second language may be relevant only in the
classroom setting where such tasks are required; it may or may not be applicable to success in second language tasks which are required in the world outside the classroom. Another point which complicates the generalization of results is that researchers have not uniformly shown FI students to excel in all academic tasks put to them. Interestingly, no one has yet discovered tasks on which FD students are advantaged due to cognitive style. This gap in the literature is undoubtedly due in part to the inevitable limitation of scope inherent in empirical studies. No one study has as yet been able to test the relationship of FD/I to more than a few of the vast number of tasks which learners face in either the academic or the real world of second language acquisition; for this reason, it is difficult to obtain results which can be generalized to all situations. Thus, the empirical process moves slowly, with progress toward complete understanding being hampered by conflicting or inconsistent results.

Most early studies on the relationship of FD/I to second-language acquisition were done using native English speakers in a foreign language learning context. In an effort to describe the good language learner, these studies examined FD/I as one of the multiple cognitive, affective, and social factors which affect second language acquisition;
most found a positive relationship between FI and proficiency on second language tasks. One of the earliest and most extensive studies of this kind was conducted by Naiman, Frohlich, Stern, and Todesco (1978). As summarized by Abraham (1979), the purpose of this study was to identify strategies and techniques of good language learners. In addition, they examined the influence of such cognitive styles as FD/I, constricted/flexible control, breadth of categorization, and an aspect of personality called tolerance of ambiguity. Their subjects were English-speaking Canadian high school students (eighth, tenth, and twelfth graders) studying French as a second language. The subjects were administered two tests of oral French ability; one was a listening comprehension test, and the other was an imitation task in which the student had to repeat ten sentences of nine to fifteen syllables each. Their results showed FI and tolerance of ambiguity to be significant predictors of success for students at some levels on both instruments, and suggested that FI may play a more important role in the more advanced stages of second language acquisition. An error analysis of the imitation test revealed that the perception of FD learners was dominated by complete sentences such that they would omit entire segments rather than imitate the smaller constituent units they were
able to perceive within the larger segments. They also seemed to make more errors attributable to the distraction of other elements within the field.

Other studies which found a positive relationship between FI and second language acquisition include Tucker, Hamayan, and Genesee (1976), Genesee and Hamayan (1980), and d'Anglejan and Renaud (1985). All three of these studies were conducted in a French-as-a-second-language setting in Canada; with the exception of the subjects in the d'Anglejan and Renaud study (who were immigrants from diverse backgrounds), all the subjects were native English speakers. Tucker, Hamayan, and Genesee (1976) conducted a pilot study of the relationship of cognitive, affective, and social factors to performance on a variety of achievement measures. Their subjects were English Canadian seventh graders who had been studying French since kindergarten. Tucker et al. discovered that FI was related to performance on a standardized multiple choice achievement test, which included sections for spelling, listening comprehension, vocabulary, and grammar.

While the next two studies found a positive relationship between FI and second language acquisition, they qualified their conclusions by calling attention to an apparent overlap between the effects of intelligence (IQ)
and FI, two constructs purported to be functionally independent. Genesee and Hamayan (1980) investigated the influence of multiple cognitive and affective variables on the performance of English Canadian first graders on four measures of French proficiency. A factorial analysis was used to reduce the number of potential predictors to be entered into a subsequent regression analysis; thus, the effect of FI was examined as part of a factor which included nonverbal reasoning (IQ) and attitude toward continued schooling in French. Regression analyses revealed that this factor was a positive significant predictor of performance on a general proficiency test, a listening comprehension test, and a test of reading achievement in English. However, Genesee and Hamayan noted that "achievement in French cannot be interpreted to reflect the field independence trait exclusively since this factor was made up of two other items with high loadings" (1980, p. 106).

In a similar study, d'Anglejan and Renaud (1985) related various cognitive and affective factors to the French achievement of adult immigrants. An analysis of

---

1This instrument was used to test the transfer of reading proficiency in French to reading proficiency in English. Though they were native English speakers, the first graders in this study were schooled exclusively in French prior to receiving literacy training in English. Thus, in this case the direction of transfer of reading skills was from the second language to the first.
multivariate variance revealed a positive significant relationship between FI and nonverbal reasoning (IQ), and performance on a French achievement test (which included a reading comprehension subtest). However, a multiple regression analysis showed that IQ was the only cognitive variable under investigation which was a significant predictor of performance. The subjects were subsequently classified as either good or poor learners on the basis of teacher evaluations. A multivariate analysis of variance showed that FI, independent of IQ, was a positive significant predictor of membership in the group judged to be good language learners. In light of this finding (as well as a significant correlation found between FI and IQ), d'Anglejan and Renaud concluded that the interaction of the two variables obscured the true contribution of FI in the initial analysis.

Another study which found a positive relationship between FI and aspects of second language acquisition was conducted by Hansen and Stansfield (1981). Rather than study a group of cognitive and affective variables, these investigators focused specifically on the role of FD/I. In a study of adult learners of Spanish in an American university, Hansen and Stansfield hypothesized that students who were more FI would perform better on a test of
linguistic competence, while students who were more FD would perform better on a test of communicative competence. In addition to being tested for these competencies, students were tested with a measure of integrative competence, a cloze test. Following a correlational analysis, the authors concluded that despite their hypotheses, FI played a positive but minor role in performance in all three areas of competence; FD was not associated with communicative competence as predicted. The strongest correlations occurred between FI and performance on the cloze test, leading the authors to conclude that the inferencing competence required for cloze performance must derive from a more FI cognitive style.

Stansfield and Hansen (1983) conducted a follow-up analysis of the above data to determine if a test bias existed for the cloze test in favor of students with a more FI cognitive style. Additional analyses revealed that FI and the cloze test scores shared an extraordinarily similar correlational pattern with measures of academic aptitude and were "more related to quantitative ability than to verbal aptitude" (p. 36). When academic aptitude was partialled out, the correlations between FI and the measures of linguistic and communicative competence were rendered nonsignificant. However, the relationship between FI and
performance on the cloze test remained significant, leading
the authors to conclude that the cloze test may be biased in
favor of FI students.

Despite the evidence presented in the above studies,
the literature does not completely support a positive
relationship between FI and second language acquisition.
One example of this lack of consistency is demonstrated in
the results of a study reviewed earlier. While Tucker,
Hamayan, and Genesee (1976) found a positive association
between FI and a test of general French proficiency, they
did not find such a relationship between FI and separate
measures of listening comprehension, reading comprehension,
and oral production. In another example, Bialystok and
Frohlich (1978) investigated the role of cognitive and
affective variables in the French acquisition of English
Canadian high school students. They cite a previous study
(Bialystok & Frohlich 1977), in which they found a lack of
evidence for a positive relationship between FI and reading
comprehension. Their analysis in the 1978 study produced
similarly inconclusive results regarding FI; aptitude
correlated with FI, and FI correlated with performance on
formal and functional tests of oral and written ability (the
test of functional written ability was a reading test which
measured comprehension of coherent passages). However, in a
multiple regression analysis, aptitude, not FI, emerged as a significant predictor of performance on the proficiency measures. While the above studies do not appear to be in the majority, they represent to many researchers sufficient cause for skepticism about the claim that students who are more FI are better second language learners than those who are more FD.

In more recent years the investigation of FD/I in second language acquisition has extended from the foreign language learning context of native English speakers to the ESL learning context of nonnative English speakers. The latter is more directly relevant to the present investigation because it has focused on adult ESL learners in academic settings.

Some of these ESL studies have corroborated earlier findings from foreign language acquisition research in a fairly uncomplicated fashion. Chapelle and Roberts (1986) conducted a study similar to early good language learner studies; they investigated the role of several cognitive and affective variables in performance on various tests of ESL proficiency. They found FI to be a positive significant predictor of performance on the TOEFL (overall and part scores, including the reading section), a multiple choice grammar test, a dictation test, a cloze test, and an oral
test of communicative competence. Abraham (1983) investigated the relationship of several cognitive style variables to use of the strategy of monitoring on three writing tasks. Spanish-speaking ESL students were tested for their ability to monitor use of the third-person-singular *s* on fill-in-the-blank, proofreading, and composition tests. FI was found to relate positively to the amount of monitoring on all three tests, but Abraham points out that the relationship is not so great as to preclude the contribution of other learner variables to strategy use. Hansen (1984) examined the relationship between FI and performance on cloze tests by ESL students from six Pacific island cultures. Overall, this study confirmed the findings obtained by Stansfield and Hansen (1983), which indicated a positive relationship between FI and cloze test performance. However, Hansen cautioned that cross-cultural generalization of the role of FI in language proficiency may not be appropriate.

Two studies previously discussed found a strong positive (and thus empirically problematic) relationship between IQ and FI (d'Anglejan & Renaud, 1985; Genesee & Hamayan, 1980) despite the assertion of Witkin, Oltman, Raskin, and Karp (1971) that FD/I is a construct which is distinct from general intelligence. As with the second
language acquisition studies noted above, IQ has clouded the results of two studies in FD/I and ESL acquisition. Watkins and Astilla (1980) studied the relationship between FD/I, intelligence, and academic achievement in female Filipino high school students. Like the studies mentioned above, they found a positive correlation between IQ and FI. In subsequent regression analyses they also found that IQ was by far the greatest predictor of achievement. FI was found to share "a small but not insignificant amount of variance with achievement after the variance attributable to intelligence was removed" (Watkins & Astilla, 1980, p. 594). Another investigation which incorporated measures of intelligence was a comparative study of the performance of Maltese-Australian and Anglo-Celtic Australian high school students (Gauci, 1983). The author found that FD/I accounted for very little variance for either ethnic group on a test of English reading ability and two cognitive skills tests (social science and mathematics/science); the contribution of FD/I was greatly overshadowed by that of IQ. Of interest for this study is that despite its meager contribution after the removal of IQ, FD was the operative cognitive style which predicted reading test scores for all subjects but the Maltese-Australian males.
While the above ESL studies have generally produced evidence for a positive relationship between FI and second language learning, others have provided numerous examples of tasks in which FI did not appear to affect performance. In fact, ESL research on the effect of FD/I may be characterized to some extent by the mixed findings obtained within and between studies.

For example, Seliger (1977) examined the extent to which the active pursuit of oral practice in the classroom contributes to oral proficiency in adult ESL students. Observers recorded speech acts (defined as any type of utterance in English) to classify students as either high input generators (HIGs) or low input generators (LIGs). Once this classification had been made, three HIGs and three LIGs were tested using a number of proficiency measures. The subsequent analysis revealed that HIGs achieved higher scores than LIGs on end-of-semester measures of structure and listening comprehension, demonstrated higher levels of language contact and motivation than LIGs, and were more FI than LIGs. Seliger did not, however, find a significant relationship between HIGs and cloze test performance. Although these results are interesting, they must be evaluated with skepticism in light of the extremely small sample size.
Day (1984) conducted an investigation modeled on Seliger's study which extended Seliger's hypotheses to examine the performance of middle input generators (MIGs). Day improved upon Seliger's study in two ways. He tested a much larger sample than Seliger did (13 HIGs, 13 LIGs, and 26 MIGs). He also made finer distinctions in the coding of students' speech acts; he coded self-initiated speech acts separately to distinguish them from speech acts performed in response to a teacher's general or personal solicit. Day found no support in his analysis for confirmation of Seliger's findings (with the exception of the lack of a significant relationship found between HIGs and cloze test performance, which Day's data corroborated). Furthermore, Day found no support for his own hypotheses that the classroom participation of HIGs, MIGs, or LIGs was related to FI or performance on proficiency measures.

The results of several FD/I and ESL acquisition studies which have used the cloze as a proficiency measure have produced contradictory results. Earlier it was mentioned that Stansfield and Hansen (1983) and Hansen (1984) found evidence for a positive relationship between FI and cloze test performance. In a more recent study, however, Chapelle (in press) obtained results which do not entirely support the findings of previous researchers. Chapelle investigated
the relationship of FI to performance on four language proficiency measures (which included a cloze test) using native speakers of English in regular and remedial English classes as well as nonnative speakers in ESL classes as subjects. While there was a positive relationship found for FI and the cloze for the native speakers, no such relationship was noted for the nonnative speakers in ESL classes. Thus, the issue of test bias on the cloze continues to be clouded. However, for the purposes of this study it is interesting to note that Chapelle found a significant correlation between FI and performance on another of the proficiency measures used, a multiple choice reading test.

The above discussion should suffice as evidence of the current muddle in FOil and ESL acquisition research. Despite assertions to the contrary, it seems that the results of studies in this area are in many instances inconclusive. It is important to note, however, that the absence of a predictive relationship between FI and test performance in the above studies was not construed by researchers as a predictive relationship between FD and test performance. Chapelle (in press) suggests that this may well be a function of the way the FD/I construct is operationalized. A high score on a FD/I test is indicative
of field independence; a low score does not necessarily indicate field dependence, but rather relatively less field independence. Thus, it is difficult to find research which is able, directly or indirectly, to identify tasks in which a more FD person is advantaged.

This state of affairs may now be changing. One recent study has given attention to students at the FD as well as the FI end of the continuum. This study has confirmed what has been suggested by previous researchers: FD learners may derive greater benefit from less rule oriented and more inductive instructional methods. Abraham (1985) focused on discovering specific lesson types which promoted high performance in FD as well as FI students. Students at both ends of the FD/I continuum were assigned to deductive and example (more inductive) CALL lessons on the formation of participial phrases; pre- and post-test scores were used to test the effectiveness of each type of lesson. The findings indicated that the more FI students performed better with the deductive lesson and the more FD students performed better with the example lesson. Thus, it would appear that when given lessons matched to cognitive style, FD as well as FI students demonstrate high performance.
FD/I and Reading Comprehension

The research reviewed thus far has explored the relationship between FD/I and second language acquisition. This next section reviews what is known about the relationship of FD/I to reading comprehension. Unlike the cognitive style research previously discussed, these studies seem to have been conducted exclusively by researchers in elementary and secondary education. Furthermore, their subjects have been child, adolescent, and adult native speakers of English.

In a review of cognitive style and reading research, Rasinski (1983) observes that "one of the first overarching and consistent findings in this line of research has been that good readers tend to score higher on measures of field independence than poor readers" (p. 4). According to Rasinski, numerous studies have found a positive relationship between FI and general reading ability (e.g., Blaha & Chomin, 1982; Readence, Baldwin, Bean, & Disher, 1980; Smith & Standal, 1981). The same positive relationship was found for FI and specific reading skills such as utilizing semantic, syntactic, and graphophonic cues; recognizing grammatical relationships; employing predictive strategies by making use of background knowledge; and recalling and retelling the content of texts (e.g.,
Christiansen, Annesley, & Scott, 1980; Scott, Annesley, Maher, & Christiansen, 1980). Further evidence for the advantage FI students have over FD students in background knowledge utilization was cited in Spiro and Tirre (1980).

Rasinski's (1983) observations tend to be corroborated by the findings of studies not included in his review. The positive relationship of FI to reading comprehension was found in several studies of lower and upper elementary school children. Pitts and Thompson (1982) found that FI correlated with performance on a general reading comprehension test. Blaha (1982) found that FI consistently predicted variance not only for a test of general reading ability but also for basic skills tests of arithmetic concepts and computation. Dermott, McIntire, and Roberts (1979) found that FI correlated with performance on tests of specific reading skills (phoneme/grapheme relationships, semantic/syntactic clues, and word part clues). Rounds (1979) investigated the relationship between FD/I and grammatical awareness ("the ability to reflect upon the syntactic structure of sentences" p. 1), reading ability, and training in grammatical awareness. Children who were more FI were found to have higher levels of grammatical awareness and higher scores on two tests of reading ability (an eye-voice span test of use of grammatical cues in recall
and a series of cloze tests). Children with higher levels of grammatical awareness were found to be better readers, and FI was a factor in developing that awareness.

Not all reading researchers have obtained findings which support Rasinski's assertion, quoted above. For example, Roberge and Flexer (1984) investigated the relationship of FOI and level of operational development to performance on word knowledge and reading achievement tests. Their subjects were sixth, seventh, and eighth grade students. They found that FOI was not related to performance on reading measures; instead, a high level of operational development was associated with high performance on the reading tests. Roberge and Flexer concluded that these results support the claim (first asserted by Witkin and Goodenough, 1977) that "the ability to manipulate verbal materials is primarily a function of traditional school experience rather than individual differences in field dependence-independence" (1984, p. 232).

Blake (1985) conducted a study using sixth grade subjects to determine if FOI had any effect on the comprehension test scores for expository and literary text types. It was hypothesized that differences in cognitive processing (as determined by FOI) might contribute to significant variation in the comprehension of different text
types as measured by multiple choice tests. It was discovered that the students who worked with the literary text type attained higher comprehension test scores than the students who worked with the expository text type. However, no evidence emerged to support the conclusion that extreme FI or FD influenced the comprehension test scores of students assigned to either text type.

The above studies have investigated the relationship of FD/I to reading comprehension in tasks that are at a level particularly relevant to elementary and secondary school students. The next group of studies in this section was conducted using college level students as subjects; as such they involve learners, learning contexts, and reading tasks which are more directly related to those encountered in this study. In addition, they have explored the manner in which a student's background knowledge of content and text structure facilitates the organization of information in memory and thus affects comprehension and recall.

Spiro and Tirre (1980) conducted a study using college students to discover how FD/I might affect the ability to use background knowledge (or knowledge schemata) in comprehension. Half the sample read a passage about buying certain food items in a supermarket; the other half read a passage about buying the same food items in a restaurant.
The restaurant passage was thought to be more structured as the range of foods one can purchase at a restaurant is restricted to several specific categories (e.g., appetizer, entree, side order, beverage, dessert); thus, it is a more constrained context than that of a supermarket where the range of foods for purchase is wider and not so strictly categorized. It was hypothesized that comprehension of the restaurant passage would depend on the ability to recognize its more constrained organization, which in turn would require a cognitive predisposition for perceiving and organizing information in an articulated fashion. Thus, the more FI students were expected to demonstrate greater differences in the recall of the two passages than the more FD students. This expectation was confirmed; FI students' recall increased dramatically from the supermarket to restaurant texts, while FD students' recall increased very little from one to the other. Spiro and Tirre concluded that the recall of FI students was greatly enhanced by their ability to "detect the relevance and applicability of preexisting knowledge schemata and superimpose those structures on the text in interaction with the text's inherent structure" (1980, p. 207).

Several studies have examined the effects on comprehension of the manipulation of text structure and
training in study techniques. Wilcox, Richards, and Merrill (1977) studied the effects of text condensation and FD/I in the comprehension of a chapter from a psychology text. They found that the use of an abridged chapter (comprising, for the most part, superordinate information) improved the recall of the subordinate information included. They also discovered that FD students found this chapter to be clearer than an unedited chapter with a summary. In addition, FI students were moderately more successful in answering questions which required the application of learned information (from the text) to a previously unencountered situation (the post-test), a task which the authors saw as analogous to perceptual disembedding. Interestingly, the FI students found the unedited chapter clearer than the abridged chapter, perhaps because it left them greater latitude to impose their own structure upon it.

Annis (1979) studied the effects of FD/I and well organized versus scrambled text structure on the effectiveness of two study techniques: reading a passage only, and reading and taking notes. After assessing reading comprehension with free recall and fill-in-the-blank tests, Annis found that FI students were better able to recall information of high structural importance (abstract or general information, such as topic sentences) than FD
students, regardless of whether the passage they had studied was organized or scrambled. Furthermore, FD students were less able to analyze information in terms of high structural importance and low structural importance (concrete or specific information, such as examples and details). Finally, study technique did not seem to significantly affect the comprehension of FI or FD students of either organized or scrambled passages.

Along the same lines, Brooks and Dansereau (1981) conducted a study to determine if training in the use of knowledge schemata would facilitate the processing of an unfamiliar scientific text. They hypothesized that FI students, by virtue of their more articulated style of processing, would show greater gains on several comprehension measures (essay, short answer, multiple choice and cloze tests) as a result of training than FD students, who do not have as great an ability to process information in a structured manner. They obtained somewhat mixed results. The students in the trained group achieved higher scores than the untrained group on the essay test alone (both groups contained FI and FD students). However, FI students outperformed FD students on all the comprehension measures regardless of whether they had received training. Thus, while training in the use of knowledge schemata had
positive effects for FI and FD students on one measure, it appears from this study that cognitive style (FI) has more to do with performance on comprehension measures than training in schema utilization.

Smith and Standal (1981) obtained similar results from their study of FD/I and training in study techniques. Students were trained to study college level passages using one of two techniques: paraphrasing or mapping (a visual, diagramatic representation) salient ideas from the text. Like Brooks and Dansereau (1981), they found that FI students scored higher on a reading comprehension test than FD students regardless of training. What is of interest for this study is that Smith and Standal's reading comprehension test included subtests which assessed the ability to understand main ideas, as well as understand direct statements and make inferences. FI students scored higher than FD students on each of the subtests at a statistically significant level.

The research reviewed in this section appears to be fairly consistent in terms of the results obtained. The literature which focuses on the reading comprehension of elementary and secondary school children shows, with a few exceptions, that children who are more FI are more proficient readers than children who are more FD. The
research on adult learners presents similar results. Spiro and Tirre's (1980) study provides evidence to support the notion that FI is related to an ability to recognize and utilize schemata in comprehension and recall; thus, adults who are more FI have an advantage in reading comprehension over their more FD peers. The natural practical application of these findings is to attempt to improve the comprehension and recall of less FI, or FD, students by directly teaching schema utilization or other organizational study techniques. However, the research in the pragmatic use of Spiro and Tirre's findings generally showed that training in these techniques had little to no apparent effect on comprehension test performance; FI students continued to outperform FD students even when the FD students had received training in study techniques. Thus, this body of research confirms the expectation that FI is associated with and perhaps predictive of a high level of reading comprehension.

Schema Utilization in Reading Comprehension

In addition to FD/I, one of the learner variables which may affect the ability to find the main idea in reading is the prior knowledge an individual has of the subject matter being read. Because the use of prior or background knowledge has been widely recognized as a critical factor in
successful reading comprehension, this section reviews research in this area, particularly as it relates to the comprehension process of nonnative speakers of English.

The active use of background knowledge, a skill also known as schema utilization, is a pivotal element in current models of reading comprehension. During the time that reading comprehension has received empirical attention, three principal theories of the process have been offered: bottom-up, top-down, and interactive. Of these theories, the later two incorporate the notion that the reader's prior knowledge plays a part in comprehension. The earliest of these is the top-down theory put forth by Kenneth Goodman (1970a, 1970b). In what has become a classic paper, Goodman defined the prevailing theory of the time, the bottom-up model, as one in which "reading is a precise process [involving] exact, detailed, sequential perception and identification of letters, words, spelling patterns and large language units" (1970b, p. 259). In place of this "common sense notion" he offered an opposing view which has become known as the top-down model. Goodman saw reading as a "psycholinguistic guessing game" which involved "an interaction between thought and language" (1970b, p. 260). More specifically, Goodman proposed that reading comprehension occurs as the result of the interaction
between a reader's background knowledge and the writer's text. The process he described is a perpetual cycle comprising several steps: sampling, in which the reader relies on the redundancy of language to perceive and identify the fewest linguistic cues necessary to begin guessing at meaning; predicting, in which the reader tries to anticipate new information based on what is known from the selective sampling; testing, in which the reader checks to see if the prediction is compatible with semantic and syntactic cues in the text; and confirming, in which the predictions are either proven correct, are refined, or are rejected. As explained by Eskey (1986), the significant features of this model are that reading is viewed as a "reconstruction of meaning based on a skillful sampling of the text" (p. 12); linguistic redundancy is a tool for perceiving text more efficiently; background knowledge plays a crucial role in prediction; and reading should take place at a reasonable rate in which the units of processing are meaningful chunks of words, not single words or letters.

