AN ASSESSMENT OF THE GEOGRAPHIC KNOWLEDGE AND UNDERSTANDINGS OF FIFTH GRADE STUDENTS IN MICHIGAN

By

Norman C. Bettis

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Geography

1974

ABSTRACT

AN ASSESSMENT OF THE GEOGRAPHIC KNOWLEDGE AND UNDERSTANDINGS OF FIFTH GRADE STUDENTS IN MICHIGAN

By

Norman C. Bettis

The major purpose of this study was to determine the geographic knowledge and understandings of fifth grade students in Michigan. Several tasks were carried out to achieve this purpose. First, significant geographic content and cognitive behaviors were identified. Geographic content areas assessed included physical geography, human geography, and geographic skills. Cognitive behaviors assessed included remembering facts and understanding of concepts and principles. Second, a standardized assessment instrument was constructed. The test consists of forty-nine multiple-choice items focused on one of nine maps or other figures. Third, a proportional stratified 1 percent sample of Michigan's fifth grade population completed the test (N=1689). The sample represented twenty school districts from three urban regions of lower Michigan and was selected proportionally across five types of school districts and across student achievement levels.

This study concluded that:

1. The Michigan Intermediate Geography Test (MIGT) functioned adequately as an assessment instrument.

2. Raw score distributions indicate that understandings of the geographic facts, concepts, and principles on MIGT are not well developed among Michigan's fifth grade population.

3. Students' knowledge and understandings are best developed in human geography, less developed in geographic skills and least developed in physical geography.

4. Michigan's fifth graders are more adept at answering questions depending on recall of specific information than questions requiring understanding of concepts and principles.

5. Achievement in geographic learning is highest among urban fringe students, less well developed among students in the city, town, and rural districts, and least well developed in the metropolitan core students.

It was recommended that additional research be conducted to ascertain what students at other grade levels know about geography and which factors contribute to geographic learning. Greater emphasis should be placed on pre-service and in-service training of teachers of geography; curriculum materials suitable for teaching and learning geography in the elementary school should be developed jointly by geographers and educators.

To Sandi and the future

ACKNOWLEDGMENTS

Appreciation is expressed to the members of my guidance committee for their assistance and support in the pursuit of my educational goals. The members of the committee were Dr. Gary A. Manson, chairman, Dr. Daniel Jacobson, Dr. Stanley Brunn, Dr. Harold A. Winters, and Dr. Stanley P. Wronski, Professor of Education. A special degree of indebtedness is owed Dr. Gary A. Manson for his professional advice and assistance with the manuscript and with other aspects of my program. His personal counsel and friendship have been invaluable.

Appreciation is expressed to the personnel of the Computer Institute for Social Science Research, the Data Processing Department, and the Office of Evaluation Services at Michigan State University. Thanks is extended also to the educators who served as judges and in whose classrooms the testing was carried out.

Finally, to my family, the prime motivating force, I express my deepest appreciation: to Sandi, for her patience and understanding, for her endless labor throughout the years of my professional training and in the typing of the manuscript; and to Lori, Michele, and Todd for their tolerance.

TABLE OF CONTENTS

APPENDICES

Page

Appendix

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Many geographic educators believe that the geography taught in elementary schools, high schools, and colleges should reflect accurately and adequately the status and progress of the discipline. This belief has been expressed in geographic education literature for decades. For instance, in 1902 Davis, discussing geography's representation in colleges, stated:

> There can be no question that the disciplinary side of geography deserves more emphasis than it has usually received in school teaching. There is every reason to hope that, commensurate with the development of a body of higher learning in geography, there will be an increase of the disciplinary value of school geography.[1]

Decades later Scarfe also emphasized the importance of school geography curriculum reflecting "ideas, concepts and attitudes" from the discipline. Discussing the teaching of geography and teacher preparation, he said that

---

[1]William Morris Davis, "The Progress of Geography in the Schools," First Year Book National Society for the Scientific Study of Education (1902), Part II, 7-49, in Geographical Essays, ed. by Douglas Wilson Johnson (New York: Dover Publications, Inc., 1954), p. 63.

> . . . knowledge of geography grows from small immature ideas and concepts to larger and more complex ideas and eventually to a fully rounded philosophy of the subject. It is these embryo ideas and concepts that are particularly significant criteria in the lower school, but it is very difficult, if not impossible, for anyone who has not achieved a fully rounded philosophy of the subject to pick out and select the embryonic ideas that eventually grow to geographic expertness.[2]

In a later paper he expresses this viewpoint again:

> . . . school geography should be concerned with the essence of the subject only. If it is to justify its place in the curriculum it should promote ways of thinking that are distinctly geographical.[3]

Advocating transfer of understandings from the discipline to the school is not confined to higher levels of education. Whipple adopted the same philosophical attitude when she discussed a proper role for geography in the elementary social studies program.

> It is imperative that recent geographic findings . . . be interpreted for elementary curriculums with a view to clarifying the essential concepts and eliminating those that are outmoded. Important concepts and skills should be initiated in the first grade and applied, and extended from one grade to the next higher, all the way through the secondary

---

[2]Neville V.Scarfe, "Designing the Curriculum to Develop Geographic Concepts," The Journal of Geography, LII (March, 1953), 100.

[3]Neville V. Scarfe, "Geography as an Autonomous Discipline in the School Curriculum," The Journal of Geography, LXIII (October, 1964), 297.