While the top-down model was recognized as a clear improvement over its predecessor, later theorists became uncomfortable with the emphasis the top-down model placed on interpretive skills almost to the exclusion of identification skills (Eskey, 1986). In place of the top-
down model they offered the interactive model (McClelland & Rumelhart, 1981; Rumelhart, 1977; Stanovich, 1980). This theory espouses the view that reading is a bidirectional process which requires a balance between the ongoing accurate perception of words and letters, and the higher level skills which permit interpretation in a larger context. Like the top-down model, a crucial part of this interaction is the role of background knowledge. In order to process text actively and efficiently, the reader must call upon what she or he already knows about language (e.g., phonological, syntactic, semantic, and pragmatic knowledge) and what she or he knows about the world (e.g., spatial, temporal, role, causal, and episodic knowledge). Thus, the interactive model sees comprehension as a process in which the use of background knowledge is essential. Further, the prerequisite knowledge for comprehending text comes from both the bottom-up and the top-down domains; in fact, efficient processing is dependent on a balance in the use of identification and interpretation skills.

The interactive model appears to have become the model of choice among many ESL researchers. As a result, a great deal of research has been done in the last five to ten years on the schema utilization of nonnative speakers of English. Because of the predominant focus on the schematic aspect of
the interactive model, a good many ESL researchers tend to speak in terms of a schema-theoretic approach to understanding ESL reading comprehension. Schemata are defined as "interacting knowledge structures stored in hierarchies in long term memory" (Carrell, 1983a, p. 82). In schema theory, comprehension is contingent upon being able to relate new or unfamiliar information to what is already known or stored in one's schematic structures. More specifically, this process "is guided by the principle that every input is mapped against existing schema and that all aspects of that schema must be compatible with the input information" (Carrell, 1983a, p. 82). Input in this case includes representations of form (e.g., letters, words, texts) as well as content (e.g., scenes, events, activities, concepts); these two types of schemata represent the bottom-up and top-down features, respectively, of the interactive model. As Carrell (1983a) explains:

Bottom-up processing ensures that the listener or reader will be sensitive to information that is novel or that does not fit her or his ongoing hypotheses about the content or structure of the text; top-down processing helps the listener or reader to resolve ambiguities or to select between alternative possible interpretations of the incoming data. (p. 82)
As this process continues during reading, new information is perpetually added to one's existing schemata and is reapplied to the comprehension of new material.

Carrell (1983a) identified two types of schemata which are particularly relevant to the reading comprehension of ESL students: formal and content schemata. Formal schemata comprise background knowledge of the "formal, rhetorical, organizational structures of different kinds of texts" (Carrell, 1983a, pp. 83-84). Carrell found evidence for the role of formal schemata in the comprehension of ESL learners in two studies. In one study, Carrell (1984b) investigated the effects of story structure on the recall of ESL learners. She found that the quantity of recall was improved when the rhetorical organization of the story conformed to the reader's schema for stories. Further, even when students read stories which violated typical story structure, they recalled the events of the story in a way which matched their schema for stories rather than the order in which the events were listed in text. In another study, Carrell (1984a) investigated the effects of different expository patterns on recall. She found that recall is improved when ESL readers recognize and utilize the rhetorical pattern of the text in the structuring of their written recalls. She also found that the more tightly
organized patterns of comparison, causation, and problem/solution tend to facilitate recall better than the more loosely organized pattern known as collection of descriptions. Finally, she found that the recalls of certain native language groups were differentially affected by the above rhetorical patterns.

The second type of schemata identified by Carrell (1983a) was content schemata, which comprise background knowledge about the subject matter or substance of a text. Numerous studies have demonstrated the necessity of adequate content schemata for reading comprehension. Steffensen, Joag-dev, and Anderson (1979) gave students from both America and India two letters to read; one was about an American wedding, and the other was about an Indian wedding. They found that the students read the passage on their native culture more rapidly, recalled more information from the native passage, and elaborated more on the content of the native passage in culturally appropriate ways; in addition, students distorted the cultural content of the foreign passage more than the native passage. Thus, adequate, culturally appropriate content schemata influenced what was recalled and how much was recalled.

Johnson (1981) studied the effects of English linguistic complexity and the cultural origin of narrative
texts on the reading comprehension of Iranian ESL students. Half of the students read unadapted English texts of an Iranian and an American folktale; the other half read the same texts in adapted or simplified English. Johnson found that the students recalled the situation and supporting details of an Iranian folktale better, regardless of whether they read the original or adapted version; further, they made more errors in the recall of both versions of the American story. Thus, the cultural origin of the text had a greater effect on the comprehension of the Iranian students than the semantic and syntactic complexity of the text.

In a similar study, Nelson (1987) gave Egyptian ESL students four paired readings; each pair was written on the same topic but used two cultural contexts (e.g., one pair of readings were the changing role of women in America and the changing role of women in Egypt). After testing their comprehension of all four pairs of passages, Nelson found that the students recalled the four passages written from the Egyptian cultural context better than those couched in American cultural terms.

In a later study, Johnson (1982) investigated the effects of teaching cultural content schemata on reading comprehension. ESL students were given a reading passage on Halloween two weeks after having participated in a city-wide
Halloween celebration. Students were tested for comprehension of familiar information (elements of Halloween the students had experienced at the celebration), and unfamiliar information (historical background on the holiday assumed to be obscure even for native speakers). Johnson found that the students recalled the familiar cultural material better than the unfamiliar information, presumably as a result of the schema-building which certainly occurred when the students attended the Halloween celebration.

Finally, Alderson and Urquhart (1985a, 1985b) studied the effect of students' academic majors on their comprehension of passages taken from their own and other academic majors. They classified their sample of ESL students into four areas: development administration, finance, and economics; engineering; liberal arts; and science and mathematics. After testing their comprehension of passages in each of these areas, Alderson and Urquhart found that background knowledge, as evidenced by the student's major, led to better comprehension of the passage which corresponded to that major.

The studies reviewed above validate the existence and influence of formal and content schemata in ESL reading comprehension. In the process of identifying the schemata which are relevant for comprehension, researchers have noted
that ESL learners seem to have special problems in schema utilization which are distinct from those encountered by native English speakers. These often lead to what Carrell (1984c) has termed a unidirectional mode of text processing in which the interference of one or more factors causes an overreliance on either the text-based (bottom-up) or the knowledge-based (top-down) mode of processing. Carrell identified five factors which could cause this problem: (a) the absence of either formal or content schemata to apply to written text, (b) the failure of texts to provide sufficient lexical cues to activate the student's existing schema, (c) linguistic and reading skill deficiencies, (d) culturally-based misconceptions of reading in English, and (e) individual differences in cognitive style. This final factor has obvious relevance to this study. Carrell (1984c) explains:

Text is an external stimulus with a structure; interactive reading requires that relevant internal knowledge structures be superimposed on the text. Those who are overly text-bound in reading situations may tend to be stimulus-bound in general. (p. 16)

The principal characteristic of a more FD cognitive style is that perception is dominated by the totality of a stimulus such that perception of discrete elements within the
totality is very difficult. One may infer from Carrell's explanation above that being "stimulus-bound in general" is analogous to being FD; therefore, a more FD cognitive style may contribute to text-boundedness in reading. This hypothesis is supported by the work reviewed earlier by Spiro and Tirre (1980), who found that FD students tend to underutilize their existing schemata, while FI students tend to apply their schemata rather systematically.

Like Carrell (1984c), Eskey (1986) points out that ESL students may have greater difficulty with text-based identification skills than native speakers. To illustrate this observation, Eskey cites a study by Cohen, Glassman, Rosenbaum-Cohen, Ferrara, and Fine (1980). Cohen et al. found that ESL students have difficulty in identification at the lexical level (a tendency to attribute only one meaning to a given word) and the lexical-syntactic level (comprehension of extended noun phrases), as well as with lexical rhetorical markers (assigning the correct meaning to transition words or logical connectors). Furthermore, Eskey notes that ESL students may well have a confidence problem when it comes to reading English and undoubtedly need overt support and encouragement from teachers in order to succeed.

Both Carrell (1984c) and Eskey (1986) have suggested that when an ESL reader resorts to unidirectional
processing, she or he tends to cling to processing in a linguistically literal, text-bound fashion; the student relies almost exclusively on bottom-up processing rather than allowing bottom-up and top-down processing to work interactively. However, recent research has questioned this conclusion. Wolff (1987) conducted a study which is analogous to one carried out by Carrell (1983b); Wolff obtained different results and came to different conclusions. Carrell (1983b) investigated the individual and interactive effects of three components of background knowledge (context, text transparency, and familiarity) on the comprehension of ESL students and native English speakers. Each component had two values: context/no context, text transparent/text opaque, and content familiar/content novel. After reading two passages in which the combination of the values of context and transparency was varied for a familiar and a novel text, students were asked to write recalls of each passage. The results of the study showed that all three components play a role in the comprehension and recall of native speakers; thus, their comprehension process was shown to be a balance between bottom-up and top-down processing. However, the processing of the ESL students did not reflect an interactive use of the two; in fact, ESL students did not appear to use either
bottom-up or top-down processing in an efficient manner. Carrell concluded that the ESL students were not able to move past the literal language of the text to begin "making the necessary connections between the text and the appropriate background information" (p. 200).

In Wolff's (1987) study, ESL students listened to stories on videotape that were either contextualized with an illustration on the screen or were presented without such an illustration. The students then recalled on tape the version of the story (illustrated or nonillustrated) each had seen; the recall was carried out in their first language (German). Wolff found that when students are permitted to recall a story using their first language, it becomes apparent that they rely on top-down strategies as a means of transcending their deficiencies in bottom-up processing. Thus, Wolff's results confirm those that Carrell found for the imbalance in ESL students between bottom-up and top-down processing; however, Wolff asserts that this imbalance is due to the student's overemphasis on top-down processing and is not because of the text-boundedness that Carrell concludes is the problem. Wolff's conclusion offers an interesting alternate perspective on the problem of unidirectional processing but does not account for the possibility that text-boundedness in a written text is
different from that in an oral text. To generalize Wolff's conclusion to reading it would be necessary to repeat the study using written rather than oral texts and recalls.

The discussion thus far has served to describe current models of the reading process and elaborate on the role of schema utilization in ESL reading comprehension. While schema research of this relatively general nature is abundant, little schema research has been done in relationship to the other two constructs in this study, FD/I and finding the main idea in reading. Two studies reviewed in the previous section have explored the relationship between FD/I and schema utilization (Brooks & Dansereau, 1981; Spiro & Tirre, 1980). In some of the FD/I and schema research, main idea questions are occasionally noted as being included in the instrument a researcher has used to test overall reading comprehension (e.g., Groebel, 1980; Smith & Standal, 1981). However, only one study was found which specifically examined the processing of main ideas from a schema-theoretic point of view. Kimmel and MacGinitie (1984) studied upper elementary school children who employ a perseverative text processing strategy; that is, their reading comprehension is impaired by "a particular type of overemphasis on top-down processing" (p. 163). Such readers tend to formulate hypotheses about the meaning of a
text after only a brief sampling of its material, and then persist in their initial interpretation despite incoming information to the contrary. The investigators found that children with this perseverative strategy had greater difficulty in comprehending main ideas in inductively organized paragraphs (main idea last) than in deductively organized paragraphs (main idea first). In other words, perseverative readers found deductive paragraphs easier to comprehend because "the reader's first hypothesis about the main point of the paragraph is more likely to be correct" (p. 165). Conversely, the main idea in an inductive paragraph was more difficult to understand because the perseverative reader would tend to accept the initial sentence as the main idea and would ignore the correct main idea encountered at the paragraph's end.

How or if this finding may be integrated with the research already reviewed here can only be speculated. Some FD/I studies have suggested that FI learners benefit from deductively structured lessons while FD learners benefit from inductively structured lessons (e.g., Abraham, 1985). At least one study has shown that FI learners are better able to activate and apply content schemata in reading (Spiro & Tirre, 1980). These two findings, taken with Kimmel and MacGinitie's findings, may suggest that FI
learners, who prefer a deductive mode of learning and are better able to utilize high-level schemata, may have greater difficulty in finding the main idea in inductive passages because they in fact overemphasize the top-down mode of processing. The complementary supposition might be that FD learners, who prefer an inductive mode of learning and are less able to utilize high-level schemata, would have less difficulty in finding the main idea in inductive passages. Needless to say, a great deal of research would need to be done to test the validity of such speculations.

Assessment of Background Knowledge

As mentioned earlier, background knowledge is being examined here as a predictor of performance on a main idea comprehension test. In order to operationalize background knowledge as a predictor, an instrument was developed to assess the student's prior knowledge of each topic included on the test. Carrell (1983a) notes that empirical investigations of the role of content schemata have assessed an individual's background knowledge in one of two ways. In the first method, the text is kept constant in form and content while the background knowledge given to two or more groups of subjects is experimentally controlled. In the second method, the text is again kept constant in form and
content, and preexisting differences in prior knowledge are assessed with an instrument of some kind. (The research design of this study was such that the second method of background knowledge assessment was used; further discussion of this instrument is included in the next chapter.)

Carrell (1983a) reports that experiments which attempt to measure a subject's prior knowledge have been performed mainly by first language researchers. A review of means for assessing background knowledge yielded several methods. Anderson, Reynolds, Schallert, and Goetz (1977) measured the prior knowledge of their subjects with an autobiographical inventory administered after the subjects completed the reading task in the study. The inventory contained items "intended to tap matters which could be expected to relate to the interpretations given to the passages" (p. 373). The answers given on these items were used as predictors in a multiple regression analysis of performance on corresponding comprehension test items. However, the report does not state how the inventory answers were evaluated (some were yes/no questions, others more open-ended) or quantified for use in the regression.

Stevens (1980) used a 100 item multiple choice test on 25 factual topics (four questions per topic) to determine areas in which students had high and low knowledge. High
knowledge topics were defined as those for which the student answered all four questions correctly; low knowledge topics were defined as those for which the student answered no questions correctly. Each student was then assigned two reading passages, one on a topic from a high knowledge area and the other on a topic from a low knowledge area. The effect of background knowledge was then evaluated based on each student's performance on comprehension questions for the passages each had read.

Following the lead of Stevens (1980), Levine and Haus (1985) assessed prior knowledge of baseball by administering a 9-item multiple choice questionnaire derived from an analysis of the single reading passage used in the study. Students were then identified for participation in the study as either having high knowledge (8 or 9 items correct) or low knowledge (1 to 3 items correct). The effect of content schemata was evaluated by entering background knowledge as an independent variable into an analysis of variance of performance on a comprehension test.

Zakaluk, Samuels, and Taylor (1986) assessed background knowledge using a simple free association task. Students were given a stimulus word, a key word, or a phrase which encompassed the main idea of a topic. They then had three minutes to write down as many words as they could think of
in association with the stimulus word. The quantity and quality of their background knowledge were evaluated by assigning points to each set of responses; the totalled points for each free association were then compared to ranges of points which indicated low, average, and high prior knowledge. A more elaborate and time consuming version of this procedure, called PReP (Pre REading Plan), was developed by Langer (1984) for use as a classroom prereading exercise as well as an empirical measure of background knowledge.

Each of the above methods for assessing background knowledge is problematic in some way when evaluated in terms of the research situation anticipated for this study. Factors such as time available for student testing, ease of administration and scoring, and adaptability of results for statistical analysis make some of the approaches discussed above impractical. For example, the multiple choice tests used by Stevens (1980) and Levine and Haus (1985) seem to be a good way to determine prior knowledge but could be very time consuming if the reading task included more than one topic (as is the case in this study). Apart from these considerations, the testing of ESL learners presents some special constraints imposed by the linguistic proficiency level of the students. The accuracy of an assessment method
requiring productive competency in English (i.e. written English) would most certainly be affected by the proficiency level of the subjects involved. For example, the open-ended questions in the Anderson et al. (1977) autobiographical inventory presuppose an intermediate to high level of writing competency on the part of the ESL student. The accuracy with which the free association task (described by Zakaluk et al., 1986) reflects an ESL student's background knowledge is most certainly predicated on the scope of the student's vocabulary.

In addition to these drawbacks, at least one potential problem could arise in the use of these measures which has not been accounted for in the studies above. It is conceivable that the use of any of these instruments would begin to activate students' schemata prior to the reading task, thus allowing them to anticipate and begin making predictions about what they were going to read. This is unquestionably desirable in a classroom prereading exercise but may be less so in an empirical setting if the researcher wants to control advance knowledge of the purpose of the reading task, as is the case in this study. Thus, while the above methods have proven themselves useful in a first language context, they present potential difficulties for use in the ESL research setting of this study.
Main Idea Comprehension

In contrast to the areas of study already reviewed (FD/I and schema theory), research into the subconstruct of reading comprehension loosely known as finding the main idea has surfaced rather recently. However, given its relative novelty, a sizeable amount of research has been done on main idea, principally by first language researchers using native English speakers as subjects.

The most salient and pressing concern of researchers in this area is the definition and operationalization of the term main idea. No one seems to know precisely what is meant by main idea or its sibling, important information; further, no one has as yet been able to construct a model of main idea comprehension which explains how main ideas are perceived and/or constructed during the reading process. In short, the fundamental questions involved in this line of research are as yet equivocal and unresolved. The result is that main idea researchers are necessarily preoccupied at this time with product rather than process, with describing the results of main idea tasks rather than defining the cognitive and metacognitive processes which make them possible.

As might be expected, much of the extant research has focused on defining the term main idea; all of the research
to date has been confounded to some extent by the lack of a uniform definition. Three studies have addressed this issue by attempting to discover what fluent readers' perceptions of the main ideas are in assigned readings. The studies were performed using sixth graders (Moore, Cunningham, & Rudisill, 1983); tenth graders (Graham, Cunningham, & Moore, 1985); and teachers and undergraduate education majors (Cunningham, Graham, Moore, & Moore, 1984). From the perceptions of these three groups of informants, Cunningham and Moore (1986) delineated nine types of main idea responses. These nine types are differentiated by the amount of invention, or generation of new information, required of the reader to complete the response. For example, main idea questions which require the student to pick out a key word, identify an appropriate title, identify a topic sentence or thesis statement, or summarize the most superordinate points in a passage can be answered by pulling this information directly from text; very little creative generation is required of the student. Main idea questions which require the student to produce a topic phrase (a phrase which labels the topic of a passage without specifying its content) or the gist of a passage necessitate some generation of ideas beyond those which are explicit in the passage. Finally, main idea questions which elicit the
production of a topic issue (a label which establishes a conceptual context for the passage), a theme, or an interpretational summary require a great deal of invention from the student. In addition to advocating the recognition of nine types of main idea responses, Cunningham and Moore (1986) suggest that the evaluation of a student's response to a main idea question should depend on how well it meets the requirements of a particular main idea type. For example, a student might phrase her or his perception of the main idea as a topic issue, which according to Cunningham and Moore requires a high level of generative ability to produce; however, if the teacher was expecting a different type of response, such as a thesis statement (which requires less invention to produce), the student's answer might well be counted wrong, even though the topic issue represents a higher level of generative ability in the student.

Williams (1986) and Winograd and Bridge (1986) report that one cause for confusion in the definition of main idea lies in the distinction between textually important information and contextually important information. Textually important information is that which is considered important in the eyes of the author and which is generally elicited in classroom testing situations. Contextually important information is that which is considered important
by the reader and which is determined by the idiosyncratic purpose the reader may have for reading. Again, the definition of main idea becomes a matter of perspective: the main idea according to whom?

Williams (1986) cites a second cause of confusion in the definition of main idea as the differences among text types; what is considered important varies depending on whether the text is narrative or expository. Williams reviewed three approaches to analyzing text processing, each of which attempts to specify what information is important in expository text: Kintsch and van Dijk's (1978) model of proposition processing, Meyer's (1985) hierarchical content structures, and Kieras' (1985) model of thematic processing. Williams saw each model as a valuable contribution to a better understanding of text processing, but found each approach lacking when it came to defining what constitutes important information.

The above discussion has touched on some of the difficulties inherent in defining what main ideas are; even greater difficulties, and thus fewer studies, have been encountered in describing how main idea comprehension takes place. Winograd and Bridge (1986) have suggested that this ability is acquired developmentally in first language readers, noting that as they mature, "they increase their
knowledge of the world and of text structure and also become more efficient at identifying the methods authors use to mark important information" (p. 25). Brown and Smiley (1977) studied this developmental dimension by testing the ability of elementary, secondary, and university students to identify important information. While all age groups recalled superordinate information more frequently than subordinate information, only the secondary and university students were able to report why the information they recalled was important. Williams (1986) deduces from these results that "tacit and intuitive knowledge...is the basis for the ability to recall information, but...the ability to identify important information is more a matter of conscious awareness—in a sense, it represents metacognitive knowledge" (p. 21).

Empirical work which seeks to illuminate the process of comprehending main ideas is just beginning to surface. One study which has attempted to elucidate the process of main idea comprehension used think aloud protocols to discover how expert readers constructed main ideas from difficult texts (Afflerbach & Johnston, 1986). They found that their readers employed three strategies in constructing main ideas. One strategy was to generate a hypothesis of the main idea before reading based on prior knowledge gleaned
from skimming or from previous experience with the topic; the hypothesis was then refined as needed as reading of the text progressed. A second strategy, called crunching, represented a more passive approach in which the reader simply read and waited for some automated cognitive process to act on the information entered into working memory. A third strategy was to reread the text as necessary to pinpoint key words or phrases and subsequently use these terms to construct a main idea. Expert readers were found to employ additional comprehension techniques such as making use of knowledge-based and text structure cues to activate relevant schemata, hypothesizing about the author's intent and biases in writing, and remaining flexible in determining what was important in text. Furthermore, expert readers consistently monitored their main idea construction process at the end of meaning units and at junctures where comprehension was stalled or problematic. Finally, Afflerbach and Johnston found that affective factors such as the reader's attitudes, opinions, and beliefs influenced how a text was read; additionally, the reader's level of prior knowledge affected the extent to which she or he concurred with the author.

Though studies like Afflerbach and Johnston's (1986) are beginning to show how schema utilization enters into the
ability to identify main ideas, there appears to have been no attempt as yet to determine the effects of FD/I on the process. Consequently, and as is the case with the bulk of main idea research, more good questions are raised than can be supplied with good answers. For example, is main idea comprehension an analytical or a holistic process? Afflerbach and Johnston's study seems to show that finding the main idea involves both analytical and holistic processing. The third strategy they identified (rereading the text to find key terms) seems analogous to disembedding, an analytical skill that FI learners characteristically employ in processing information. The second strategy they identified, crunching, is a more passive and possibly intuitive approach to information processing which may be analogous to holistic perception, a principal characteristic of FD learners. If both cognitive styles can be associated with main idea comprehension, do they both predict high performance on a test of finding the main idea? Or is test performance influenced by an interaction between cognitive style and such factors as text organization or position of the main idea in text? Afflerbach and Johnston also found that expert readers made efficient use of formal and content schemata; Spiro and Tirre (1980) discovered a similar connection between FI readers and schema utilization. From
these findings, is it possible to deduce that FI students are better able to identify main ideas in part because of a greater ability to utilize formal and content schemata? Brown and Smiley's (1977) study revealed a developmental factor in the ability to consciously identify important information; young children could recall main ideas but could not identify why they were important. Assuming these children followed the developmental pattern known for FD/I and were thus more FD than the older subjects, is it possible that they are analogous to FD adults learning a second language? In other words, could a more FD cognitive style be related to a tendency to recall information but not be able to identify it as a main idea? Likewise, are the adolescents and adults in Brown and Smiley's study, who were presumably more FI than the children, analogous to FI adult second language learners? Could a more FI cognitive style be related to an ability to recall and consciously identify main ideas in a second language? These are all questions which the research community has yet to answer.

Assessment of Main Idea Comprehension

The method used to assess main idea comprehension in this study was developed after reviewing the ways in which other researchers accomplished the same objective. In this
review, method of assessment refers to the type of test question or task which the student was asked to complete after reading a passage.

Numerous methods have been used to assess the comprehension of main ideas. Following Cunningham and Moore's (1986) typology of main idea types, they can be viewed on a continuum of the extent to which they require the use of selective versus generative production. For example, tasks such as multiple choice questions (Van Blaricom & White, 1976) and title selection tasks (Williams, Taylor, & Ganger, 1981) simply require the student to select the main idea or an appropriate title for a passage from several choices. Rating tasks (Brown & Smiley, 1977) entail more involved selective production. The student reads a passage and crosses out a quarter of the least important sentences; the exercise is repeated until approximately one fourth of the original number of sentences remain. These sentences are taken as the information the student considers most important in the passage.

Other types of tasks require the student to generate main ideas after reading a passage. For example, in a cued recall (Wilhite, 1984) a subject is asked a question "generated by replacing segments of sentences presented in the passages with interrogatory terms" (p. 44). The subject
then records her or his recall of the main point prompted by the cue question. Free recall (Brown & Smiley, 1978) requires greater generative production; students are told to write down the gist of the passage in as many words as necessary. Finally, summaries (Brown & Day, 1983; Garner, 1982; Winograd, 1984) also require great generative production from the student; summaries can range from one sentence to several paragraphs depending on the length and complexity of the passage.

The choice of a methodology for assessing main idea comprehension in this study was subject to constraints similar to those for choosing a method to assess background knowledge. A principal concern was the linguistic proficiency of the students being tested. The validity of results from both the selective and generative types of main idea tasks could be limited by the level of language required for students to comprehend and to produce responses, respectively. Difficulties with the use of selective measures such as multiple choice questions could arise if students could not comprehend the language of the choices; similarly, problems could result from the use of generative measures if students did not have the production skills to express themselves accurately in writing. A further consideration was the amount of time which could be
allotted for assessing comprehension. Generative tasks could require more time for completion than selective tasks; the time needed for completion could also vary unpredictably among students. In the end, the choice of assessment methodology reflected a balance between instrument validity and experimental practicality. More will be said about the assessment of main idea comprehension in this study in the next chapter.

Summary

The research reviewed in this chapter has come from several interrelated categories: FD/I and second language acquisition, FD/I and reading comprehension, schema utilization in ESL reading comprehension, and main idea comprehension.

Many researchers in FD/I and second language acquisition have found that students who are more FI are more proficient in numerous second language tasks than students who are more FD. However, other researchers have found a lack of evidence to support this assertion, and no one seems to have identified tasks in which FD learners are more proficient than FI learners. As long as such inconsistencies and gaps are inherent in the literature, the generalization of findings is problematic; it is not yet
possible to unequivocally state that FI learners are more successful than FD learners in second language acquisition.

Researchers in FD/I and reading comprehension have generally worked with native English speakers from a broad age spectrum. Those who have studied elementary and secondary school children have found with little exception that FI learners make better readers. Those who have studied adults have also found that FI learners are more proficient readers, particularly in their ability to utilize content and formal schemata in the processing of text. Attempts to help adult FD learners compensate for a less FI cognitive style (such as training in study techniques) have not shown great success; it appears that FI learners continue to outperform FD learners on proficiency measures regardless of any special training the FD learners receive.

Research in ESL reading comprehension has focused to a great extent on schema utilization. Researchers have found that content and formal schemata both play a significant role in facilitating comprehension in ESL readers. They have also found that ESL readers may be hampered by an overreliance on either bottom-up or top-down processing strategies. Beyond such relatively general research in schema utilization, little work has been done to relate the schema utilization of ESL students to FD/I; however, the
research done on adult native English speakers (noted above) has shown that FI learners are more active and efficient in schema utilization than FD learners. It appears that no research has yet been done to relate the schema utilization of ESL learners to main idea comprehension; in fact, this particular research question has received scant attention even from investigators of native English speakers. Despite these gaps in the literature, it has yielded numerous methods of measuring prior knowledge which have suggested several important considerations in the development of such a method in this study.

Main idea comprehension research is still very much at the fledgling stage in its development. It is currently hampered by the lack of a definition for the terms "main idea" and "important information" and is only beginning to examine how the construct is operationalized in the comprehension process. Consequently, little work has addressed the role of schema utilization in main idea comprehension, and apparently no work has been done on the potential role of cognitive style (FD/I) in main idea comprehension. In spite of these theoretical problems, main idea comprehension has long been an important part of reading instruction and has been measured using a variety of methods; these methods differ depending on whether the student is required to select or generate the main idea.
METHOD

The first section of this chapter provides a description of the subjects who participated in this study. The second section discusses the tests and other information-gathering instruments used in the collection of data. The final section deals with the testing procedures used.

Subjects

The forty-nine subjects in this study were drawn from the high intermediate and advanced levels of the Intensive English and Orientation Program (IEOP) at Iowa State University. The subjects in these levels who had taken the Test of English as a Foreign Language, or TOEFL (n=43), had scores ranging from 360 to 583, with a mean of 480.47. However, not all the subjects who had taken the TOEFL had reading section scores on record; the reported reading section scores (n=38) fell in a range from 30 to 61, with a mean of 48.08. Twenty-six of the subjects were taken from the second eight-week IEOP session of Spring 1987; the other twenty-three subjects came from the first eight-week IEOP session of Fall 1987. Students from the lower levels in IEOP were not sampled because of their presumed lower level of reading proficiency; these students were not expected to
have the reading proficiency necessary to process the length, vocabulary, and grammatical and rhetorical complexity of the reading test passages.

Some descriptive characteristics of the sample worth noting are the subjects' classification as graduates or undergraduates, their first languages, and their academic majors. Table 1 shows how the sample was divided using these three characteristics. Of the forty-nine subjects, there were twenty-eight graduate students and twenty-one undergraduate students. Three major first language groups were evident in the sample: Arabic (n=10), Japanese (n=15), and Spanish (n=10) speakers. The remainder of the sample (n=14) comprised miscellaneous speakers, mostly Asian. (For a complete list of the first languages represented in the sample, see Appendix A.) Of further note is the proportion of graduate to undergraduate students within the four language groups. While this proportion is fairly even for the Arabic, Japanese, and miscellaneous language groups, all of the Spanish speakers were graduate students.

The subjects in the sample represented a wide variety of academic majors. These were assigned to four major subject areas similar to those constructed by Alderson and Urquhart (1985a, 1985b): business/economics, engineering, science/mathematics, and liberal arts. (For a complete list
TABLE 1. Distribution of subjects by language group, graduate/undergraduate status, and major academic area

| Language group | G/U | Major academic area | | | |
|---|---|---|---|---|
| Arabic | G=5 | 0 | 1 | 3 | 1 |
| | U=5 | 1 | 3 | 1 | 0 |
| Japanese | G=7 | 2 | 0 | 1 | 3 |
| | U=8 | 2 | 2 | 1 | 1 |
| Spanish | G=10 | 0 | 0 | 9 | 1 |
| | U=0 | 0 | 0 | 0 | 0 |
| Misc. | G=6 | 0 | 2 | 2 | 2 |
| | U=8 | 4 | 1 | 3 | 0 |
| Total | 49 | 9 | 9 | 20 | 8 |

*a* Bus. = business/economics; Eng. = engineering; Sci. = science/mathematics; Lib. = liberal arts.

*b* Three Japanese students reported their majors as undecided and so do not appear in the academic major columns.

Overall, nearly half of the subjects (44%) reported majors in science or mathematics; the remaining subjects were fairly evenly distributed among the other three
TABLE 2. Distribution of subjects by language group and major academic area

<table>
<thead>
<tr>
<th>Language group</th>
<th>Bus. a</th>
<th>Eng. a</th>
<th>Sci. a</th>
<th>Lib. a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Japanese b</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Spanish</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Misc.</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9</strong></td>
<td><strong>9</strong></td>
<td><strong>20</strong></td>
<td><strong>8</strong></td>
</tr>
</tbody>
</table>

aBus. = business/economics; Eng. = engineering; Sci. = science/mathematics; Lib. = liberal arts.

bThree Japanese students reported their majors as undecided and so do not appear in the academic major columns.

academic areas. Several patterns were noted in the majors of the subjects from the three principal language groups in the sample. The majority of the Arabic-speaking subjects (eight out of ten) were evenly divided between engineering and science/mathematics majors; of the remaining two subjects one was a business/economics major, the other a liberal arts major. This pattern was somewhat reversed for the Japanese speakers for whom only one third (n=5) were enrolled in engineering and science/mathematics. The two-thirds majority (n=10) were evenly distributed between
business/economics and liberal arts majors. Nearly all of the Spanish-speaking subjects (nine out of ten) were enrolled as science/mathematics majors. As such, they accounted for 45% of the entire sample majoring in science/mathematics. No particular pattern was noted for majors of the subjects who were assigned to the miscellaneous language group; perhaps all that can be said about this group is that the science/mathematics category claimed the most subjects while liberal arts claimed the least.

Materials

Three principal test instruments were used to collect the data. The Group Embedded Figures Test, or GEFT (Oltman, Raskin, & Witkin, 1971), was used to test for field dependence/independence. Two other instruments, a background knowledge survey and a reading test for finding the main idea, were developed by the investigator. A personal data form was also developed to collect information such as first language, academic major, graduate or undergraduate status, and amount of experience in the study of English (see Appendix C). Finally, measures of the subjects' reading proficiency were obtained from their reading section scores from the TOEFL and the Michigan English Placement Test.
Group Embedded Figures Test

The GEFT was used to determine the extent of each subject's field independence. This test consists of eighteen complex geometric figures in which simple figures are embedded. The task is for the subject to demonstrate perception of the simple figure by tracing it in pencil. The subject has a limited amount of time to complete the eighteen items. When the test is scored, only those items where the subject has identified the simple form clearly and completely are counted as correct. A subject who scores high on the GEFT is said to be relatively more field independent; conversely, a subject who scores low is said to be relatively more field dependent.

Background Knowledge Survey

Researchers in reading, particularly those interested in schema theory, have shown that a subject's level of background knowledge contributes positively to her or his level of reading comprehension. Thus, it was important to find out how much previous knowledge the subjects of this study had of the topics included in the main idea test.

A number of practical constraints were considered in the development of this instrument. First, the information would have to be collected prior to the administration of the main idea test. To minimize the possibility of missing
data, it seemed best to distribute the survey during the same period as that of the main idea test. This led to an additional requirement: economy of time. Because the class time available for data collection would be limited, the length of time spent on the survey would have to be minimal.

Beyond these practical requirements, some more substantive issues were examined. In order to indicate level of background knowledge, subjects would obviously need to be told what the topics in the main idea test were; at the same time, the detail of these descriptions would have to be carefully modulated. Enough information would have to be given for the subjects to place each topic within their own conceptual frameworks, and yet the information would have to be restricted so as not to give away the main idea of the passage.

Once this information about a topic was presented, the form of response given by the subjects would have to fulfill two requirements. First, the response must allow the subjects to report their background knowledge as accurately as possible. Second, it must be in a form that could be used in the statistical analysis of the data.

In the end, the topic information for each background knowledge item was drawn from the title of its corresponding reading passage. These titles were quite general; they
stated the topic discussed in each reading passage but gave no indication of the author's point of view or of the passage's main idea.

Several methods for quantifying background knowledge were discussed in the previous chapter (e.g., autobiographical questions, multiple choice tests, free association tasks). Because each of these methods seemed problematic for use with the learners and testing situation in this study, other methods were considered. One of these was a multiple choice format in which the subjects would choose a statement describing their level of background knowledge. These statements would be weighted in some fashion so as to provide a more objective picture of the subjects' knowledge of each topic. However, variation in the nature of the topics made it difficult to standardize the descriptive statements and the weights assigned to each statement. For example, a descriptive statement of background knowledge for the passage "Using a Computer" could be: "I know a lot about this topic because my major is computer science." This statement of background knowledge recognizes a correspondence between a specific academic major and the topic. This correspondence could not be stated as clearly for the passage "Changes in the Structure of the American Family". Being a sociology or anthropology
major would not necessarily endow a subject with great knowledge of the recent changes in American family life, whereas being a computer science major most certainly would lead to great knowledge of computer use. Because of such practical problems, this method of assessment was also rejected.

The difficulties associated with the above assessment options led to the development of a more objective survey format (see Appendix C). In this case, the subjects would indicate their level of background knowledge by circling a number on a scale from one to five. A "1" indicates no background knowledge, a "2" shows some but "not very much" knowledge, a "3" communicates "an average amount" of knowledge, and so on.

There seemed to be a number of advantages inherent in this approach. The use of a numerical scale would mean that only minimal reading would be required for a subject to fill out the survey, thus reducing the time necessary for completion. It was hoped that a numerical scale would elicit the subjects' perception of their background knowledge with the greatest clarity; the reliance of the other methods on textual responses might increase the chances of misinterpretation on the part of the subject as well as that of the investigator. Finally, the data
obtained would be in a form readily usable in statistical analysis.

**Main Idea Test**

Field dependence/independence is a cognitive style whereby the subject is able, to a greater or lesser extent, to disembed discrete parts from a complex whole. To test whether the extent of a subject's field independence contributes to finding the main idea in reading, it seemed necessary to provide a sufficiently complex whole (a reading passage) from which to disembed the discrete part (the main idea). Thus, it followed that the passages in the main idea test should be made challenging in length and level of language; otherwise, it would be possible for the subjects to make intelligent guesses about the main idea without really understanding the content of the passage. For example, if the subjects in this study were given a single paragraph and were asked to find the main idea, they might come up with the correct answer by guessing the first sentence of the paragraph. This could be deduced by recalling from English class that the main idea is usually found in the topic sentence, which often occurs at the beginning of the paragraph. They might also guess the main idea by process of elimination; a solitary paragraph usually has fewer sentences or ideas to consider for the main idea.
than an entire passage. The point is that in decreasing the length of the reading passage, the complexity of the "field" is in most cases significantly reduced. This is not to say that the difficulty of a passage is determined by its length; rhetorical, grammatical, and lexical complexity are among the other factors which must be considered. However, when these factors are controlled to bring the level of a passage into line with the reading proficiency of high intermediate and advanced ESL students, a longer passage can provide a more complex field than a single paragraph. Without the provision of a complex field, it would be difficult to know whether performance in finding the main idea was influenced by the subject's level of field independence or was produced by an educated guess.

With these factors to consider, the objective was to acquire a reading test in which the task was to find the main idea in a passage of about 400 words. Since no such test could be located in the literature, it was necessary to develop a reading test specifically for this study. This was accomplished in three main steps: (1) test development, (2) test piloting on nonnative speakers similar to those used in the study, and (3) review of the test by native speakers of English.
Test Development The first step in developing the test was to search authentic sources for reading passages suitable as test items. Several criteria were used to determine the usefulness of a given passage. An appropriate length was estimated to be approximately 400 words, so passages of between 350 and 450 words were sought. As most authentic passages are rarely that short, many potential test items were excerpted from an original longer piece. Another obvious requirement was that the passage indeed have a main idea. A surprising amount of published writing takes the form of a collection of descriptions or ideas apparently lacking an identifiable main idea. It was also necessary that the final test comprise passages which contained main ideas in more than one position in the text. That is, some passages would be chosen with the main idea in the initial position, some with the main idea in a more medial position, and some with the main idea in a final position. The last criterion was that the lexical, grammatical, and rhetorical complexity of the passage be at a level comprehensible to the subjects in the study. Failing this, the passage had to be able to withstand editing and revision to bring it more into line with the proficiency of the subjects.

It was difficult to find passages that met these criteria; it was even more difficult to find passages with
content that would be appropriate for the subjects taking the test. One concern was that the passages represent a broad spectrum of topic areas. As the subjects were university students, it seemed appropriate to choose articles from a variety of academic areas. While it was important that a prospective passage embody the content and language used in the field of study, it had to stop short of being so technical or using such profession-specific vocabulary that nonspecialists would find it incomprehensible. Passages meeting these criteria were found mainly in periodicals published for an audience which included nonspecialists such as The Futurist, Science Digest, and Human Nature. Such publications yielded a selection of passages from a range of contemporary research topics. About half of the topics included in the final version of the test were science and technology-oriented (such as medical research and computer science), while the other half represented more humanistic areas (such as sociology and history).

It was estimated that about ten passages would be needed for the final version of the test. This number seemed to be the minimum necessary to judge a subject's performance in finding the main idea. Twice as many items would obviously have been more desirable, but the limit on
how much reading a subject could reasonably be expected to accomplish in a test situation precluded the addition of more items. It was expected that the subjects would find the reading passages challenging and that it would take considerable concentration and effort for them to comprehend the content and find the main idea. Test fatigue was thus an important factor in the decision to limit the number of test items.

Once the ten reading passages had been assembled, they were edited and revised as necessary to adapt them for the proficiency level of the subjects. This often required the simplification of grammatical structures, the insertion of transition words or other rhetorical markers, and the omission of material which became extraneous when it did not support the main idea of the passage. Such omissions were made when an original source was excerpted; in these instances material was occasionally drawn from noncontiguous parts of a source, and some material was eliminated to ensure continuity in the adapted passage. After these revisions the length of the passages ranged from 294 to 474 words, with a mean length of 381 words. Titles were used to sustain the authenticity of the reading task but were modified in some cases so as to provide only topical information. The modifications were made because in some
passages the original title gave away the main idea, effectively relieving the subject of the necessity of reading the passage at all.

Another issue to be addressed in the development of the test was the glossing of vocabulary. Even after the initial editing and simplification of some passages, it was apparent that key terms needed to be defined in order to facilitate comprehension. Although intelligent guesses could be made about the vocabulary that the subjects would find troublesome, it was important to be more precise. To determine the vocabulary that required definition, the test was piloted on subjects similar to the study sample (discussed in the section "Test Piloting") and submitted for review to a jury of TESL instructors and graduate students (discussed in the section "Test Review").

The remaining task in the preparation of the test was to arrive at some means of eliciting the subject's perception of the main idea. For the sake of objectivity, it was decided that a multiple choice format would be used. Four choices would be offered for each reading passage: one correct answer and three distractors. Steps were taken to make the distractors plausible in the eyes of the subjects. First, the content of the distractors was drawn directly from the passages, often comprising details and supporting
ideas found in single paragraphs. Secondly, no distractors contradicted the information in the passages since choices that were obviously false could be easily rejected. Third, all choices for a particular main idea were written so that they were the same length when printed; length was controlled to prevent a subject from guessing an answer because it stood out visually. Fourth, the position of the correct answer among the distractors was varied from item to item. Finally, the choices were not prefaced by either a letter or number, as is common for most multiple choice tests. This was done in an effort to discourage the practice some test-wise subjects make of deducing an answer based on a pattern of previously chosen letters or numbers.

As the test began to take shape, some decisions were made regarding its administration. One of these was that the amount of time for reading the passage and choosing the main idea would be limited. This was considered necessary so as to discourage the tendency of some subjects to painstakingly read at a word-by-word rate. Another decision concerned the ordering of items in the test. Ideally, the presentation of the passages would vary on different forms of the test so that subjects would not be working the items in the same order. The advantage of such variation would be to prevent poor performance on the same items which could
result as students tired toward the end of the test. However, it was not possible to put this measure into practice because the passages varied in their length. This produced time allotments which varied among the passages. Thus, in the interest of controlling the timing of the reading, it was necessary for the subjects to work the items in a uniform order.

At this point in the process, a number of aspects of the test needed to be reexamined. One concern was the construction of the multiple choice questions for each passage. Did the main idea statement accurately reflect that of the passage? Were the distractors plausible and yet distinct from the correct choice? Another aspect needing further input was the selection of vocabulary to be glossed in each passage. What terms would prove difficult for the subjects and possibly crucial for their comprehension of the passages? Finally, the time allotment for reading each passage and answering the multiple choice item needed to be established.

To obtain feedback in these areas the test was submitted to two audiences. It was first piloted on ESL students whose reading proficiency was similar to that of the study subjects. With the information gained from this exercise, the test was revised and then presented to a jury
of TESL instructors and graduate students. This provided the feedback used to produce the final version of the test.

Test Piloting The most desirable group to pilot the test on would have been students at the same proficiency level as those of the study sample. Unfortunately, the small number of students enrolled in the upper levels of IEOP were needed for the actual study and so were not candidates for pilot testing. The next best pilot group was students enrolled in English 100B, a composition class for nonnative speakers at a level only slightly higher than that of the IEOP students in the study sample.

With the permission of their instructors, twenty-two students from two sections of English 100B were asked to take the main idea test. Due to the limited amount of class time which could be devoted to this exercise, each section took half of the main idea test. Thus, the subsequent modifications of the reading passages and the multiple choice items were based on the input from eleven students per item.

One objective in piloting the test on nonnative speakers was to get some approximation of how much time would be required for them to read each passage and find the main idea. The students were asked to read for comprehension, not necessarily speed, and raise their hands
when they had finished. The same procedure was employed to
determine the time interval needed for the students to read
each multiple choice item and select the main idea. The
ranges of times gathered for each passage and multiple
choice item were then used to help determine the time
allowed for each task on the final version of the test. The
times of the average and slower readers among the pilot
subjects were assumed to approximate the reading rates of
the IEOP subjects. This assumption was supported by a
personal communication from one IEOP teacher; she related
that her students (who were prospective study subjects) were
reading at the slower end of the range expected for high
intermediate and advanced students. Thus, a conservative
reading rate of seventy-five words per minute was assumed
for the main idea test; the time allowed for each passage
was calculated from this rate and varied depending on the
length of the passage.

A second objective for the pilot test was to discover
which vocabulary items were troublesome for nonnative
speakers. The students were asked to circle words they did
not understand in the course of reading each passage. These
terms were compiled later and used in the determination of
vocabulary to be glossed in the final test version.
A third objective in piloting the test was to gather information to help fine-tune the multiple choice items. As part of the test scoring, item information was collected to see if the pilot subjects selected the main idea choice with greater frequency than that of the distractors. One outcome of this analysis was the elimination of two passages from the test. The item results for these two passages indicated that the pilot subjects were generally unable to identify the main idea; it was concluded that the nature of the passage and the choices presented for the main idea was such that the items would be unreliable. Two more passages were found to replace those which were eliminated. Care was taken to make sure that a main idea was more clearly in evidence for these passages than for the ones they replaced.

The item information was also used to help improve the distractors presented in each multiple choice item. Some distractors were apparently too plausible to the pilot subjects. The distractors that were chosen with a higher frequency than the correct answer were revised to provide a clearer contrast to the main idea.

Test Review In addition to the information gleaned from pilot testing, it was important to get feedback from native speakers. For example, it was necessary to confirm that the wording of the main idea in each multiple choice
item accurately reflected that of the passage. A jury of seven TESL instructors and two graduate students who reviewed the test returned feedback on this and other aspects of the test. They provided comments on the relative difficulty of passages, suggested vocabulary that needed to be glossed, and gave constructive criticism on the composition of the answers and distractors of the multiple choice items.

Each of the instructors and graduate students were given a copy of the test and asked to give feedback in several areas. First, they were asked to circle vocabulary items which they thought were problematic for the subjects of the study. Secondly, before looking at the multiple choice items, they were asked to write in their own words their perception of the main idea. Finally, after recording their own assessment, they were asked to select the answer which best stated the main idea from the choices provided. When the tests were collected, the results of the reviewers' responses were compiled for each item. Where there was not a consensus among at least 75% of the reviewers as to the correct choice for the main idea, the answer and sometimes the distractors were revised.

The process of piloting and jurying the test provided essential information needed to produce the final version of
the test. The responses of the pilot subjects and the comments of the reviewers on the multiple choice items revealed answers and distractors in need of adjustment. The reports each group made of difficult vocabulary helped to determine which terms needed glossing. The field of vocabulary terms was narrowed by eliminating those items whose meaning could be derived from context. A maximum of ten items were glossed in the margins of each reading passage; in some cases additional simplification of a passage was needed to avoid overwhelming the subjects with marginal notations. Finally, the time intervals recorded for the pilot subjects established the time limits for each passage and its corresponding multiple choice question. These times were printed at the top of each page to give the subjects an idea of how much time they would have for each task.

Reading Proficiency

With the permission of each subject two measures of reading proficiency were obtained. One was the reading section score from the Michigan English Placement Test. The other was the reading section score from the Test of English as a Foreign Language (TOEFL). There were reasons for and against using each of these scores in the analysis of the data from the study.
The Michigan Test

There were two factors to consider in using the Michigan Test scores as the measure of reading proficiency for the subjects. The most important of these was that this score was available for all forty-nine subjects in the sample. The other was that all but eight of the subjects took the Michigan Test within three weeks after the date of the main idea test, giving a reasonable indication of their reading proficiency at the time the data were collected for this study. (Six of the exceptions had test scores from six weeks prior to the main idea test, while the most recent scores for the remaining two predated this study by just over three months.)

The main problem with the use of these scores was their questionable validity. The construct tested by the Michigan Test appears to be the comprehension of a single decontextualized sentence. There are twenty items in the reading section, each of which consists of one or two sentences followed by a comprehension question; the test-taker then selects the answer from four choices. How closely the construct of this task is related to the construct of main idea comprehension is questionable. The comprehension of the main idea in a passage is certainly predicated on the comprehension of constituent sentences in a text; however, it also requires a high level type of
processing which does not seem to be tested by the Michigan Test reading section items. Thus, the Michigan Test reading scores may not be the best predictors of performance on a test of reading a long passage for its main idea.

Test of English as a Foreign Language  The factors involved in using the TOEFL reading section scores were virtually the reverse of those cited for the Michigan Test. The advantage in using the TOEFL scores as a measure of reading proficiency was that the construct being tested more closely resembled that of the main idea test. Half of the TOEFL reading section items test the comprehension of vocabulary terms in a single sentence context and thus are probably no more closely related to main idea comprehension than items in the Michigan Test reading section. However, the second half of the TOEFL reading section tests comprehension at a level more closely related to that inherent in main idea comprehension. In this part, students are given a series of different types of reading material, including a few passages ranging from approximately 120 to 150 words in length. The multiple choice questions which follow these passages do not explicitly ask the student to choose the main idea of the passage; instead, they are sentence completion items, many of which require the comprehension of the main idea or implication of the passage
to complete the sentence correctly. Thus, the TOEFL reading section contains longer reading samples than the Michigan Test and items which test the comprehension of main ideas or important information.

The disadvantage in relying on the TOEFL scores was that only thirty-eight of the original forty-nine subjects had TOEFL scores on record. Another problem was that the most recent scores for eighteen of these thirty-eight subjects were dated two to five months prior to this study. The remaining twenty subjects had scores from a TOEFL taken four weeks after the testing in this study. So while the use of TOEFL scores was preferable on the basis of construct validity, the interpretation of analyses using these scores would have to include consideration of their potential for inaccuracy.

**Testing Procedure**

Early in the spring semester of 1987, a project proposal was submitted to the Human Subjects Committee of Iowa State University. Permission to conduct the study using human subjects was granted on February 9, 1987. Permission was also granted by Dr. Roberta Vann, Director of the Intensive English and Orientation Program, to contact the instructors and students of the program to seek their
cooperation in the study. In the fall, permission to once again solicit the cooperation of IEOP instructors and students was granted by Dr. Barbara Matthies, Acting Director of the Intensive English and Orientation Program.

The procedure for collecting the data was the same for the subjects tested in the spring and fall IEOP sessions. With the instructors' permission, students in the upper two levels of the program (five and six) were approached and invited to participate in the study. The purpose of the study, the tasks the subjects would be expected to complete, and the use to be made of the data were explained verbally and in writing with a consent form (see Appendix C). Students who agreed to participate in the study signed and returned the forms to the investigator.

Subjects from the spring and fall were tested during the fifth and sixth weeks of the eight-week IEOP term. A two-hour block of class time was allotted to administer the GEFT and the main idea test, as well as distribute and collect the personal data form and background knowledge survey. There were practical advantages in collecting all the data in one meeting. In doing so, the amount of time diverted from regular class activities was minimized. In addition, the complication of possible absenteeism from a second meeting was avoided. The main disadvantage was that
the largely unrelieved two-hour span undoubtedly contributed to test fatigue, especially during the hour required to complete the main idea test.
ANALYSIS AND RESULTS

This chapter details the statistical analysis of the data collected with the tests and other instruments described in the previous chapter. The analysis begins with a series of descriptive statistics used to characterize the subjects in the study. This is followed by an analysis of the reliability of the principal test instruments. The final part of the analysis details the statistical procedures used to address the main research question: does field dependence/independence (FD/I) predict performance in finding the main idea beyond what can be accounted for by reading proficiency and background knowledge? This last section also includes a reanalysis of the main question in light of the findings which arise from the initial procedures.

The data were analyzed using a number of programs from SPSSx (Norusis, 1983, 1985). The program FREQUENCIES was used to obtain means, standard deviations, and other descriptive statistics. ONEWAY was used to obtain analyses of variance. CROSSTABS was used to perform chi-square tests. Reliability estimates were obtained using the program RELIABILITY (models SPLIT and ALPHA). Correlations were performed using PEARSON CORR. Finally, multiple regression analyses were performed using the program REGRESSION.
Characteristics of the Sample

The first step in the analysis of the data was to characterize the subjects in terms of their collective performance on the measures. Means, standard deviations, ranges, and kurtosis and skewness statistics were obtained for the entire sample (n=49) and for various subgroups within the sample on seven measures: overall TOEFL scores (TOFLTOTL); TOEFL reading section scores (TOFLRDG); overall Michigan Test scores (EPTTOTL); Michigan Test reading section scores (EPTRDG); overall scores on the Group Embedded Figures Test (GEFT); overall scores on the background knowledge survey (BKGRDTOT); and overall scores on the main idea test (MIT).

Descriptive Statistics for the Entire Sample

Table 3 displays the descriptive statistics on the above measures for the entire sample.\(^2\) In addition to obtaining the means, standard deviations, and ranges for each measure, kurtosis and skewness statistics were computed to determine whether the distribution of scores for each measure was normal or near normal. Kurtosis and skewness

\(^2\)It should be noted that for this and all subsequent analysis the sample size was variously reduced when the overall and reading section scores from the TOEFL were used. This was due to the fact that not all subjects had both of these scores on record.
are considered normal at 0 and close to normal if they fall within the range of -1.00 and +1.00.

TABLE 3. Frequency statistics on seven measures for the entire sample

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL a</td>
<td>43</td>
<td>480.47</td>
<td>46.42</td>
<td>360-583</td>
<td>.56</td>
<td>.24</td>
</tr>
<tr>
<td>TOFLRDG a</td>
<td>38</td>
<td>48.08</td>
<td>6.98</td>
<td>30-61</td>
<td>.36</td>
<td>-.12</td>
</tr>
<tr>
<td>EPTTOTL a</td>
<td>49</td>
<td>75.94</td>
<td>10.58</td>
<td>55-96</td>
<td>-.88</td>
<td>-.11</td>
</tr>
<tr>
<td>EPTRDG a</td>
<td>49</td>
<td>68.33</td>
<td>20.10</td>
<td>0-95</td>
<td>3.80</td>
<td>-1.71</td>
</tr>
<tr>
<td>GEFT a</td>
<td>49</td>
<td>11.41</td>
<td>5.03</td>
<td>1-18</td>
<td>-.99</td>
<td>-.39</td>
</tr>
<tr>
<td>BKGRDTOT a,b</td>
<td>49</td>
<td>22.88</td>
<td>5.59</td>
<td>11-37</td>
<td>.23</td>
<td>.40</td>
</tr>
<tr>
<td>MIT a,c</td>
<td>49</td>
<td>4.53</td>
<td>1.91</td>
<td>1-9</td>
<td>-.67</td>
<td>.09</td>
</tr>
</tbody>
</table>

aTOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

bIndividual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

cIndividual scores were calculated as the number correct out of a total of 10 items.
The overall TOEFL mean was 480.47 with a standard deviation of 46.42 and a range of 360 to 583. The TOEFL reading mean was 48.08 with a standard deviation of 6.98 and a range of 30 to 61. To see if the subjects of this study behaved similarly to other TOEFL test-takers, the means and standard deviations on the overall and reading scores from the sample were compared to the same statistics obtained by the Educational Testing Service (ETS). The sample data were broken down by sex to facilitate a direct comparison with the ETS data, which did not present statistics for a mixed group (see Table 4 and Table 5).

**TABLE 4.** Comparison of means and standard deviations on overall TOEFL and reading section scores for females

<table>
<thead>
<tr>
<th>Score</th>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL b</td>
<td>ETS a</td>
<td>226,635</td>
<td>515.0</td>
<td>65.0</td>
</tr>
<tr>
<td></td>
<td>Study sample</td>
<td>13</td>
<td>467.0</td>
<td>30.6</td>
</tr>
<tr>
<td>TOFLRDG b</td>
<td>ETS</td>
<td>226,635</td>
<td>50.9</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>Study sample</td>
<td>11</td>
<td>45.6</td>
<td>5.3</td>
</tr>
</tbody>
</table>

a Based on examinees tested from July 1984 through June 1986 who responded to sex group membership on answer sheets. The ETS data were taken from the TOEFL Test and Score Manual (1987).

b TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score.
TABLE 5. Comparison of means and standard deviations on overall TOEFL and reading section scores for males

<table>
<thead>
<tr>
<th>Score</th>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL b</td>
<td>ETS a</td>
<td>449,654</td>
<td>511.0</td>
<td>66.0</td>
</tr>
<tr>
<td>Study sample</td>
<td></td>
<td>30</td>
<td>486.0</td>
<td>51.1</td>
</tr>
<tr>
<td>TOFLRDG b</td>
<td>ETS</td>
<td>449,654</td>
<td>51.3</td>
<td>7.4</td>
</tr>
<tr>
<td>Study sample</td>
<td></td>
<td>27</td>
<td>49.1</td>
<td>7.4</td>
</tr>
</tbody>
</table>

a Based on examinees tested from July 1984 through June 1986 who responded to sex group membership on answer sheets. The ETS data were taken from the TOEFL Test and Score Manual (1987).

b TOFLTOTL = overall TOEFL score; TOFLRDG = TOEFL reading section score.

This comparison revealed that the TOEFL performance of females and males in this study was somewhat lower than that of other TOEFL test-takers. This was not unexpected considering the fact that the subjects of this study were at a pre-admission level of proficiency, and the ETS data includes students who are at or above the level required for university admission.

The overall Michigan Test mean was 75.94 (see Table 3) with a standard deviation of 10.58 and a range of 55 to 96. The Michigan Test reading mean was 68.33 with a standard deviation of 20.10 and a wide range of 0 to 95. The
kurtosis statistic was high (3.80), indicating a peaked distribution; the skewness statistic (-1.71) was low, indicating that the scores were skewed toward the high end of the range. Thus, the distribution for the Michigan Test reading scores was distorted.

To see how similar this sample's performance was to that of other students who took the Michigan Test, the means and standard deviations on the overall and reading scores from this study were compared to those of the entire IEOP sample from which they were drawn. Because the study sample comprised students from two IEOP sessions, their means and standard deviations were calculated separately.

The statistics on the overall scores for both spring and fall members of the study sample were fairly comparable, though the fall group's mean was higher than the spring group's mean by 3.5 points (see Table 6). The statistics on the overall scores for both the spring and fall IEOP samples were also comparable; the means were separated by a much narrower margin (1.15 points) than the spring and fall study samples. When the spring and fall samples were compared, the subjects of this study had higher overall score means and lower overall score standard deviations than the respective IEOP samples of which they were a part.
TABLE 6. Comparison of means and standard deviations on the Michigan Test overall scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>Study</td>
<td>18</td>
<td>73.72</td>
<td>10.90</td>
</tr>
<tr>
<td></td>
<td>IEOPa</td>
<td>54</td>
<td>64.39</td>
<td>17.14</td>
</tr>
<tr>
<td>Fall</td>
<td>Study</td>
<td>23</td>
<td>77.22</td>
<td>10.64</td>
</tr>
<tr>
<td></td>
<td>IEOP</td>
<td>70</td>
<td>63.24</td>
<td>16.84</td>
</tr>
</tbody>
</table>

aIntensive English and Orientation Program.

The means and standard deviations on the reading scores were similar for the spring and fall members of the study sample; the fall group had a slightly higher mean (1.41 points higher) on the reading section than the spring group (see Table 7). The means and standard deviations on the reading scores were very similar for the spring and fall IEOP samples; the means were separated by only a fraction of a point. When the spring and fall samples were compared, the subjects of this study had higher reading means and lower standard deviations than students from the respective IEOP samples.

The results of the preceding comparisons between the study samples and the IEOP samples from which they were drawn showed that the subjects of this study occupied the
TABLE 7. Comparison of means and standard deviations on the Michigan Test reading scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>Study</td>
<td>18</td>
<td>13.50</td>
<td>3.56</td>
</tr>
<tr>
<td></td>
<td>IEOP&lt;sup&gt;a&lt;/sup&gt;</td>
<td>54</td>
<td>11.52</td>
<td>4.10</td>
</tr>
<tr>
<td>Fall</td>
<td>Study</td>
<td>23</td>
<td>14.91</td>
<td>2.76</td>
</tr>
<tr>
<td></td>
<td>IEOP</td>
<td>70</td>
<td>11.21</td>
<td>4.26</td>
</tr>
</tbody>
</table>

<sup>a</sup>Intensive English and Orientation Program.

The upper range of the distributions for both overall and reading scores on the Michigan Test. The fact that the study sample had higher means on the overall and reading scores than the IEOP sample was not unexpected; it only reflected the sample's advanced level within the program.

Following the measures of English proficiency in Table 3 is the measure of field dependence/independence (FD/I), the Group Embedded Figures Test (GEFT). The mean score on the GEFT was 11.41 with a standard deviation of 5.03 and a wide range of 1 to 18 (eighteen is the highest score possible on the GEFT). To see if the subjects of this sample performed similarly to other students who took the GEFT, the mean and standard deviation were compared to those found by the developers and researchers of the GEFT (Witkin,
Oltman, Raskin, & Karp, 1971). Here again the sample data were broken down by sex to permit a direct comparison of the two sets of results (see Table 8).

<table>
<thead>
<tr>
<th>Source of data</th>
<th>Females</th>
<th></th>
<th>Males</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>S.D.</td>
<td>N</td>
</tr>
<tr>
<td>This study</td>
<td>16</td>
<td>12.4</td>
<td>5.3</td>
<td>33</td>
</tr>
<tr>
<td>GEFT(^a) Manual</td>
<td>242</td>
<td>10.8</td>
<td>4.2</td>
<td>155</td>
</tr>
</tbody>
</table>

\(^a\)Group Embedded Figures Test.

The statistics for females and males from Witkin et al. (1971) were almost a complete reverse of the findings for the study sample. Witkin et al. found males to be more field independent than females whereas the reverse was true for the study sample. Witkin et al.'s males had a mean of 12.0 with a standard deviation of 4.1, while the females in the present study had a like mean of 12.4 with a standard deviation of 5.3. Witkin et al.'s females had a mean of 10.8 with a standard deviation of 4.2, while the males in this study had a similar mean of 10.9 with a standard deviation of 4.9.
The next measure for which statistics are reported in Table 3 is the background knowledge survey (BKGRDTOT). The group as a whole did not seem to perceive themselves as very knowledgeable of the topics represented on the background knowledge survey. The possible range of scores on this measure was 10 to 50, and the subjects in this study rated their overall background knowledge at the lower end of this range. Their mean score on the survey was 22.88, with a standard deviation of 5.59 and a range of 11 to 37.

The final measure for which statistics are given in Table 3 is the main idea test (MIT). The mean score on the MIT was 4.53 with a standard deviation of 1.91. The range of 1 to 9 reflected a diversity of performance on the test (ten was the highest score possible for the MIT).

Descriptive Statistics for Sample Subgroups

The statistics discussed thus far were descriptive of the study sample as a whole. On the basis of the three student characteristics noted in the previous chapter, it was possible to also examine the sample by subgroups. These subgroups were classified by graduate/undergraduate status, membership in one of four language groups, and membership in one of four major academic areas. In addition to obtaining descriptive statistics for each of these subgroups, between-groups variance was examined using analyses of variance (ANOVAs).
Graduate/Undergraduate Subgroups

The first of the subgroups to be compared was the graduate and undergraduate subjects in the sample. Table 9 and Table 10 display the means, standard deviations, ranges, and kurtosis and skewness statistics on the measures for these two groups.

TABLE 9. Frequency statistics on seven measures for graduates

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>27</td>
<td>493.11</td>
<td>47.54</td>
<td>413-583</td>
<td>-.56</td>
<td>.34</td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>22</td>
<td>51.27</td>
<td>6.31</td>
<td>37-61</td>
<td>-.41</td>
<td>-.12</td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>28</td>
<td>76.64</td>
<td>10.54</td>
<td>58-96</td>
<td>-.85</td>
<td>-.01</td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>28</td>
<td>65.11</td>
<td>22.14</td>
<td>0-95</td>
<td>3.95</td>
<td>-1.79</td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>28</td>
<td>9.96</td>
<td>5.08</td>
<td>1-18</td>
<td>-1.02</td>
<td>-.08</td>
</tr>
<tr>
<td>BKGRDTOT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>28</td>
<td>23.46</td>
<td>5.70</td>
<td>14-37</td>
<td>.06</td>
<td>.34</td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,c&lt;/sup&gt;</td>
<td>28</td>
<td>5.39</td>
<td>1.50</td>
<td>3-9</td>
<td>-.06</td>
<td>.33</td>
</tr>
</tbody>
</table>

<sup>a</sup> TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>b</sup> Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>c</sup> Individual scores were calculated as the number correct out of a total of 10 items.
TABLE 10. Frequency statistics on seven measures for undergraduates

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>16</td>
<td>459.13</td>
<td>36.66</td>
<td>360-517</td>
<td>2.73</td>
<td>-1.19</td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>16</td>
<td>43.69</td>
<td>5.36</td>
<td>30-50</td>
<td>1.69</td>
<td>-1.29</td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>21</td>
<td>75.00</td>
<td>10.82</td>
<td>55-91</td>
<td>-.95</td>
<td>-.23</td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>21</td>
<td>72.62</td>
<td>16.55</td>
<td>35-90</td>
<td>.41</td>
<td>-1.17</td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>21</td>
<td>13.33</td>
<td>4.37</td>
<td>4-18</td>
<td>-.33</td>
<td>-.87</td>
</tr>
<tr>
<td>BKGRDTOT&lt;sup&gt;a&lt;/sup&gt;,b</td>
<td>21</td>
<td>22.10</td>
<td>5.47</td>
<td>11-34</td>
<td>1.05</td>
<td>.51</td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,c&lt;/sup&gt;</td>
<td>21</td>
<td>3.38</td>
<td>1.80</td>
<td>1-7</td>
<td>-.33</td>
<td>.78</td>
</tr>
</tbody>
</table>

<sup>a</sup>TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>b</sup>Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>c</sup>Individual scores were calculated as the number correct out of a total of 10 items.
To determine whether the differences between graduates and undergraduates were statistically significant, an ANOVA was run for each of the seven measures (see Table 11). The variance between the groups on the overall and reading section scores of the TOEFL was significant (at \( p = 0.018 \) and \( p = 0.000 \), respectively). Thus, the graduates had a higher level of English proficiency than the undergraduates on the basis of TOEFL scores. The variance in mean scores on the MIT was also statistically significant \( (p = 0.000) \), further substantiating the higher level of performance found for the graduate students. Finally, the variance between graduate and undergraduate means on the GEFT was significant \( (p = 0.019) \); the graduate students were significantly more field dependent than the undergraduate students.

**First Language Subgroups** The second set of subgroup comparisons was carried out using first language as the classifying principle. As mentioned before, three distinct language groups were evident in the sample: Arabic \( (n=10) \), Japanese \( (n=15) \), and Spanish \( (n=10) \) speakers. The remainder of the sample \( (n=14) \) consisted of mostly Asian language speakers and was placed in a miscellaneous category. The means, standard deviations, ranges, and kurtosis and skewness statistics on the measures for each language group may be found in Table 12 through Table 15.
TABLE 11. ANOVA between graduate and undergraduate performance on seven measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Source of var.</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;b&lt;/sup&gt;</td>
<td>43</td>
<td>BG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1</td>
<td>11604.28</td>
<td>6.03</td>
<td>.018*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG&lt;sup&gt;c&lt;/sup&gt;</td>
<td>41</td>
<td>1924.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOFLRDLG&lt;sup&gt;b&lt;/sup&gt;</td>
<td>38</td>
<td>BG</td>
<td>1</td>
<td>532.96</td>
<td>15.13</td>
<td>.000***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>36</td>
<td>35.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPTTOL&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49</td>
<td>BG</td>
<td>1</td>
<td>32.39</td>
<td>.29</td>
<td>.596</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>47</td>
<td>113.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49</td>
<td>BG</td>
<td>1</td>
<td>677.15</td>
<td>1.70</td>
<td>.199</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>47</td>
<td>398.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49</td>
<td>BG</td>
<td>1</td>
<td>136.21</td>
<td>5.93</td>
<td>.019*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>47</td>
<td>22.97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Variance between groups.

<sup>b</sup>TOFLTOTL=overall TOEFL score; TOFLRDLG=TOEFL reading section score; EPTTOL=overall Michigan test score; EPTRDG=Mic...ction score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>c</sup>Variance within groups.

*<sup>p</sup><.05; ***<sup>p</sup><.001.
Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

Individual scores were calculated as the number correct out of a total of 10 items.

A second ANOVA was run to make meaningful comparisons between the four language groups. Table 16 displays the results of the ANOVA among the four language groups' means on the seven measures.

Statistically significant differences between the groups were found for three of the measures: the overall TOEFL means, the TOEFL reading means, and the GEFT means. The variance between groups on the overall TOEFL means was significant at the $p=.013$ level. A Scheffé test was run (see Table 17) which showed that the source of variance was
TABLE 12. Frequency statistics on seven measures for Arabic speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9</td>
<td>485.33</td>
<td>43.67</td>
<td>430-580</td>
<td>2.21</td>
<td>1.28</td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7</td>
<td>48.00</td>
<td>6.38</td>
<td>41-61</td>
<td>3.39</td>
<td>1.62</td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>79.50</td>
<td>11.00</td>
<td>62-95</td>
<td>-.42</td>
<td>-.62</td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>63.30</td>
<td>26.90</td>
<td>0-90</td>
<td>2.97</td>
<td>-1.72</td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>11.20</td>
<td>5.61</td>
<td>2-18</td>
<td>-1.07</td>
<td>-.39</td>
</tr>
<tr>
<td>BKGRDTOT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>10</td>
<td>23.00</td>
<td>7.26</td>
<td>11-34</td>
<td>.11</td>
<td>.07</td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,c&lt;/sup&gt;</td>
<td>10</td>
<td>4.10</td>
<td>2.18</td>
<td>1-7</td>
<td>-1.46</td>
<td>-.24</td>
</tr>
</tbody>
</table>

<sup>a</sup>ToFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>b</sup>Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>c</sup>Individual scores were calculated as the number correct out of a total of 10 items.
TABLE 13. Frequency statistics on seven measures for Japanese speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL(^a)</td>
<td>12</td>
<td>458.33</td>
<td>50.18</td>
<td>360-547</td>
<td>.19</td>
<td>-.26</td>
</tr>
<tr>
<td>TOFLRDG(^a)</td>
<td>11</td>
<td>44.46</td>
<td>6.02</td>
<td>30-53</td>
<td>2.84</td>
<td>-1.31</td>
</tr>
<tr>
<td>EPTTOTL(^a)</td>
<td>15</td>
<td>75.67</td>
<td>11.05</td>
<td>55-90</td>
<td>-1.05</td>
<td>-.36</td>
</tr>
<tr>
<td>EPTRDG(^a)</td>
<td>15</td>
<td>70.00</td>
<td>14.64</td>
<td>40-90</td>
<td>-.34</td>
<td>-.69</td>
</tr>
<tr>
<td>GEFT(^a)</td>
<td>15</td>
<td>13.73</td>
<td>4.43</td>
<td>4-18</td>
<td>.08</td>
<td>-1.04</td>
</tr>
<tr>
<td>BKGRDTOT(^a,b)</td>
<td>15</td>
<td>22.07</td>
<td>5.26</td>
<td>14-34</td>
<td>.60</td>
<td>.36</td>
</tr>
<tr>
<td>MIT(^a,c)</td>
<td>15</td>
<td>3.87</td>
<td>1.89</td>
<td>2-8</td>
<td>.25</td>
<td>1.03</td>
</tr>
</tbody>
</table>

\(^a\)TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Mic...ual background knowledge score; MIT=overall main idea test score.

\(^b\)Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

\(^c\)Individual scores were calculated as the number correct out of a total of 10 items.
TABLE 14. Frequency statistics on seven measures for Spanish speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>517.40</td>
<td>43.90</td>
<td>447-583</td>
<td>-.80</td>
<td>.15</td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>55.20</td>
<td>4.52</td>
<td>48-61</td>
<td>-1.21</td>
<td>-.46</td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>72.70</td>
<td>10.66</td>
<td>58-91</td>
<td>-.70</td>
<td>.12</td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>60.00</td>
<td>24.61</td>
<td>0-85</td>
<td>3.80</td>
<td>-1.76</td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>8.00</td>
<td>3.80</td>
<td>2-14</td>
<td>-1.00</td>
<td>-.05</td>
</tr>
<tr>
<td>BKGRDTOT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>10</td>
<td>25.40</td>
<td>6.35</td>
<td>16-37</td>
<td>-.39</td>
<td>.33</td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,c&lt;/sup&gt;</td>
<td>10</td>
<td>5.70</td>
<td>1.49</td>
<td>4-9</td>
<td>1.74</td>
<td>1.14</td>
</tr>
</tbody>
</table>

<sup>a</sup>TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>b</sup>Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>c</sup>Individual scores were calculated as the number correct out of a total of 10 items.
TABLE 15. Frequency statistics on seven measures for speakers of miscellaneous languages

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLOTTOTLa</td>
<td>12</td>
<td>468.17</td>
<td>28.07</td>
<td>427-520</td>
<td>.17</td>
<td>.77</td>
</tr>
<tr>
<td>TOFLOTRDGa</td>
<td>10</td>
<td>45.00</td>
<td>5.42</td>
<td>35-53</td>
<td>.37</td>
<td>-.79</td>
</tr>
<tr>
<td>EPTTOTLa</td>
<td>14</td>
<td>76.00</td>
<td>10.08</td>
<td>59-96</td>
<td>-.01</td>
<td>.37</td>
</tr>
<tr>
<td>EPTRDGa</td>
<td>14</td>
<td>76.07</td>
<td>14.17</td>
<td>45-95</td>
<td>.12</td>
<td>-.80</td>
</tr>
<tr>
<td>GEFTA</td>
<td>14</td>
<td>11.5</td>
<td>5.03</td>
<td>1-18</td>
<td>-.32</td>
<td>-.70</td>
</tr>
<tr>
<td>BKGRD TOTA,b</td>
<td>14</td>
<td>21.86</td>
<td>3.78</td>
<td>17-29</td>
<td>-.84</td>
<td>.39</td>
</tr>
<tr>
<td>MITa,c</td>
<td>14</td>
<td>4.71</td>
<td>1.73</td>
<td>2-7</td>
<td>-.95</td>
<td>-.22</td>
</tr>
</tbody>
</table>

aTOFLOTTOTL=overall TOEFL score; TOFLOTRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRD TOTL=overall background knowledge score; MIT=overall main idea test score.

bIndividual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

Individual scores were calculated as the number correct out of a total of 10 items.
### TABLE 16. ANOVA among the performance of the language groups on seven measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Source of var.</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;b&lt;/sup&gt;</td>
<td>43</td>
<td>BG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3</td>
<td>7182.66</td>
<td>4.06</td>
<td>.013*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG&lt;sup&gt;c&lt;/sup&gt;</td>
<td>39</td>
<td>1768.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;b&lt;/sup&gt;</td>
<td>38</td>
<td>BG</td>
<td>3</td>
<td>248.81</td>
<td>8.02</td>
<td>.000***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>34</td>
<td>31.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49</td>
<td>BG</td>
<td>3</td>
<td>77.63</td>
<td>.68</td>
<td>.569</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>45</td>
<td>114.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49</td>
<td>BG</td>
<td>3</td>
<td>609.25</td>
<td>1.56</td>
<td>.212</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>45</td>
<td>390.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49</td>
<td>BG</td>
<td>3</td>
<td>65.94</td>
<td>2.92</td>
<td>.044*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>45</td>
<td>22.62</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Variance between groups.

<sup>b</sup>TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>c</sup>Variance within groups.

*<sup>p</sup><.05; ***<sup>p</sup><.001.
TABLE 16. Continued

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Source of var.</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>BKGRDTOT&lt;sup&gt;b,d&lt;/sup&gt;</td>
<td>49</td>
<td>BG</td>
<td>3</td>
<td>29.41</td>
<td>.94</td>
<td>.430</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>45</td>
<td>31.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT&lt;sup&gt;b,e&lt;/sup&gt;</td>
<td>49</td>
<td>BG</td>
<td>3</td>
<td>7.54</td>
<td>2.24</td>
<td>.100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>45</td>
<td>3.37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>d</sup> Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>e</sup> Individual scores were calculated as the number correct out of a total of 10 items.

between the Japanese and Spanish speakers (p=.05); no other two groups were different at a statistically significant level for this measure.

The variance between groups on the TOEFL reading means was greater than for the overall means and occurred at a higher level of significance (p=.000). The Scheffé test run for this measure (see Table 18) showed that the significant variance again occurred between the Japanese and Spanish speakers (p=.05); significant variance was also found between the Spanish speakers and the miscellaneous language group (p=.05).
TABLE 17. Scheffé test for source of variance among language groups on overall TOEFL scores

<table>
<thead>
<tr>
<th>Mean</th>
<th>Group</th>
<th>Japanese</th>
<th>Misc.</th>
<th>Arabic</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>458.33</td>
<td>Japanese</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>468.17</td>
<td>Misc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>485.33</td>
<td>Arabic</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>517.40</td>
<td>Spanish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Denotes a pair of groups which are significantly different at the p=.05 level.

TABLE 18. Scheffé test for source of variance among language groups on TOEFL reading scores

<table>
<thead>
<tr>
<th>Mean</th>
<th>Group</th>
<th>Japanese</th>
<th>Misc.</th>
<th>Arabic</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>44.46</td>
<td>Japanese</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45.00</td>
<td>Misc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48.00</td>
<td>Arabic</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>55.20</td>
<td>Spanish</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

* Denotes a pair of groups which are significantly different at the p=.05 level.

Finally, the variance between groups on the GEFT means was also statistically significant (p=.044). A Scheffé test (see Table 19) showed the locus of significant variance once again to be between the Japanese and Spanish speakers (p=.05).
TABLE 19. Scheffé test for source of variance among language groups on Group Embedded Figures Test scores

<table>
<thead>
<tr>
<th>Mean</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.00</td>
<td>Spanish</td>
</tr>
<tr>
<td>11.20</td>
<td>Arabic</td>
</tr>
<tr>
<td>11.50</td>
<td>Misc.</td>
</tr>
<tr>
<td>13.73</td>
<td>Japanese *</td>
</tr>
</tbody>
</table>

* Denotes a pair of groups which are significantly different at the p=.05 level.

The qualitative significance of these findings may be better understood with a visual ranking of the language groups' means on the measures. Table 20 provides a linear depiction of the language groups on each measure according to mean scores.

This table shows that the Spanish speakers had the highest TOEFL means and were also the most field dependent language group in the sample; conversely, the Japanese speakers had the lowest TOEFL means and were the most field independent language group. The ranking of means on the background knowledge survey and the MIT was also notable. While differences between groups were not statistically significant for these measures, it was interesting to again observe the opposition of these means for Japanese and
TABLE 20. Linear ranking of mean scores of the language groups on seven measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Lowest mean</th>
<th>Highest mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Japanese 458.33</td>
<td>Arabic 485.33</td>
</tr>
<tr>
<td></td>
<td>Misc. 468.17</td>
<td></td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Japanese 44.46</td>
<td>Arabic 48.00</td>
</tr>
<tr>
<td></td>
<td>Misc. 45.00</td>
<td></td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Spanish 72.70</td>
<td>Arabic 79.50</td>
</tr>
<tr>
<td></td>
<td>Japanese 75.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Misc. 76.00</td>
<td></td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Spanish 60.00</td>
<td>Arabic 70.00</td>
</tr>
<tr>
<td></td>
<td>Japanese 63.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arabic 70.00</td>
<td></td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Spanish 8.00</td>
<td>Arabic 11.20</td>
</tr>
<tr>
<td></td>
<td>Japanese 11.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arabic 11.50</td>
<td></td>
</tr>
<tr>
<td>BKGRDTOT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>Misc. 21.86</td>
<td>Arabic 23.00</td>
</tr>
<tr>
<td></td>
<td>Japanese 22.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arabic 23.00</td>
<td></td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,c&lt;/sup&gt;</td>
<td>Japanese 3.87</td>
<td>Arabic 4.71</td>
</tr>
<tr>
<td></td>
<td>Misc. 4.10</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>b</sup>Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>c</sup>Individual scores were calculated as the number correct out of a total of 10 items.
Spanish speakers. The Japanese speakers had the lowest means for these two measures while the Spanish speakers had the highest, a trend which echoed the pattern found for the TOEFL means.

**Major Academic Area Subgroups** The third and final set of subsample comparisons was performed using major academic area to establish four subgroups. As noted previously, the spectrum of academic majors reported by individual subjects were collapsed into four categories: business/economics, engineering, science/mathematics, and liberal arts. The means, standard deviations, ranges, and kurtosis and skewness statistics on the measures for each major academic area are found in Table 21 through Table 24.

A third ANOVA was run (Table 25) to discover whether statistically significant differences in performance on the measures occurred between these subgroups.

The only measure for which significant group differences emerged was the GEFT ($p=.027$). A subsequent Scheffé test (Table 26) revealed that the only significant source of variance on the GEFT means occurred between the business/economics majors and the science/mathematics majors ($p=.05$). It was curious to find that the business/economics majors were more field independent with a GEFT mean of 15.44 than the science/mathematics majors, who were more field
TABLE 21. Frequency statistics on seven measures for business/economics majors

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7</td>
<td>456.71</td>
<td>46.38</td>
<td>413-547</td>
<td>2.01</td>
<td>1.33</td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7</td>
<td>46.29</td>
<td>4.03</td>
<td>40-53</td>
<td>1.10</td>
<td>.22</td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9</td>
<td>71.67</td>
<td>11.54</td>
<td>55-91</td>
<td>-.26</td>
<td>.36</td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9</td>
<td>70.00</td>
<td>16.01</td>
<td>45-90</td>
<td>-.94</td>
<td>-.29</td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9</td>
<td>15.44</td>
<td>2.51</td>
<td>11-18</td>
<td>-.30</td>
<td>-.90</td>
</tr>
<tr>
<td>BKGRDTOT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>9</td>
<td>21.67</td>
<td>6.10</td>
<td>15-34</td>
<td>.81</td>
<td>.88</td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,c&lt;/sup&gt;</td>
<td>9</td>
<td>4.00</td>
<td>2.05</td>
<td>1-8</td>
<td>1.41</td>
<td>.94</td>
</tr>
</tbody>
</table>

<sup>a</sup>TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Micigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>b</sup>Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>c</sup>Individual scores were calculated as the number correct out of a total of 10 items.
TABLE 22. Frequency statistics on seven measures for engineering majors

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9</td>
<td>475.56</td>
<td>50.20</td>
<td>360-520</td>
<td>3.39</td>
<td>-1.67</td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8</td>
<td>44.75</td>
<td>7.09</td>
<td>30-53</td>
<td>2.33</td>
<td>-1.33</td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9</td>
<td>79.33</td>
<td>11.61</td>
<td>62-96</td>
<td>-.54</td>
<td>-.45</td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9</td>
<td>75.00</td>
<td>16.39</td>
<td>40-95</td>
<td>1.86</td>
<td>-1.15</td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9</td>
<td>12.22</td>
<td>4.68</td>
<td>5-18</td>
<td>-1.19</td>
<td>-.28</td>
</tr>
<tr>
<td>BKGRDTOT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>9</td>
<td>22.33</td>
<td>6.04</td>
<td>11-34</td>
<td>2.50</td>
<td>.10</td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,c&lt;/sup&gt;</td>
<td>9</td>
<td>4.22</td>
<td>2.17</td>
<td>1-7</td>
<td>-1.34</td>
<td>.01</td>
</tr>
</tbody>
</table>

<sup>a</sup>TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>b</sup>Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>c</sup>Individual scores were calculated as the number correct out of a total of 10 items.
TABLE 23. Frequency statistics on seven measures for science/mathematics majors

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL(^a)</td>
<td>18</td>
<td>495.22</td>
<td>47.81</td>
<td>427-583</td>
<td>-.38</td>
<td>.73</td>
</tr>
<tr>
<td>TOFLRDG(^a)</td>
<td>17</td>
<td>50.53</td>
<td>7.81</td>
<td>35-61</td>
<td>-.43</td>
<td>-.40</td>
</tr>
<tr>
<td>EPTTOTL(^a)</td>
<td>20</td>
<td>74.80</td>
<td>10.56</td>
<td>58-95</td>
<td>-.77</td>
<td>.15</td>
</tr>
<tr>
<td>EPTRDG(^a)</td>
<td>20</td>
<td>59.65</td>
<td>24.62</td>
<td>0-85</td>
<td>1.84</td>
<td>-1.45</td>
</tr>
<tr>
<td>GEFT(^a)</td>
<td>20</td>
<td>9.65</td>
<td>4.58</td>
<td>1-17</td>
<td>-.77</td>
<td>-.26</td>
</tr>
<tr>
<td>BKGRDTOT(^a,(^b)</td>
<td>20</td>
<td>23.45</td>
<td>5.55</td>
<td>14-37</td>
<td>.49</td>
<td>.55</td>
</tr>
<tr>
<td>MIT(^a,(^c)</td>
<td>20</td>
<td>5.10</td>
<td>1.65</td>
<td>2-9</td>
<td>.49</td>
<td>.21</td>
</tr>
</tbody>
</table>

\(^{a}\)TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Micigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

\(^{b}\)Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

\(^{c}\)Individual scores were calculated as the number correct out of a total of 10 items.
### TABLE 24. Frequency statistics on seven measures for liberal arts majors

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8</td>
<td>473.63</td>
<td>38.05</td>
<td>430-547</td>
<td>.98</td>
<td>.81</td>
</tr>
<tr>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6</td>
<td>47.67</td>
<td>5.68</td>
<td>41-58</td>
<td>2.78</td>
<td>1.28</td>
</tr>
<tr>
<td>EPTTOTL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8</td>
<td>76.25</td>
<td>8.35</td>
<td>62-87</td>
<td>-.39</td>
<td>-.70</td>
</tr>
<tr>
<td>EPTRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8</td>
<td>78.13</td>
<td>10.67</td>
<td>60-90</td>
<td>-.55</td>
<td>-.88</td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8</td>
<td>10.13</td>
<td>6.77</td>
<td>2-18</td>
<td>-2.30</td>
<td>.05</td>
</tr>
<tr>
<td>BKGRDTOT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>8</td>
<td>23.13</td>
<td>6.47</td>
<td>14-34</td>
<td>-.36</td>
<td>.38</td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,c&lt;/sup&gt;</td>
<td>8</td>
<td>5.13</td>
<td>1.73</td>
<td>2-7</td>
<td>-.10</td>
<td>-.70</td>
</tr>
</tbody>
</table>

<sup>a</sup> TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>b</sup> Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>c</sup> Individual scores were calculated as the number correct out of a total of 10 items.
TABLE 25. ANOVA among the performance of academic major groups on seven measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Source of var.</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLTOTL</td>
<td>42</td>
<td>BG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3</td>
<td>2819.95</td>
<td>1.31</td>
<td>.287</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG&lt;sup&gt;c&lt;/sup&gt;</td>
<td>38</td>
<td>2159.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOFLRDG</td>
<td>38</td>
<td>BG</td>
<td>3</td>
<td>71.42</td>
<td>1.53</td>
<td>.224</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>34</td>
<td>46.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPTTOTL</td>
<td>46</td>
<td>BG</td>
<td>3</td>
<td>92.47</td>
<td>.82</td>
<td>.492</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>42</td>
<td>113.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPTRDG</td>
<td>46</td>
<td>BG</td>
<td>3</td>
<td>897.01</td>
<td>2.28</td>
<td>.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>42</td>
<td>393.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEFT</td>
<td>46</td>
<td>BG</td>
<td>3</td>
<td>75.84</td>
<td>3.37</td>
<td>.027*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>42</td>
<td>22.51</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Variance between groups.

<sup>b</sup>TOFLTOTL=overall TOEFL score; TOFLRDG=TOEFL reading section score; EPTTOTL=overall Michigan test score; EPTRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; BKGRDTOT=overall background knowledge score; MIT=overall main idea test score.

<sup>c</sup>Variance within groups.

*<sup>p</sup>&lt;.05.
TABLE 25. Continued

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Source of var.</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>BKGRDTOT&lt;sup&gt;b&lt;/sup&gt;&lt;sup&gt;,d&lt;/sup&gt;</td>
<td>46</td>
<td>BG</td>
<td>3</td>
<td>7.60</td>
<td>.22</td>
<td>.884</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>42</td>
<td>34.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT&lt;sup&gt;b&lt;/sup&gt;&lt;sup&gt;,e&lt;/sup&gt;</td>
<td>46</td>
<td>BG</td>
<td>3</td>
<td>4.77</td>
<td>1.40</td>
<td>.258</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>42</td>
<td>3.42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>d</sup> Individual scores for this measure comprised the summation of the numerical ratings (on a scale of 1 to 5) each subject gave of their background knowledge of the ten topics included in the main idea test. Overall scores on this instrument fell within a range of 10 to 50 points.

<sup>e</sup> Individual scores were calculated as the number correct out of a total of 10 items.

Summary of Sample Characteristics

The statistical analysis thus far served to characterize the study sample and three subgroups within the sample. In the process some subgroup characteristics were identified which have some bearing on the analysis of the main research question in the study. In the comparison of graduates to dependent with a GEFT mean of 9.65. This runs counter to previous research which has shown subjects in science and mathematics majors to be very field independent (Witkin, Moore, Goodenough, & Cox, 1977).
TABLE 26. Scheffé test for source of variance among academic major groups on Group Embedded Figures Test scores

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>9.65</td>
<td>Sci.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.13</td>
<td>Lib.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.22</td>
<td>Eng.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.44</td>
<td>Bus.</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

aSci.=science/mathematics; Lib.=liberal arts; Eng.=engineering; Bus.=business/economics.

* Denotes a pair of groups which are significantly different at the p=.05 level.

undergraduates, the graduate students had higher TOEFL means and a higher MIT mean than the undergraduate students, thus demonstrating a higher level of English proficiency and performance in finding the main idea. At the same time, the graduates had a lower GEFT mean, indicating they were more field dependent than the undergraduates. When subgroups were compared on the basis of first language, significant differences were found mainly between the Japanese and Spanish speakers. The measures for which these differences were significant were again the TOEFL and the GEFT; though not statistically significant, differences between the same two language groups were evident in the MIT means as well as the background knowledge survey means. It appeared that the
Spanish speakers were the most proficient readers in the sample on the basis of their TOEFL and MIT scores and were the most field dependent language group in the sample. Conversely, the Japanese speakers appeared to be the least proficient readers on the basis of TOEFL and MIT scores and were the most field independent language group in the sample. Finally, when subgroups classified according to major academic area were compared, science/mathematics majors appeared to be the most field dependent while business/economics majors seemed to be the most field independent.

As these findings were unexpected, the distribution of the language groups in the categories of graduate/undergraduate and major academic area was reexamined. What emerged from this review was that the Spanish speakers as a group exhibited a number of characteristics which could complicate subsequent analysis: they formed a subgroup consisting solely of graduate students rather than a balance of graduates and undergraduates, they had the highest TOEFL means and MIT mean, and they were the most field dependent of the language groups. Further, all but one of the ten Spanish speakers identified themselves as science/mathematics majors; as such, they comprised nearly half of all the
science/mathematics majors in the sample. Because of these distinctions, the presence of the Spanish speakers had the potential for skewing the analysis of the role of FD/I in reading for the main idea.

Reliability of Measures

The strength of the findings from a study is dependent in part on the extent to which the measures produce consistent results. Test consistency was determined for this study by obtaining a statistical estimate of reliability for each of the measures used (see Table 27 and Table 28).

Information on students' responses to individual items on the TOEFL was not available for computing a reliability estimate. However, the Educational Testing Service (1987) estimated the reliability of the overall test to be .95 and that of the reading section to be .90 (see Table 27). Given these statistics and the continued and accepted use of the TOEFL as a test of English proficiency, it was assumed that the two measures were reliable for the students who comprised the study sample.

As noted in the previous chapter, one consequence of sampling students from two different sessions of IEOP was that their Michigan Test scores came from two different test
TABLE 27. Reliability estimates for the TOEFL: overall and reading section

<table>
<thead>
<tr>
<th>Score</th>
<th>Reliability</th>
<th>S.E. of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>.95</td>
<td>14.1</td>
</tr>
<tr>
<td>Reading</td>
<td>.90</td>
<td>2.4</td>
</tr>
</tbody>
</table>

TABLE 28. Reliability estimates for dependent and independent variables

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Split-half</th>
<th>K-R 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michigan</td>
<td>17</td>
<td></td>
<td>.76</td>
</tr>
<tr>
<td>Test</td>
<td>23</td>
<td></td>
<td>.60</td>
</tr>
<tr>
<td>GEFT(^a)</td>
<td>49</td>
<td>.88</td>
<td></td>
</tr>
<tr>
<td>MIT(^a, b)</td>
<td>49</td>
<td>.43</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)GEFT=Group Embedded Figures Test; MIT=main idea test.
\(^b\)Individual scores were calculated as the number correct out of a total of 10 items.

versions. Thus, reliability estimates were calculated for each version using the K-R 20 reliability formula (see Table 28). Neither test version had exceptionally high reliability for this sample, though the version used in the spring produced a higher estimate (.76) than the version used in the fall (.60).
The reliability of the GEFT was estimated using the split-half formula (see Table 28). An acceptable estimate of .88 was found for this measure.

The reliability of the main idea test (MIT) was estimated using the K-R 20 formula (see Table 28). The estimate of .43 raises some questions about the reliability of the test and must be considered when drawing conclusions from the findings yet to be discussed.

Estimating the reliability of the background knowledge survey seemed inappropriate as it was not intended to comprise a measure of a single construct. It was not expected that students would share the same level of experience with the topics on the MIT; in fact, it was hoped that the instrument would elicit variation in background knowledge from item to item, thus behaving "unreliably." However, it was important to know if students' responses on a background knowledge item predicted their performance on the corresponding MIT item. In other words, was the performance on each MIT item dependent on or independent of reported background knowledge in that subject area? To answer this question a chi-square test was performed between each background knowledge item and its corresponding MIT item (see Table 29). The chi-squares were computed from background knowledge surveys and main idea tests completed
by the entire sample (n=49). For the purposes of the chi-square, the five-point background knowledge scale was collapsed into two categories: little to no knowledge (1,2) and average to a lot of knowledge (3-5).

TABLE 29. Chi-square tests between background knowledge survey items and corresponding main idea test items

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-square</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>BKNOWL 1 by MIT 1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.00005</td>
<td>0.995</td>
</tr>
<tr>
<td>BKNOWL 2 by MIT 2</td>
<td>0.0</td>
<td>1.000</td>
</tr>
<tr>
<td>BKNOWL 3 by MIT 3</td>
<td>0.0</td>
<td>1.000</td>
</tr>
<tr>
<td>BKNOWL 4 by MIT 4</td>
<td>0.90304</td>
<td>0.342</td>
</tr>
<tr>
<td>BKNOWL 5 by MIT 5</td>
<td>0.61054</td>
<td>0.435</td>
</tr>
<tr>
<td>BKNOWL 6 by MIT 6</td>
<td>0.33828</td>
<td>0.561</td>
</tr>
<tr>
<td>BKNOWL 7 by MIT 7</td>
<td>0.00311</td>
<td>0.956</td>
</tr>
<tr>
<td>BKNOWL 8 by MIT 8</td>
<td>0.02720</td>
<td>0.869</td>
</tr>
<tr>
<td>BKNOWL 9 by MIT 9</td>
<td>0.0</td>
<td>1.000</td>
</tr>
<tr>
<td>BKNOWL 10 by MIT 10</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

<sup>a</sup>BKNOWL 1 by MIT 1=background knowledge survey item number 1 by main idea test item number 1.

* Chi-square could not be computed as no students circled 3-5 (average to a lot of knowledge) on the background knowledge survey; choices 3-5 made up the second column of the chi-square.
In all ten cases, performance on the MIT appeared to be independent of background knowledge as determined by this instrument. This finding runs counter to much of current reading research which has shown that previous knowledge of a subject enhances a student's comprehension of a passage about that subject (e.g., Carrell, 1983b; Carrell & Eisterhold, 1983; Johnson, 1981, 1982). Therefore, it would seem that the background knowledge survey provided an inadequate measure of the subjects' previous knowledge of the ten MIT topics.

Accounting for Variance on the Main Idea Test

Having determined the reliability of the measures in the study, the next step was to address the main question proposed in the research: is FOil related to the ability to find the main idea in reading? Factors other than FOil most certainly affect reading comprehension, among them reading proficiency and background knowledge. Because the background knowledge survey was found to be an inadequate measure, the consideration of background knowledge was foregone in the main research question for the remainder of the analysis. Thus, the original question was modified: does FOil predict performance in finding the main idea beyond what is accounted for by reading proficiency?
This revised question suggested the use of two principal types of measures which could be considered as predictors of performance on the main idea test. The first of these was the measure of FD/I, which was obtained from the GEFT scores. The second type of measure was that of reading proficiency: the reading section scores from the TOEFL and Michigan Test. (The overall scores from these two tests were eliminated from use in further analysis as the construct being tested here was only that of reading comprehension.)

Because the study sample comprised groups of students from two IEOP sessions, the question arose of whether they should be treated separately in the analysis of the main research question. A multiple regression was run to find out whether membership in either the spring or fall IEOP session contributed to variance on the MIT beyond that accounted for by FD/I (determined by the GEFT scores) and reading proficiency (determined by the Michigan Test reading scores). After FD/I and reading proficiency were entered into the regression, group membership did not contribute to variance on the MIT at a significant level (p=.285). Thus, the two groups could be treated as one in the subsequent analysis.
Following the above preliminaries, the first step in answering the main research question was to find out if the above measures were indeed related to performance on the MIT; that is, did high performance on each of these three independent variables coincide with high performance on the dependent variable? Pearson product moment correlations were obtained for the MIT scores with each of the independent variables (see Table 30).

TABLE 30. Pearson product moment correlations between dependent and independent variables for the entire sample

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TOFLRDG(^a)</td>
</tr>
<tr>
<td>MIT(^a,b)</td>
<td>.5229    (n=38)</td>
</tr>
<tr>
<td></td>
<td>p=.000***</td>
</tr>
</tbody>
</table>

\(^a\)TOFLRDG=TOEFL reading section score; EPRDG=Michigan Test reading section score; GEFT=Group Embedded Figures Test score; MIT=main idea test score.

\(^b\)Individual scores were calculated as the number correct out of a total of 10 items.

\(*p<.05; \***p<.001\).
The TOEFL reading scores correlated positively with the MIT scores at the p=.000 level of significance. The Michigan Test reading section scores did not correlate with the MIT scores at an acceptable level of significance (p=.301). This was not an unexpected result in light of the distorted distribution of scores on this measure as well as its lower reliability. Interestingly, field independence correlated negatively with the MIT scores at the p=.016 level of significance; that is, field dependence, rather than field independence, correlated with high MIT performance.

The correlational analysis produced two potential predictors for consideration in the main research question: field dependence as determined by the GEFT and reading proficiency expressed by the TOEFL reading score. The use of the Michigan Test was ruled out on the basis of its unreliability for this sample and its lack of correlation with the MIT. The unavoidable drawback in using the TOEFL reading scores as the measure of reading proficiency was that it reduced the sample size available for analysis; only 38 of the original 49 subjects had TOEFL reading scores on record.

A multiple regression analysis was used to determine what variance in MIT performance field dependence could
predict beyond what was accounted for by reading proficiency (see Table 31). In this procedure, the MIT scores served as the dependent variable with the TOEFL reading scores and the GEFT scores entered stepwise as independent variables.

**TABLE 31.** Multiple regression analysis using the main idea test scores as the dependent variable and the TOEFL reading and Group Embedded Figures Test scores as independent variables

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>N</th>
<th>Step</th>
<th>Variable entered</th>
<th>Partial</th>
<th>T value</th>
<th>Sig. of T</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>38</td>
<td>1</td>
<td>TOFRLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.503947</td>
<td>3.452</td>
<td>.002**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-.311986</td>
<td>-1.943</td>
<td>.060</td>
</tr>
</tbody>
</table>

<sup>a</sup>MIT=main idea test scores; TOFRLRDG=TOEFL reading section scores; GEFT=Group Embedded Figures Test scores.

<sup>b</sup>Individual scores were calculated as the number correct out of a total of 10 items.

**p<.01.

This procedure showed that the TOEFL reading scores contributed positively to variance at the p=.002 level of significance. Field independence was negatively correlated with performance on the MIT but did not contribute to variance at an acceptable level of significance (p=.060).
The finding that field independence was negatively correlated with MIT performance was an unexpected outcome given previous research in which field independence is generally associated with greater success in second language acquisition and in reading in a second language (e.g., Bialystok & Frohlich, 1978; Genesee & Hamayan, 1980). Thus, it seemed possible that factors other than reading proficiency and field dependence were involved in the results of this analysis. Given the nature of the subgroups within the sample, the next question became: in addition to field dependence, what student factors might contribute to variance in MIT performance beyond that accounted for by reading proficiency?

Though it would be interesting to examine the effect of several student factors on MIT performance, the small sample size precluded the use of more than a few independent variables in the regression analyses. Thus, of the three student factors discussed so far (graduate/undergraduate status, first language, and academic major), only graduate/undergraduate status was selected as an additional independent variable. This was because it was the only factor of the three for which significant between-groups variance in MIT performance was found in the subgroup ANOVAs. Therefore, an extension of the main research
question was: does field dependence predict performance on the MIT beyond what is accounted for by reading proficiency and graduate/undergraduate status?

A multiple regression was run with the MIT scores as the dependent variable and the TOEFL reading scores, graduate/undergraduate status, and the GEFT scores entered stepwise as independent variables (see Table 32).

**TABLE 32.** Multiple regression analysis with the main idea test scores as the dependent variable and the TOEFL reading scores, graduate/undergraduate status and the Group Embedded Figures Test scores as independent variables

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>N</th>
<th>Step</th>
<th>Variable entered</th>
<th>Partial</th>
<th>T value</th>
<th>Sig. of T</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT(^a,b)</td>
<td>38</td>
<td>1</td>
<td>TOFLRDG(^a)</td>
<td>.317624</td>
<td>1.953</td>
<td>.059</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>GU(^a)</td>
<td>.338829</td>
<td>2.100</td>
<td>.043*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>GEFT(^a)</td>
<td>-.180114</td>
<td>-1.068</td>
<td>.293</td>
</tr>
</tbody>
</table>

\(^a\)MIT=main idea test scores; TOFLRDG=TOEFL reading section scores; GU=graduate/undergraduate status; GEFT=Group Embedded Figures Test scores.

\(^b\)Individual scores were calculated as the number correct out of a total of 10 items.

\(^*\)p<.05.
This analysis showed that graduate status was the only positive significant predictor of variance ($p=.043$). The TOEFL reading scores made a positive contribution to variance, though not at an acceptable level of significance ($p=.059$). The contribution of field dependence was not significant ($p=.293$) when variance had been accounted for by the other two independent variables.

The unexpected finding that field dependence was correlated with MIT performance may well have been enhanced, if not produced, by several of the student factors previously mentioned. In other words, it was likely that the Spanish speakers were skewing the results as was suspected earlier. The fact that the graduate students across the sample achieved the highest means on the TOEFL reading section and the MIT as well as the lowest mean on the GEFT was coincident with the performance pattern of the Spanish speakers on these same measures. Considering that the all-graduate Spanish speakers ($n=10$) formed a substantial proportion of the sample graduate students ($n=28$), it seemed plausible that they played a significant part in determining the performance of graduate students in the sample as a whole. A similar situation existed for the science/mathematics majors on the GEFT means. The fact that nearly half of all the science/mathematics majors were
Spanish speakers, the most field dependent of the language groups, undoubtedly depressed the mean GEFT score for that major academic area.

Reanalysis of Variance on the MIT

In order to maximize the validity of the regression analyses, it was necessary to investigate the imbalance created by the presence of the Spanish speakers in the sample. It was obvious that to present even tentative conclusions at the end of the analysis it was necessary to see if the same results would be obtained with a sample not weighted by a single subgroup. Thus, the same series of questions and statistical analyses were applied to the sample minus the Spanish speakers.

Descriptive Statistics for the Reduced Sample

Descriptive statistics were obtained for the dependent variable (the MIT) and the independent variables which remained after the correlational analysis done on the original sample (the TOEFL reading and GEFT scores). The means, standard deviations, ranges, and kurtosis and skewness statistics for each of these measures using the reduced sample are found in Table 33.

When compared to the same statistics for the original sample (see Table 3), it was found that the TOEFL reading mean fell from 48.08 to 45.54; the MIT mean fell from 4.53
TABLE 33. Frequency statistics on dependent and independent variables for the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLRDG(^a)</td>
<td>28</td>
<td>45.54</td>
<td>5.87</td>
<td>30-61</td>
<td>2.18</td>
<td>-.23</td>
</tr>
<tr>
<td>GEFT(^a)</td>
<td>39</td>
<td>12.28</td>
<td>4.97</td>
<td>1-18</td>
<td>-.60</td>
<td>-.69</td>
</tr>
<tr>
<td>MIT(^a,b)</td>
<td>39</td>
<td>4.23</td>
<td>1.90</td>
<td>1-8</td>
<td>-1.03</td>
<td>.16</td>
</tr>
</tbody>
</table>

\(^a\) TOFLRDG=TOEFL reading section score; GEFT=Group Embedded Figures Test score; MIT=overall main idea test score.

\(^b\) Individual scores were calculated as the number correct out of a total of 10 items.

to 4.23; and the GEFT mean rose from 11.41 to 12.28. The standard deviations and ranges for the three measures from the reduced sample were comparable to those from the original sample. The skewness for each measure increased in the same direction established for the original sample but was still within the range considered close to normal. Differences were noted for the kurtosis of the TOEFL reading and MIT distributions. In both cases the kurtosis increased dramatically in the directions established in the original sample statistics, producing abnormality in the height of the distribution curves. This distortion in the two distributions places restrictions on the interpretation of the correlational analysis which follows.
The next step in characterizing the reduced sample was to examine it in terms of the previously established subgroups. Because the removal of the Spanish speakers involved graduate students only, the statistics on the undergraduate group remained the same. A new set of descriptive statistics were obtained for the altered graduate sample. In addition, an ANOVA was performed for the graduates and undergraduates to see if there was any change in sources of variance due to the reduction of the graduate sample. The statistics for the subgroups classified according to first language were also unaffected by the sample reduction, so a new set of descriptive statistics was not needed; however, a statistical comparison of the remaining three language groups was conducted within the reduced sample by means of an ANOVA. The final subgroups altered by the sample reduction were two of the four major academic areas: the science/mathematics and liberal arts majors. Descriptive statistics were rerun for these two groups; in addition, a post-sample-reduction ANOVA among all four major academic areas was performed.

The descriptive statistics on the graduate students from the reduced sample may be found in Table 34. A comparison of the graduate students from the two samples (see also Table 9) showed that when the Spanish speakers
were removed, the TOEFL reading mean fell from 51.27 to 48.00; the MIT mean fell slightly from 5.39 to 5.22; and the GEFT mean rose from 9.96 to 11.06. The differences in TOEFL reading and GEFT means paralleled those found when the reduced and original samples were compared as two overall groups above. That the MIT mean remained much the same reinforced the earlier finding that graduate students had greater success in identifying the main idea on the MIT.

TABLE 34. Frequency statistics on dependent and independent variables for graduates minus Spanish speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLRDGa</td>
<td>12</td>
<td>48.00</td>
<td>5.80</td>
<td>37-61</td>
<td>2.19</td>
<td>.58</td>
</tr>
<tr>
<td>GEFTa</td>
<td>18</td>
<td>11.06</td>
<td>5.46</td>
<td>1-18</td>
<td>-.94</td>
<td>-.44</td>
</tr>
<tr>
<td>MITa,b</td>
<td>18</td>
<td>5.22</td>
<td>1.52</td>
<td>3-8</td>
<td>-.91</td>
<td>.03</td>
</tr>
</tbody>
</table>

aTOFLRDG=TOEFL reading section score; GEFT=Group Embedded Figures Test score; MIT=overall main idea test score.

bIndividual scores were calculated as the number correct out of a total of 10 items.
An ANOVA of the performance of graduate and undergraduate students from the reduced sample was run to determine if there were significant sources of variance between the two groups (see Table 35). This analysis showed that the MIT was the only measure on which the two groups varied significantly ($p=.002$). This finding was in contrast to the same ANOVA run on the original sample (see Table 11) which showed significant variance between graduates and undergraduates on the TOEFL reading section and the GEFT as well as the MIT. Thus, removing the Spanish speakers from the sample served to reduce the variance accounted for by the performance of graduates and undergraduates.

As mentioned earlier, the basic descriptive statistics on the remaining language subgroups did not change when the Spanish speakers were removed; however, the ways in which the performance of these subgroups varied with respect to each other could have been altered. An ANOVA was run using the three remaining language groups (see Table 36); it showed no significant source of variance among these groups on the dependent and independent variables.

The last comparison between the reduced and original samples was made using the subgroups classified by major academic area. This comparison was made only for the science/mathematics and liberal arts majors; because the
TABLE 35. ANOVA between graduate and undergraduate performance on dependent and independent variables for the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Source of var.</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLRDG&lt;sup&gt;b&lt;/sup&gt;</td>
<td>28</td>
<td>BG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1</td>
<td>127.53</td>
<td>4.14</td>
<td>.052</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG&lt;sup&gt;c&lt;/sup&gt;</td>
<td>26</td>
<td>30.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;b&lt;/sup&gt;</td>
<td>39</td>
<td>BG</td>
<td>1</td>
<td>50.29</td>
<td>2.09</td>
<td>.157</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>37</td>
<td>24.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT&lt;sup&gt;b,d&lt;/sup&gt;</td>
<td>39</td>
<td>BG</td>
<td>1</td>
<td>32.86</td>
<td>11.68</td>
<td>.002**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>37</td>
<td>2.81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Variance between groups.
<sup>b</sup>TOFLRDG=TOEFL reading section score; GEFT=Group Embedded Figures Test score; MIT=overall main idea test score.
<sup>c</sup>Variance within groups.
<sup>d</sup>Individual scores were calculated as the number correct out of a total of 10 items.

**p<.01.

Spanish speakers reported their majors in these two areas alone, their removal from the sample did not alter the basic statistics for the business/economics and engineering majors. The descriptive statistics for the two pertinent majors in the reduced sample are reported in Table 37 and
TABLE 36. ANOVA among the performance of language groups on dependent and independent variables for the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Source of var</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>F prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLRDG&lt;sup&gt;b&lt;/sup&gt;</td>
<td>28</td>
<td>BG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2</td>
<td>29.12</td>
<td>.84</td>
<td>.445</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG&lt;sup&gt;c&lt;/sup&gt;</td>
<td>25</td>
<td>34.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEFT&lt;sup&gt;b&lt;/sup&gt;</td>
<td>39</td>
<td>BG</td>
<td>2</td>
<td>25.93</td>
<td>1.05</td>
<td>.360</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>36</td>
<td>24.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT&lt;sup&gt;b,d&lt;/sup&gt;</td>
<td>39</td>
<td>BG</td>
<td>2</td>
<td>2.72</td>
<td>.74</td>
<td>.483</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>36</td>
<td>3.65</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Variance between groups.

<sup>b</sup>TOFLRDG=TOEFL reading section score; GEFT=Group Embedded Figures Test score; MIT=overall main idea test score.

<sup>c</sup>Variance within groups.

<sup>d</sup>Individual scores were calculated as the number correct out of a total of 10 items.

Table 38; these are compared with the same statistics obtained for the original sample (see Table 23 and Table 24).

The Spanish speakers made up nearly half of all the science/mathematics majors, so their absence from the sample was expected to alter the descriptive statistics for that
TABLE 37. Frequency statistics on dependent and independent variables for science/mathematics majors from the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLRDGA</td>
<td>8</td>
<td>45.63</td>
<td>7.91</td>
<td>35-61</td>
<td>1.61</td>
<td>.71</td>
</tr>
<tr>
<td>GEFTA</td>
<td>11</td>
<td>10.64</td>
<td>5.12</td>
<td>1-17</td>
<td>-.55</td>
<td>-.63</td>
</tr>
<tr>
<td>MITA,B</td>
<td>11</td>
<td>4.73</td>
<td>1.74</td>
<td>2-7</td>
<td>-1.34</td>
<td>-.19</td>
</tr>
</tbody>
</table>

  aTOFLRDG=TOEFL reading section score; GEFT=Group Embedded Figures Test score; MIT=overall main idea test score.

  bIndividual scores were calculated as the number correct out of a total of 10 items.

TABLE 38. Frequency statistics on dependent and independent variables for liberal arts majors from the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>Range</th>
<th>Kurt.</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLRDGA</td>
<td>5</td>
<td>45.60</td>
<td>2.88</td>
<td>41-48</td>
<td>1.33</td>
<td>-1.22</td>
</tr>
<tr>
<td>GEFTA</td>
<td>7</td>
<td>11.00</td>
<td>6.81</td>
<td>2-18</td>
<td>-2.20</td>
<td>-.31</td>
</tr>
<tr>
<td>MITA,B</td>
<td>7</td>
<td>4.86</td>
<td>1.68</td>
<td>2-7</td>
<td>.05</td>
<td>-.58</td>
</tr>
</tbody>
</table>

  aTOFLRDG=TOEFL reading section score; GEFT=Group Embedded Figures Test score; MIT=overall main idea test score.

  bIndividual scores were calculated as the number correct out of a total of 10 items.
group. The TOEFL reading mean fell from 50.53 to 45.63 with the standard deviations and ranges for the two subsamples remaining nearly the same. The kurtosis and skewness statistics for the reduced sample changed considerably. The kurtosis reversed direction (from -.43 to 1.61), which produced an abnormal height in the distribution in the reduced sample; the skewness also changed direction (from -.40 to .71), but was still considered close to normal. The GEFT mean rose from 9.65 to 10.64. The MIT mean fell slightly from 5.10 to 4.73; the standard deviation increased slightly from 1.65 to 1.74; and the range decreased from 7 to 5. The kurtosis and skewness for the MIT were different than in the original subsample. The kurtosis dropped from .49 to -1.34, thus producing an abnormally flattened curve; the skewness dropped from .21 to -.19, but was still near normal.

There was only one Spanish speaker in the liberal arts major area, but the removal of this one group member altered the statistical picture of the subgroup substantially. The TOEFL reading mean fell from 47.67 to 45.60; the standard deviation decreased from 5.68 to 2.88; and the range dropped from 17 to 7. The kurtosis fell from 2.78 to 1.33, which lowered the height of the distribution, but was still higher than could be considered normal. The skewness changed
direction completely, falling from 1.28 to -1.22; the distribution was still abnormally skewed though in the opposite direction. Thus, the removal of one Spanish speaker changed the TOEFL reading score distribution considerably.

The GEFT mean for the reduced liberal arts subsample rose from 10.13 to 11.00, while the standard deviation, range and kurtosis remained much the same. The skewness decreased from .05 to -.31, which shifted the direction of the skew but could still be considered as close to normal.

The MIT mean for the reduced subsample fell from 5.13 to 4.86, with the standard deviation and range changing little from the original liberal arts subgroup. The kurtosis shifted slightly from -.10 to .05 but could still be considered as close to normal. The skewness also shifted, from -.70 to -.58, and was still within the range considered close to normal.

An ANOVA was performed to see if significant variance occurred between the major academic subgroups without the Spanish speakers (see Table 39). No significant variance between groups was found for the dependent and independent variables. In the original subsample, there was significant variance between groups on the GEFT (see Table 25); the removal of the more field dependent Spanish speakers
appeared to have leveled the variance between academic major groups on the GEFT.

TABLE 39. ANOVA among the performance of major academic groups on dependent and independent variables for the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Source of var.</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOFLRDG(^b)</td>
<td>28</td>
<td>BG(^a)</td>
<td>3</td>
<td>2.99</td>
<td>.08</td>
<td>.971</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG(^c)</td>
<td>24</td>
<td>38.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEFT(^b)</td>
<td>36</td>
<td>BG</td>
<td>3</td>
<td>43.77</td>
<td>1.83</td>
<td>.162</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>32</td>
<td>23.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT(^b, d)</td>
<td>36</td>
<td>BG</td>
<td>3</td>
<td>2.14</td>
<td>.58</td>
<td>.634</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WG</td>
<td>32</td>
<td>3.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Variance between groups.

\(^b\) TOFLRDG=TOEFL reading section score; GEFT=Group Embedded Figures Test score; MIT=overall main idea test score.

\(^c\) Variance within groups.

\(^d\) Individual scores were calculated as the number correct out of a total of 10 items.
Summary of Reduced Sample Characteristics

In general, the comparison of the descriptive statistics from the original and reduced samples showed that the removal of the Spanish speakers had the same effect on the reduced sample as a whole as for the subgroups of graduates and the science/mathematics majors. Without the high TOEFL reading scores generated by the Spanish speakers, the TOEFL reading mean dropped. When the most field dependent group in the study was eliminated (the Spanish speakers), the GEFT mean increased. Finally, when the subgroup with the highest MIT scores was removed (again the Spanish speakers), the MIT mean decreased or stayed much the same. The shape of the distribution curves for the measures changed markedly in some instances, both for the reduced sample as a whole and for some of its subgroups. The primary concern here was for the changes in kurtosis found for the TOEFL reading and MIT scores from the overall reduced sample. These distortions in the normal distributions for the two measures naturally compromise the validity of the correlational analysis to follow, but were an unavoidable function of the restricted sample size.

Accounting for Variance on the MIT: the Reduced Sample

As with the original sample, the next step was to see if performance on the independent variables correlated with
performance on the dependent variable. Pearson product moment correlations were obtained for the MIT scores with both of the independent variables (see Table 40). The findings from this set of correlations paralleled those obtained using the original sample. The TOEFL reading scores correlated positively with the MIT scores at a significant level ($p=.002$). Field independence correlated negatively with the MIT scores at the $p=.047$ level. As with the original sample, field dependence and not independence correlated with high MIT performance. It should be remembered, however, that both independent variables had abnormal kurtosis statistics (see Table 33), so the significance of the correlations must be interpreted with caution.

The next question was whether field dependence predicted performance on the MIT beyond what was accounted for by reading proficiency. A multiple regression was run with the MIT scores as the dependent variable, and the TOEFL reading and GEFT scores entered stepwise as independent variables (see Table 41). This procedure showed that like the original sample (see Table 31), the TOEFL reading scores made a positive contribution to variance at an acceptable level of significance ($p=.001$). Field dependence made a greater positive contribution to variance at a significant level ($p=.017$) than was found for the original sample.
TABLE 40. Pearson product moment correlations between dependent and independent variables for the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>MIT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>0.5348 (n=28)</td>
</tr>
<tr>
<td></td>
<td>p=0.002**</td>
</tr>
</tbody>
</table>

<sup>a</sup>TOFLRDG=TOEFL reading section score; GEFT=Group Embedded Figures Test score; MIT=main idea test score.

<sup>b</sup>Individual scores were calculated as the number correct out of a total of 10 items.

**p<.01.

TABLE 41. Multiple regression analysis with the main idea test as the dependent variable and the TOEFL reading and Group Embedded Figures Test scores as independent variables for the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>N</th>
<th>Step</th>
<th>Variable entered</th>
<th>Partial</th>
<th>T value</th>
<th>Sig. of T</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>28</td>
<td>1</td>
<td>TOFLRDG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.610103</td>
<td>3.850</td>
<td>.001**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>GEFT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.456391</td>
<td>-2.565</td>
<td>.017*</td>
</tr>
</tbody>
</table>

<sup>a</sup>MIT=main idea test scores; TOFLRDG=TOEFL reading section scores; GEFT=Group Embedded Figures Test scores.

<sup>b</sup>Individual scores were calculated as the number correct out of a total of 10 items.

*p<.05; **p<.01.
The remaining question for the reduced sample was: would field dependence predict performance on the MIT beyond what was accounted for by reading proficiency and graduate/undergraduate status? A multiple regression was run with the MIT scores as the dependent variable and the TOEFL reading scores, graduate/undergraduate status and the GEFT scores entered stepwise as independent variables (see Table 42). The results produced were different from those found when the same analysis was applied to the original sample (see Table 32).

TABLE 42. Multiple regression analysis with the MIT scores as the dependent variable and the TOEFL reading scores, graduate/undergraduate status and GEFT scores as independent variables for the sample minus Spanish speakers

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>N</th>
<th>Step</th>
<th>Variable entered</th>
<th>Partial</th>
<th>T value</th>
<th>Sig. of T</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT\textsuperscript{a,b}</td>
<td>28</td>
<td>1</td>
<td>TOFLRDG\textsuperscript{a}</td>
<td>.504540</td>
<td>2.863</td>
<td>.009**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>GU\textsuperscript{a}</td>
<td>.340025</td>
<td>1.771</td>
<td>.089</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>GEFT\textsuperscript{a}</td>
<td>-.350476</td>
<td>-1.833</td>
<td>.079</td>
</tr>
</tbody>
</table>

\textsuperscript{a} MIT=main idea test scores; TOFLRDG=TOEFL reading section scores; GU=graduate/undergraduate status; GEFT=Group Embedded Figures Test scores.

\textsuperscript{b} Individual scores were calculated as the number correct out of a total of 10 items.

**p<.01.
Unlike the original regression results, the TOEFL reading scores, not graduate/undergraduate status, were the only positive predictor with an acceptable level of significance (p=.009). The contribution of graduate/undergraduate status was again in the positive direction but not at a significant level (p=.089). The contribution of field dependence was not significant (p=.079) after the TOEFL scores and graduate/undergraduate status were entered into the regression; however, as the direction of the correlation was again negative, it would seem that field dependence rather than independence was the cognitive style which was associated with variance.

Summary of Reanalysis of Variance on the MIT

Reducing the original sample by eliminating the Spanish speakers affected the outcome of the analysis in a number of ways. The statistics used to characterize the original sample and its subgroups suggested that the Spanish speakers were weighting the data in several areas; they were the most proficient readers, the most field dependent, and skewed as a subgroup in terms of the proportion of graduates to undergraduates. When the Spanish speakers were removed, the ratio of graduates to undergraduates in the reduced sample became more balanced. Furthermore, there were fewer significant between-group differences noted within each of
the three categories of subgroups (graduates/undergraduates, first languages, and academic majors); significant differences were noted only between graduates and undergraduates in their performance on the MIT.

The removal of the Spanish speakers also had an effect on the regression analysis used to address the main research question. In the second of the original sample regressions (which included the independent variables of TOEFL reading scores, graduate/undergraduate status, and GEFT scores), graduate/undergraduate status emerged as the only positive significant predictor of MIT performance. When the Spanish speakers were excluded from the same regression, the TOEFL reading scores became the only positive significant predictor.

The interesting finding in both series of analysis was that a tendency toward field dependence, not independence, appeared to be correlated with MIT performance. Field dependence cannot be reported as a predictor of MIT performance as the contribution it made did not reach an acceptable level of significance in either the original or reduced sample regressions after the TOEFL reading scores and graduate/undergraduate status were entered. However, the fact that field dependence contributed to MIT performance at the $p=.079$ level in the final regression (see
Table 42 suggests that further research on the role of field dependence in main idea comprehension could be fruitful.
DISCUSSION

This chapter provides a discussion of the results of the data analysis. It includes a recapitulation of the statistical procedures and the results they yielded, an interpretation of the findings, a discussion of the limitations of the study, and some suggestions for further research.

Review of Statistical Analyses

Frequency statistics were obtained on the seven measures for the entire sample and the three subgroups classified by graduate/undergraduate status, first language, and major academic area. Analyses of variance were conducted for each of the subgroups. The results showed that the graduates were significantly more proficient on the TOEFL and the MIT, and significantly more field dependent (FD) than the undergraduates. Statistically significant differences were also noted for the two language groups which comprised the extremes of the range of mean scores: the Spanish and Japanese speakers. The Spanish speakers were the most proficient on the TOEFL and the most FD of the language groups. Conversely, the Japanese speakers were the least proficient on the TOEFL and the most field independent (FI) of the language groups. Finally, a statistically
significant difference was found between two of the major academic areas on the GEFT. The science/mathematics majors, of whom nearly half were Spanish speakers, were the most FD group in the sample, and the business/economics majors were the most FI.

Following these preliminary analyses, reliability estimates were obtained for several of the measures. The TOEFL (overall and reading section) and the GEFT had acceptable reliabilities, while the Michigan Test and the MIT had less acceptable reliabilities, with the MIT showing the least acceptable estimate of all. To determine if the background knowledge survey was an adequate measure of prior knowledge, a chi-square test was performed for each survey item and its corresponding MIT item. The results indicated that performance on the MIT items was independent of prior knowledge as measured by the survey.

The findings from the reliability estimates and the chi-square tests were used to narrow the field of potential predictor variables to three: the TOEFL reading scores, the Michigan Test reading scores, and the GEFT scores. Pearson product moment correlations were obtained for each of these independent variables and the dependent variable, the MIT scores. The TOEFL reading and GEFT scores correlated significantly with the MIT scores; the Michigan Test reading
scores did not show a significant correlation with the MIT scores and so were eliminated from further analysis. A notable finding from the correlational analysis was that FD, not FI, was associated with high MIT scores.

The first multiple regression analysis was run to determine if FI accounted for variance in performance on the MIT beyond what could be accounted for by reading proficiency. The results showed that the TOEFL reading scores were a positive significant predictor of variance; FD, not FI, was related to MIT performance though not at a statistically significant level. Because it was unexpected to find a positive correlation between FD and MIT performance, it was suspected that other learner variables might be affecting the outcome of the regression. Consequently, a second multiple regression was performed with the learner variable of graduate/undergraduate status entered as a third possible predictor of performance. This regression showed that graduate status was the only positive significant predictor; the TOEFL scores and FD were positively correlated with MIT performance though not at a significant level.

Taken with the results of the previous analyses of variance, these findings made an even stronger case for a skewing effect originating with the Spanish speakers.
Because they were all graduate students, were the most proficient on the measures of reading comprehension, and were the most field dependent, it seemed likely that the Spanish speakers were weighting the results of the regressions. Therefore, the Spanish speakers were dropped from the sample, and the analysis was repeated with the remaining three language groups. Frequency statistics for the reduced sample were obtained for the dependent variable (the MIT scores) and only those independent variables which were significantly correlated with MIT performance in the original sample (the TOEFL reading and GEFT scores).

Subsequent analyses of variance within the subgroups (graduates/undergraduates, the remaining three language groups, and the major academic areas) showed the effect of removing the Spanish speakers to be one of significantly reducing between-group variance; the only statistically significant source of variance which remained within any of the subgroups was between the graduate and undergraduate students on the MIT.

The Pearson product moment correlations for the reduced sample corroborated those found for the original sample; a high level of reading proficiency (measured by the TOEFL reading scores) and FD (measured by the GEFT scores) were associated with high performance on the MIT. The first
regression analysis showed that as in the original sample, the TOEFL reading scores were a significant positive predictor of MIT performance; interestingly, FD was shown to be a positive and significant predictor as well, a result which was not obtained at a significant level in the original sample. The second regression analysis showed that unlike the results for the original sample, reading proficiency (not graduate/undergraduate status) was the only positive significant predictor of performance. Again, it was interesting to find that despite a lack of statistical significance, FD was the operative cognitive style.

Interpretation of Results

The regression analyses for the entire sample produced some unexpected results. The first regression showed reading proficiency to be the only positive significant predictor of MIT performance. However, reading proficiency was rendered nonsignificant in the second regression when graduate/undergraduate status was introduced; graduate status appeared to be the only positive significant predictor of MIT performance. It was suspected that the all-graduate Spanish speakers were skewing the results, a suspicion which was confirmed by the second regression on the reduced sample; with the Spanish speakers removed,
graduate status no longer predicted MIT performance, and reading proficiency was restored as the only positive significant predictor.

The reasons why the Spanish speakers had such an impact on the sample have already been stated; their coincident all-graduate status and high MIT performance made graduate status stand out as a significant predictor. Had they comprised undergraduates as well as graduates, the skewing effect may well have been removed. If there had been Spanish-speaking undergraduates in the sample who performed the same way as the actual undergraduates in the study, their TOEFL and MIT scores would have been lower than those of their graduate counterparts. Thus, the presence of undergraduates might have lowered the means on these measures for the entire Spanish-speaking group.

Because of the skewing effect of the Spanish speakers, the results of the regressions on the reduced sample are more interesting. The first regression on the reduced sample showed that both reading proficiency and FD were positive significant predictors of MIT performance. The finding of FD as a significant predictor after the Spanish speakers were removed shows that they did not skew the results for this variable after all; the Spanish speakers were the most FD of the language groups, and yet FD emerged
as a significant predictor even when they were excluded from analysis. However, FD was once again rendered nonsignificant when graduate/undergraduate status was introduced into the analysis; the second regression on the reduced sample showed that reading proficiency was the only positive significant predictor.

It is difficult to explain why graduate status was related to MIT performance; graduate status does not automatically indicate greater ability in reading or proficiency on comprehension tests. If the graduate students of this sample had earned undergraduate degrees in universities where the language of instruction was English, then perhaps a case could be made for their being more practiced in reading English. However, the quality and quantity of reading the students in the sample had experienced is not known, so it is uncertain that such a condition was met. Perhaps the graduate students comprised a more select group academically because they were more proficient readers in their first languages. If reading proficiency in the first language transfers to reading proficiency in the second language, then this could help explain the influence of graduate status on MIT performance in the regression results.
The finding in the second reduced sample regression that reading proficiency predicted MIT performance was not unexpected; one would naturally expect students who have demonstrated a high level of general reading proficiency to perform in a similar manner on a test for a specific skill such as finding the main idea. However, it is interesting to consider that FD has consistently correlated with MIT test performance even though it was not a significant predictor of performance at the end of the regression analyses. In light of much of the previous research on FD/I and language acquisition, this is a most unexpected finding; the literature is not without studies which have found a lack of evidence for the predictive role of FI in language tasks, but no studies prior to this one have identified tasks in which FD predicts performance.

It is always risky to attempt to generalize findings from a sample as small as this one; after the Spanish speakers were removed, only twenty eight students remained for analysis. However, if the discovery of FD as a positive predictor of MIT performance had been statistically significant in the final regression analysis, it would have added a most interesting piece to the puzzle of ESL reading comprehension and main idea processing. In the review of previous research, it was noted that the process of main
idea comprehension has as yet escaped definition; no one has developed a model which explains what types of abilities are utilized to identify main ideas in expository prose. Had FD emerged as a significant predictor in this study, the implication might be that main idea comprehension is accomplished through holistic processing, that somehow a more global mode of perception enables an individual to identify a global sort of main idea inherent in an expository passage.

This speculation provides an interesting theoretical base from which to begin, but probably touches only the tip of the iceberg. For example, previous research suggests that the location of the main idea in a passage may affect readers' ability to identify it (Kimmel & MacGinitie, 1984). The main idea test in this study was constructed so as to include passages with the main idea in initial, medial, and final positions, but the analysis was not structured to consider the effect of this factor in main idea identification. It is conceivable that, depending on the location of the main idea in the passage, different or flexible modes of processing might be advantageous. Kimmel and MacGinitie looked at passages organized deductively (main idea first) and inductively (main idea last), and found that certain readers had more difficulty with one than
the other; perhaps FD is more useful for indentifying the
main idea in one type of passage and FI more useful for main
idea comprehension in another type.

Limitations of the Study

Though they suggest some interesting directions for
further research, the results of this study do not provide
conclusive evidence of the role of FD/I in performance on a
test of reading for the main idea. Even if the results were
statistically significant, they would still have to be
interpreted in light of several limitations present in the
study. As was noted earlier, there are several theoretical
and practical problems with the instruments used to measure
performance. The main idea test did not obtain a very high
reliability estimate (.43); its use as a measure of main
idea comprehension is thus of questionable value. The
background knowledge survey proved to be inadequate to
measure individual students' levels of prior knowledge of
the topics on the main idea test, so it is impossible to
know how background knowledge utilization interacted with
the other variables in predicting main idea test
performance. Furthermore, though the TOEFL was assumed to
be reliable for the sample in this study, the time interval
between the main idea test administration and the most
proximal TOEFL test results varied widely among many of the subjects; thus, it is difficult to say with certainty that the TOEFL scores were accurate measures of some students' proficiency at the time they took the main idea test.

There are also several practical limitations to consider which are related to the administration of the test instruments and the statistical analyses which followed. Because of limitations on the amount of class time which could be devoted to the data collection of this study, all the test instruments were administered in one two-hour block of time. The advantage in administering the tests and other instruments all at once was that it prevented the loss of data due to absenteeism. However, the disadvantage was that it made for a rather taxing two hours; the effect of test fatigue on test performance was undoubtedly important for some of the students. Finally, the findings of the study are constrained by the absence of data such as TOEFL scores for many of the subjects. This gap in the data was especially noticeable after the Spanish speakers were removed from the sample, as they comprised a large proportion of the total students who had TOEFL scores on record. The absence of data restricted the sample size available for the regression analyses; the small sample size in turn limits the extent to which the results can be generalized.
Suggestions for Further Research

In spite of the gaps and inconsistencies which limit the validity of the findings, this study contributes to the body of extant research by proposing further areas of investigation. For example, it would be very interesting to replicate the study with a larger sample to see if FD would predict performance in finding the main idea. Since the literature has demonstrated that background knowledge utilization has a significant effect on reading comprehension, any replication of this study should include the results of a valid measure of prior knowledge. Predictors other than those examined here should also be considered, such as intelligence and reading proficiency in the first language; both of these undoubtedly play a part in general reading comprehension and very likely would affect main idea comprehension as well. As was suggested earlier, it would be interesting to analyze the effect of the position of the main idea in text to see if FD or FI cognitive styles facilitate main idea comprehension in different types of rhetorical patterns.

These are but a few of the many conceivable variations and improvements in research design which could lead to a fuller understanding of the role of FD/I in main idea comprehension. As these avenues are explored, our
understanding of the process as well as the product of main idea comprehension will expand. This should facilitate improved instruction in finding and using main ideas, and should greatly benefit ESL students whose success in school may depend on understanding and using salient information.


Alderson, J. C., & Urquhart, A. H. (1985b). This test is unfair: I'm not an economist. In P. C. Hauptman, R. LeBlanc, & M. Bingham (Eds.), *Second language performance testing* (pp. 25-43). Ottawa: University of Ottawa Press.


ACKNOWLEDGEMENTS

This study would not have been possible without the participation and assistance of many people. First, I would like to thank the students and teachers who participated in the pilot testing and review of the main idea test; their comments and suggestions were most helpful, and I sincerely appreciate the time and effort they gave to the project. Next, I would like to thank Dr. Roberta Vann and Dr. Barbara Matthies for allowing me to seek the participation of IEOP students and teachers in the collection of the data for the study. Very special thanks go to the IEOP teachers who allowed me the class time to test their students. Most of all, I would like to thank the IEOP students who agreed to participate as the subjects of the study; I have learned a great deal from all of them and hope that the learning experience in working with the study has been mutual.

The process of carrying out this research study has truly been a team effort; for this reason, I would like to express my most heartfelt thanks to Dr. Roberta Abraham and

---

Permission to conduct this study using students at Iowa State University was granted by the University Human Subjects Review Committee in February of 1987. Permission to seek the participation of students in the Intensive English and Orientation Program (IEOP) was granted in the spring of 1987 by the Director of the IEOP, Dr. Roberta Vann; permission to again seek the participation of IEOP students was granted in the fall of 1987 by the Acting Director, Dr. Barbara Matthies.
Dr. Carol Chapelle. The time and effort they have devoted to the development and realization of this study has been invaluable; it goes without saying that this study would not have been possible without their input and participation throughout the research process. I would also like to thank Dr. Joanna Courteau for her interest, comments, and suggestions in the early stages of designing the study.

I have been very fortunate to have professors on my graduate committee who have gone beyond the requirements of guiding the progress of a research study; each of them has made a unique contribution to my education on a holistic level. I would like to thank Dr. Leon Apt for his unflagging support and confidence through the years. His belief that teaching begins with nurturing has provided me with a living model of the teacher I hope to become and has made being his student a very special joy. I would also like to thank Dr. Carol Chapelle for leading me through learning experiences that I originally thought were beyond my capabilities. With patience and persistence, she has shown me that belief in the student's ability to learn can expand the student's view of what is possible to achieve.

Finally, I would like to thank Dr. Roberta Abraham for being in the right place at the right time several years ago. As the instructor of my first course in the field of TESL, she
saw potential in me that I had been unaware of. It was chiefly because of her belief that I had something to contribute that I chose to pursue this field of study. Her encouragement, guidance, and constructive criticism have contributed much to my education, and I will always be grateful for the role she has played in who I have become.
APPENDIX A: FIRST LANGUAGES AND HOME COUNTRIES REPRESENTED IN THE SAMPLE

Arabic
Chad
Egypt
Iraq
Jordan
Saudi Arabia
Sudan
Syria
United Arab Emirates

Farsi
Iran
French
Haiti
Indonesian
Indonesia

Japanese
Japan
Korean

Spanish
Bolivia
Colombia
Dominican Republic
Ecuador
El Salvador
Mexico
Nicaragua
Panama
Paraguay
Mandarin
Malaysia
Taiwan
Taiwanese
Taiwan
Turkish
Turkey
APPENDIX B: ACADEMIC MAJORS REPRESENTED IN THE SAMPLE

Business/economics

Business administration  
Economics  
Hotel management

Engineering

Civil engineering  
Computer engineering  
Electrical engineering  
Industrial engineering  
Mechanical engineering  
Oceanographic engineering

Science/mathematics

Agronomy  
Animal science  
Climatology  
Computer science  
Entomology  
Horticulture  
Marine biology  
Mathematics  
Medicine  
Plant pathology  
Statistics  
Veterinary medicine  
Zoology

Liberal arts

Architecture  
Child development  
English  
Physical Education  
Political science  
Psychology  
Sociology
APPENDIX C: TESTING MATERIALS
Consent Form

You are invited to participate in a study which I hope will tell us more about how people learn English as a second language. I am conducting this study to fulfill the research requirement for my master's degree in English; the results will appear in my final thesis.

The results of this study should tell more about how people with different learning styles read for the main idea. Participating in this study will give you practice in reading for the main idea in a timed test situation. This type of practice may prove helpful to you in preparing to take the reading section of the TOEFL or other tests of English proficiency.

If you decide to participate, you will be asked to do the following:

-- take a test which gives an indication of your individual style of learning. This test takes about 20 minutes.

-- take a reading test in which you will be asked to read a passage and select the main idea of the passage from several choices given at the end of the reading. There will be 10 passages on this test, and it will take about an hour and a half to complete.

-- fill out a short questionnaire which will indicate how much background knowledge you have of the topics in the reading test. This will be given before the reading test and will take about 10 minutes.

In order to analyze the results of the above tests, I would like to use your English Placement Test scores and your TOEFL Reading Section scores in my statistical analysis. All of your test scores will be kept confidential. They will not in any way affect your grades in your English courses. Reports of this research will not use the names of the people participating in it.

I will be glad to share your test results with you after I have evaluated them. If you have any questions, I will be happy to answer them now or later.

Danik Wold
206 Ross Hall (mailbox)
294-4109

You are making a decision whether or not to participate in this study. Your signature indicates that you have read the information provided above and have decided to participate. You may withdraw at any time after signing this form if you choose to discontinue participation in this study.
PERSONAL DATA FORM

Date: ____________________

Name: ___________________________ Age: ________________

Nationality: ____________________________________________

First Language: ___________________________________________

Other Languages: __________________________________________

Academic Major:

Undergraduate: _______ Graduate: _______

Do you already have a college or university degree? YES NO

IF YES, what was the major area of study? ____________________________

Was this degree from a university where the language of instruction was English? YES NO

How long have you been studying English? ____________________________

When did you first come to the United States?

Month: ________________ Year: ________________
Circle the number that best describes the amount of knowledge you have of the following topics.

<table>
<thead>
<tr>
<th>Topic</th>
<th>nothing</th>
<th>not very much</th>
<th>an average amount</th>
<th>more than most people</th>
<th>a lot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robots in the Workplace</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Language Functions in the Brain</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Predicting Population Growth</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Weather Forecasting</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Changes in the Structure of the American Family</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Using a Computer</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>The Effects of Exercise on Mental Health</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>How Small Group Discussions Affect Opinion</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Animals as a Source of Power in Farming</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Early History of Chaco Canyon</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
INSTRUCTIONS

In this test, you will read ten passages and select the main idea of each passage from four choices. The amount of time you will have to read each passage and choose the main idea is given at the top of each page.

Some vocabulary terms that may be unfamiliar to you are underlined in the passage and defined in the margins to the side of each term. If you already know the meaning of the underlined words, don't waste time in reading the definitions; read the passage and focus on finding the main idea.

When I tell you to begin, turn to the passage and begin reading for the main idea. When I say "STOP," stop reading even if you are not finished with the passage. If you finish reading before I tell you to stop, do not turn the page yet. Also, do not go back to a previous passage.

When I tell you, turn the page and read the four choices for the main idea of the passage. Choose the one answer that best describes the main idea of the entire passage. You may look back at the passage if you wish. Make a check mark (✓) in the blank next to your choice. When I say "STOP," stop reading even if you are not finished choosing the main idea. Then wait for my signal to go to the next passage.
Robots in the Workplace

Robots have received a great deal of publicity recently. The enormous influx of foreign cars manufactured in part by robots has called attention to the serious effects that robots have on industrial productivity.

Most people think of a robot as something out of the movie "Star Wars," an android that walks and talks, sees and feels, and looks much like C3PO or R2D2. Real robots are much more primitive. In its simplest form, a robot is nothing more than a mechanical device that can be programmed to perform some useful act under automatic control. An industrial robot is a device that can be programmed to move some gripper or tool through space to accomplish a useful industrial task.

These robots are usually programmed by recording each task as a series of points in space. This recording is then simply replayed whenever the task is to be performed. This simple procedure is adequate to perform a surprising number of industrial tasks, such as spot welding automobile bodies, tending diecasting machines, and spray painting.

But the great majority of industrial tasks are beyond the capacities of current robot technology. Most tasks are too complex and unstructured, involve too many uncertainties, or require too much ability to see, feel, and adapt to changing circumstances. Before robots can significantly improve the productivity of the economy as a whole, they must be used in hundreds of thousands — even millions — of applications. This use will not be possible until robots are designed that are more skillful, mobile, and intelligent.

Many people today believe that the robot revolution is well under way. They believe that factories are full of armies of highly intelligent robots and that large numbers of human workers are losing their jobs to robots. The facts are quite different.

There are only about 4,000 to 5,000 robots installed in U.S. factories today, according to Robotics International, an association of manufacturers of industrial robot equipment. That’s fewer than the number of workers employed in a single factory in many companies. Today, there is a bigger market for toy robots than for real robots.

Robots are being produced in the United States at the rate of about 1,500 per year. Predictions are that this will probably grow to between 20,000 and 60,000 robots per year by the year 1990. At that rate, the United States will be lucky to have a million robots in operation before the year 2000. Thus, unless there is some drastic change in the presently projected trends, there won’t be enough robots in operation to have a significant impact on overall productivity before the turn of the century.
Check the blank beside the main idea of the whole passage.

Real robots are much more primitive than the human-like robots in the movie "Star Wars."

Industrial robots are programmed to perform a wide variety of materials-handling tasks.

Robots are not yet advanced enough to have a significant effect on industrial productivity.

Many people believe that industrial robots are already replacing many human workers in the workplace.
Scientists have long known that for most people, the left side of the brain contains important areas for certain language functions - word choice, grammar, and articulation. They have discovered the location of these areas by studying the language disorders of people who are injured in one side of the brain. Damage to the left hemisphere can produce "aphasias," which are disorders of language involving difficulties in speech fluency, word choice, and comprehension. By contrast, the contributions of the right or "minor" hemisphere to language functions have usually been considered rudimentary at best.

In 1979, however, Elliott Ross, a neurologist at the University of Texas Southwestern Medical School, and Harvard Medical School neurologist Marek-Marsel Mesulam began to study the language disorders of people who sustained brain damage in the right hemisphere. Based on this research, they proposed that the right hemisphere had a dominant role in affecting prosody (the coloring, melody, and rhythm of speech), as well as emotional gesturing. In a paper to be published this year, Ross coined the term aprosodia to refer to disorders in the emotional aspect of language that follow damage to the right hemisphere.

There are several kinds of aprosodias, depending upon location and size of damage to the right hemisphere. Aprosodias are being classified into subsets, according to the patient's disabilities. In some cases, the patient speaks flatly and monotonously, without any emotion, coloring, or gesturing, even though feeling emotion. In others, the patient speaks normally, but cannot comprehend the emotional content (happiness or anger) of other people's speech.

Ross is also finding that the aprosodias correspond anatomically to aphasias. When a patient has an aprosodia in the right side of the brain, he or she will also have an aphasia in the left side of the brain.
Check the blank beside the main idea of the whole passage.

People suffering from an aprosodia may not be able to express emotion in their speech even though they are feeling some emotion at the time.

The location of areas in the brain that affect language functions are discovered by studying people with specific language disorders.

There are several types of aprosodias which are classified by the location and size of the damage to the right hemisphere of the brain.

The right hemisphere of the brain makes an important contribution to language functioning by affecting the emotional aspects of speech.
The U.S. population will increase to 309 million in the year 2050, then start to decrease, according to a new U.S. Census Bureau projection. The projection received national publicity because it is the first time that an official zero-population-growth prediction has been given for the United States. But no sooner was the projection publicized than it received a great deal of criticism from major population organizations. These organizations say that faulty assumptions make the projection lower than it should be. They worry that the projection could cause Congress to make inappropriate decisions about population policy and immigration laws. Such laws could then cause serious problems in the areas of food production and distribution, employment, and environmental quality.

The population groups point out that the Census Bureau's projection assumes an annual immigration rate equal to the current legal levels - about 450,000 per year. But even the Census Bureau's own estimates put the illegal immigration rate at 500,000 to one million new people per year. The groups say that by ignoring this major factor, the projection could be off by millions. When The Environmental Fund, a group in Washington, D.C., added illegal immigrants to the Census Bureau's assumptions, their projections put the U.S. population at between 341 and 374 million and still rising by 2030. This number is well above the bureau's 309 million figure for 2050.

The bureau's projection also assumes that the current fertility rate - the lowest in history - will continue. The Environmental Fund says that this assumption is also invalid, since the percentage of black and Hispanic women - historically more fertile than whites - is still rising. Hispanics make up a large part of the illegal immigrants; various studies estimate that they are 50-70% of the illegal total. A report by the Population Reference Bureau projects that with an annual immigration level of one million, the number of Hispanics will double by the year 2000. This could have a large effect on the overall national fertility rate. An increase of only a few tenths of a percent in the fertility rate could result in population increases of many millions of people.

Wide margins of error are common in long-term population projections. "If projections today are no better or worse than past projections for only a 20-year period," says Carole Baker, executive director of Zero Population Growth, "we are looking at strong odds that the numbers could be off by as much as 100 million people, higher or lower than the stated projections."
Population organizations feel that the U.S. Census Bureau's projection could have serious negative effects.

The U.S. population will reach a total of 309 million in the year 2050 and then start to decrease.

Wide margins of error, though undesirable, are common in long-term population projections.

Hispanics make up a large part (50-70%) of the illegal immigrant population entering the U.S.
Weather forecasts covering the whole world are now accurate enough that multimillion-dollar business decisions can be safely based on them. In fact, the members of the European Centre for Medium-Range Weather Forecasts (ECMWF) make a lot of money by providing long-range forecasts to a wide variety of business people. Because they are able to plan for up to a week ahead, these business people can save millions of dollars a year.

Today's success in forecasting comes from applied rather than pure science. The basics of weather have been known for centuries - "energy in and energy out," is how Gilles Sommeria, a French scientist at ECMWF, summarizes it. The fundamental process consists of warming by the sun's short-wave radiation and cooling by long-wave radiation going back into space. The sun heats the Earth's air in different parts of the globe to different temperatures, setting up variable winds, pressures, and rates of evaporation. The Earth's rotation then makes the warm and cold air masses move in circles away from each other.

The combination of the Earth's rotation, its irregular surface, and differences in surface temperature as the air passes over, forces the air currents into their actual complex, irregular paths. Weather patterns are thus in continuous motion, influenced by certain physical laws.

Although this phenomenon had been observed and measured for centuries, the forecasting of weather patterns was informed guesswork until very recent times. In the late 1940s the eminent mathematician John Von Neumann, of the Institute for Advanced Study in Princeton, New Jersey, recognized that weather prediction could be seen as a set of mathematical equations that had to be solved with the use of computers. (Actually, Von Neumann was putting into practice ideas first expressed in the early 1920s by the British mathematician L.F. Richardson.) In Von Neumann's view, the observed data could be expressed in mathematical terms, the physical laws translated into formulas, and the weather forecast calculated like an arithmetic problem. Of course, literally billions of calculations would be needed, far more than mere humans could handle. But Von Neumann, a computer pioneer, saw that the job could eventually be done by advanced computers. Thus was born the age of numerical weather forecasting.
Check the blank beside the main idea of the whole passage.

- Until very recent times, the method for predicting the weather could not produce long-range forecasts.

- Long-range weather forecasts have helped many businesses in Europe to save millions of dollars a year.

- Changes in the weather are caused by the exchange of heat and cold between the Earth and outer space.

- Accurate weather forecasting using mathematical equations was made possible with the use of computers.
Changes in the Structure of the American Family

- At current rates, half of all American marriages begun in the early 1980s will end in divorce.
- The number of unmarried couples living together has more than tripled since 1970.
- One out of four children is not living with both parents.

The list could go on and on. By now, predictions of the demise of the American family are familiar to everyone. Yet the family is a resilient institution that is stronger today than its worst critics believe. There is, for example, no evidence that Americans are turning away from marriage. It is true that many young adults are living together outside of marriage, but the evidence we have about this practice suggests that it is not a life-long alternative to marriage; rather, it appears to be either another stage in the process of courtship and marriage or a transition between first and second marriages.

Young adults today do marry at a somewhat older age, on average, than their parents did. But the average age at marriage today is very similar to what it was throughout the period from 1890 to 1940. To be sure, many of these marriages will end in divorce, but three out of four people who divorce eventually remarry. Americans still seem to desire the intimacy and security that a marriage provides.

Much of the alarm about the family comes from reactions to the sheer speed at which the institution changed in the last two decades. Between the early 1960s and the mid-1970s, the divorce rate doubled, the marriage rate plunged, the birthrate dropped from a twentieth-century high to an all-time low, premarital sex became accepted, and married women poured into the labor force. But since the mid-1970s, the pace of change has slowed. The divorce rate has risen modestly and the birthrate even has increased a bit. We may have entered a period in which American families can adjust to the sharp changes that occurred in the 1960s and early 1970s. We think that, in general, accommodations will be made as expectations change and institutions are adjusted to changing family practices.
The institution of the American family has changed drastically in the past twenty years.

Despite the alarming changes of the last twenty years, the institution of the American family is surviving.

Americans still seem to want the intimacy and security that comes from marriage.

Many people predict that the American family will not survive the changes of the last two decades.
Computer users must have a firm idea of what they expect the computer to do before they can use the computer effectively. They need to "know what they know" and "what they don't know." In other words, in any given problem, certain information to be processed is available - mathematical formulas and data from experiments, a list of student names and grades, an airline schedule, and police records, to name some examples. Computer users want the computer to modify the input information to obtain some output. They must decide what that output is to be. They must have a good idea of how the computer is to serve in the processing of the information.

They then must decide whether or not the use of the computer to solve the problem is worth the effort and expense involved. A personal experience of mine may serve to illustrate this point. I came upon a contest for which the prize was a trip to Hawaii. The rules of the contest were simple: the entrant who submitted the longest list of 4-letter words in the English language to be constructed from a given set of 14 letters would win first prize. A quick calculation shows that there are 14 or 38,416 possible 4-letter combinations of the 14 given letters. I then wrote and ran a program whose input was the 14 given letters and output was all the possible combinations of the letters. This produced 2744 lines of printed output, each line consisting of 14 4-letter combinations.

What now? It was easy to eliminate "words" such as CCCCC. But how about some more questionable ones such as PITH (or is it PYTH?). Do slang or vulgar words count? The next step was to sit down with the dictionary and go through the list of words, deciding which ones were real words and which ones were not. I belatedly realized that this is exactly what I would have had to do even if I had not used the computer to generate the list of words; all the computer did was to supply a list of letters. Had a list of all valid 4-letter words been available for use by the computer, the computer could have completed the entire problem. However, this was not the case.

In other words, the example described above is an example of failure on my part to analyze the problem involved from start to finish. Had I done so, I would have realized that the cost in my and the computer's time would be really wasted. Millions of dollars and thousands of work hours are wasted every year for exactly the same reason. Computer users must determine if the use of the computer could help them and, if so, whether or not the benefits they could derive from its use justify the time and expense involved.
Check the blank beside the main idea of the whole passage.

- Computers can process a broad range of input data such as mathematical formulas, lists of student names and grades, and airline schedules.

- Computers can be programmed to perform a wide variety of tasks including generating all possible combinations from a list of fourteen letters.

- Before using the computer to solve a problem, the users must decide if the computer could be helpful and is worth the time and money involved.

- Computer users must often do a great amount of the work themselves, even after a computer program is used to help solve a difficult problem.
The Effects of Exercise on Mental Health

Regular exercise can have positive effects for the body, but is it good for the mind as well? Is a physically fit person, for instance, better able to deal with stress at work and at home? Will a program of supervised running and other exercises help to ease the mild depression that might send some people to a psychotherapist?

Following the recent "fitness boom," some physicians and psychologists have proposed that the answer to these questions is a definite "yes," but researchers are just beginning to understand the effects of exercise on mental life.

Despite conflicting results of a number of studies thus far, there is evidence that physically fit individuals have an advantage in dealing with stressful real-life events. An example is a report by University of Kansas at Lawrence psychologists David L. Roth and David S. Holmes. They found that when faced with a high percentage of life changes such as divorce, death of a loved one and changing jobs in the previous year, physically fit subjects reported fewer health problems and symptoms of depression than subjects who were less fit. However, in another study Roth and Holmes observed that once stressful changes have already occurred, supervised exercise is not particularly effective at reducing physical illness. Physical fitness, says Holmes, can help prevent the effects of stress from life changes. It is not as helpful in treating the effects of stress.

Yet the findings are far from conclusive, cautions psychologist David Sinyor of Concordia University in Montreal. Several experiments have found that heart rates of physically fit persons - in most cases, runners - do not differ from those of non-exercisers during stressful laboratory tasks. However, the exercisers return to their resting heart rates more quickly after completing a task. A few other studies have found that exercisers' heart rates are slower while they perform similar laboratory tasks. In addition, Sinyor says, people who really like to exercise a lot may do so because of personality characteristics that already protect them against stress. This factor is not explained in earlier studies. Finally, learning any skill, from rugweaving to relaxation techniques, may increase self-confidence and coping abilities in the same way as a regular exercise program.
Some doctors believe that exercise does have positive effects on mental health, but others say the results of research studies are not yet that conclusive.

Physically fit individuals report fewer health problems when dealing with stressful life changes such as divorce, death of a loved one, or switching jobs.

Learning any new skill may increase an individual's self-confidence and reduce the effects of stress in the same way that a regular physical exercise program does.

Physical fitness helps to prevent the effects of stress from life changes, but it does not reduce the effects if the stressful event has already occurred.
How Small Group Discussions Affect Opinion

We live our lives within small groups. From infancy through old age we move in friendly groups whose attitudes and interests we generally share. Their influence is powerful. Several years ago Yale University psychologist William McGuire concluded that the mass media have less of an effect on opinion than informal conversations with family, friends, neighbors, and co-workers do.

Recently, the inclination toward forming small groups has become less casual. The act of people coming together to discuss issues and solve problems seems increasingly integral to our existence. Alcoholics, dieters, and educators work in small groups to help each other, and individual psychotherapy moves more and more in the direction of various group therapies. Business people are beginning to make decisions in groups rather than give one person that responsibility.

It is no secret that people associate with others whose attitudes and values are similar to their own. Most of us need only look at our circle of friends to illustrate this point. When people come together in a group, whether it be to combat a drinking problem, to lose weight, or to make a decision regarding management policy, members are likely to carry certain shared inclinations. In the case of losing weight, they have a mutual desire to stop their excessive eating.

When group members share common attitudes, does discussion do anything more than converge their opinions? Social psychologists have found that it does. As members of diet groups discuss their mutual problem, their shared desire to lose weight may strengthen. In community disagreements, according to sociologist James Coleman, people who share the same opinions associate with one another as a conflict develops; this makes their shared opinions stronger. The President's Commission on Campus Unrest noted the same process at work during the evolution of the radical student movement in the 1960s. Similarly, investigators of gang delinquency have observed a process of growing agreement within neighborhood gangs whose members have a common social, economic, and ethnic background.

Observations of small groups such as these, together with recent experiments in social psychology, have provided new information concerning the effects of talking in small groups: From the results of these experiments, researchers now realize that discussion generally strengthens the average inclination held by group members before the discussion.
Check the blank beside the main idea of the whole passage.

- Group therapy can help alcoholics to stop drinking and dieters to lose weight more effectively.

- Small group discussions tend to increase the strength of individual group members' opinions.

- Management policy in business is increasingly being decided by small groups rather than one person.

- Informal conversations have a stronger effect on people's opinions than the mass media do.
Animals as a Source of Power in Farming

Approximately 400 million horses, cattle, oxen, donkeys, mules, camels, llamas, water buffaloes, and elephants now work for humanity. They fulfill about half of Third World agriculture's energy needs overall; in some countries their contribution approaches 90%. Yet, Third World governments and scientists overlook the significance of draft animals. Noel Vietmeyer of the National Academy of Sciences says that this is partly because "many educated people feel embarrassed at having animals rather than tractors plowing the fields."

Development planners and scientists who look down on animal power hope to replace the ox with the tractor. Vietmeyer, however, believes that farm machines will be used less in the future. The cheap and abundant petroleum and technological infrastructure needed to support mechanized agriculture are not now present in most developing countries, nor will they materialize soon. In India, for example, bullocks pull more freight than railroads do. Draft animals provide two-thirds of India's rural transport.

"Draft power has a particular place on small farms," says Vietmeyer. "Farm machines such as tractors and tillers become uneconomical when the farm is smaller than four hectares. In the Third World there are over 100 million such farms. Furthermore, in hilly terrain, narrow or waterlogged fields and other special situations, draft animals are often the only practical source of power."

"Using work animals is not backward and it is not a transient phase that will soon pass. For each country there's a special combination of animal power and mechanical power. The balance may be on one side or the other, depending on the average size of the farm, length of haulage, cost and availability of petroleum, and ability to maintain machines. Animals and machines are both legitimate power sources. One is not inferior to the other."

Draft animals, Vietmeyer concludes, played a major role in the development of Europe, North America, Australia, and South Africa, and they could play a similarly large part in the development of the Third World. The main barrier to their more effective use is the disdain of scientists, engineers, and officials. As Indian scientist and animal-power advocate N.S. Ramaswamy observes, "India has put a satellite in space and harnessed the atom, but our carts are 5,000 years old because professors are scared they may not get promoted if they work on designing better ones."
Draft animals played a major part in the development of Europe, North America, Australia, and South Africa, and they could do the same in developing countries.

Draft animal power is particularly useful on small farms of less than four hectares or where special conditions make mechanized power impractical.

Many developing countries have no choice but to use animal power because they don't have the petroleum resources or infrastructure to support mechanized power.

Work animals are an important source of draft power for developing countries but are disdained by officials and scientists who want farmers to use machines.
Early History of Chaco Canyon

New Mexico's Chaco Canyon has long fascinated and puzzled archaeologists. The Anasazi Indian civilization lived and grew there from about A.D. 1000 to A.D. 1200; then, mysteriously, they disappeared. They are believed to have been the ancestors of modern-day Hopi and other Pueblo Indian groups. They also built the largest and tallest buildings that existed in North America before the construction of skyscrapers at the end of the 19th century. Hundreds of 12th-century Chaco Canyon families lived in huge stone-and-adobe apartment buildings that stood as high as five stories and whose ruins are still impressive today.

Archaeologists consider Chaco Canyon culture to have been almost as advanced as that of the Maya in Central America, but the people of Chaco Canyon were unable to sustain the life-giving qualities of their environment. Archaeologist Julio Betancourt has discovered clues as to what happened to them, thanks to the pack rats which lived in Chaco Canyon during the time of the Anasazi Indians.

Betancourt said that throughout the Anasazi region, pack rats established nests and left garbage heaps, called middens. Scraps of seeds and small pieces of plants were left by the rats and preserved in the middens. In fact, as many as 30 different plant species are represented by preserved seeds and twigs in each midden. The age of the material is determined by analysis for radioactive carbon 14.

"This gives us a series of snapshots of what plant species were growing in the Chaco Canyon area at various times during the history of the region," Betancourt said. From this kind of evidence, he said, "you can see sudden deforestation in the Chaco Canyon. Before 1,000 years ago there was pinyon-juniper woodland in the Chaco Canyon, and for the past 1,000 years there was not. The change was irreversible."

Although the cause of the deforestation is debated by scientists, Betancourt believes that as the Chaco Canyon builders used up local woodlands for fuel and building, erosion destroyed the top soil and the irrigation system. Thus, the Anasazi Indians directly contributed to their own demise by abusing the lands on which they lived.
The Anasazi Indians were the ancestors of modern-day Hopi Indians and built the largest and tallest buildings of their time.

Archaeologists can determine what plants were alive in Chaco Canyon by analyzing the preserved seeds and plants left by rats.

The disappearance of the Anasazi Indian culture could be explained in part by their abuse of their environment.

The Anasazi Indian culture mysteriously disappeared from Chaco Canyon in New Mexico after about A.D. 1200.