> school. First grade children are capable of geographic reasoning if the subject matter is closely related to their past experiences.[4]

Similar views about the transfer of knowledge from the discipline to school curriculum have been expressed by Kohn, Thralls, James, and Gabler, among others.[5]

The High School Geography Project (HSGP) exemplifies a curriculum which includes fundamental structural components of a discipline. According to Patton:

> The materials would try to teach some of the ways the discipline looks at the world, some of the kinds of questions geographers ask about the world, and some of the methods geographers employ to answer the questions they ask.[6]

---

[4]Gertrude Whipple, "Geography in the Elementary Social Studies Program: Concepts, Generalizations, and Skills to Be Developed," in New Viewpoints in Geography, Twenty-ninth Yearbook of the National Council for the Social Studies, ed. by Preston E. James (Washington, D.C.: National Council for the Social Studies, 1959), pp. 112-13.

[5]See Clyde F. Kohn, "Foreword," in Geographic Approaches to Social Education, Nineteenth Yearbook of the National Council for the Social Studies, ed. by Clyde F. Kohn (Washington, D.C.: National Council for the Social Studies, 1948, p. vii; Zoa A. Thralls, The Teaching of Geography (New York: Appleton-Century-Crofts, Inc., 1958), p. 16; Preston E. James, "The Hard Core of Geography," in New Viewpoints in Geography, Twenty-ninth Yearbook of the National Council for the Social Studies, ed. by Preston E. James (Washington, D.C.: National Council for the Social Studies, 1959), pp. 7-8; and Robert E. Gabler, "Introduction: History, Scope, and Purpose," in A Handbook for Geography Teachers, ed. by Robert E. Gabler (Normal, Ill.: National Council for Geographic Education, 1966), p. 3.

[6]Donald J. Patton, "The Product," in From Geographic Discipline to Inquiring Student: Final Report on the High School Geography Project, ed. by Donald J. Patton (Washington, D.C.: Association of American Geographers, 1970), p. 3.

Pratt has characterized HSGP as "the course [that] allows the student to experience the same type of problems faced by professional geographers,"[7] and cites several examples of students assuming that role.

Geographic educators are not the only curriculum specialists who would have subject disciplines well represented in the school. The Anthropology Curriculum Study Project: Patterns in Human History has been developed from the rationale that knowledge and better understanding of human behavior should not be restricted to the discipline, but should be known to all. Charles goes on to point out that anthropology is the core discipline represented in these materials.[8] The Earth Science Curriculum Project (ESCP), a program for ninth grade students, draws on several of the sciences. Its developers "sought the help of scientists in many fields to make sure that the basic principles in all these fields [have] formed an integrated and up-to-date story of planet earth and its environment in space."[9]

---

[7]Robert B. Pratt, "High School Geography Project: Geography in an Urban Age," Social Education (special issue), XXXVI, No. 7 (November, 1972), 751.

[8]Cheryl L. Charles, "Anthropological Curriculum Study Project: Patterns in Human History," Social Education (special issue), XXXVI, No. 7 (November, 1972), 725.

[9]Ramon E. Bisque and Robert L. Heller, Investigating the Earth: Earth Science Curriculum Project (Boston: Houghton Mifflin Co., 1968), p. vii.

The Sociological Resources for the Social Studies (SRSS) is a "project developed out of the concern of professional sociologists for the status of their discipline in the new social studies movement."[10]

Some social science educators feel that a well-designed curriculum should reflect, not only the latest facts, concepts and principles from the disciplines, but should also evidence input from those concerned with the methods and processes by which children learn. As Wesley and Wronski point out, in the past neither the social scientists nor the educators working alone have been successful in designing a suitable social studies curriculum. The social scientist was concerned exclusively with the content in the curriculum while the educator placed too much emphasis on the learning process to the neglect of content.[11]

The High School Geography Project is an example of geography curriculum that has been constructed with both the discipline and the learner in mind.[12] Not only has it gone

---

[10]Frances Haley, "Sociological Resources for the Social Studies," Social Education (special issue), XXXVI, No. 7 (November, 1972), 765. For a review of other efforts to infuse the social sciences into the pre-collegiate curriculum, see Social Education (special issue), XXXVI, No. 7 (November, 1972).

[11]Edgar B. Wesley and Stanley P. Wronski, Teaching Social Studies in High Schools (Boston: D. C. Heath and Company, 1964), pp. 32-33.

[12]"People Taking Part in the High School Geography Project," in From Geographic Discipline to Inquiring Student: Final Report on the High School Geography Project, ed. by Donald J. Patton (Washington, D.C.: Association of American Geographers, 1970), pp. 73-79.

far in identifying "unifying strands and themes" from the discipline, but it has also utilized an inquiry teaching strategy which seems more suitable for student learning.

The "Gap Hypothesis"

Advocating the reconciliation of the discipline with what is being taught in the schools suggests that a "gap" exists and presumes that much remains to be accomplished in the effective dissemination of geographic knowledge to the public schools. Geographic educators have assumed, in a priori fashion, that the "gap" between what is known by school students and what is occurring within the discipline is great. Scarfe states categorically:

> In Europe, geography is generally well taught and universally accepted as a necessary subject in school but in America this is not the case.[13]

Kennamer makes the same assumption in discussing the implications for geography in the emerging social studies curriculums of the 1960's.

> School geography must change on two counts if it is to remain viable in the schools. First, the changes in the academic discipline itself are quite sweeping and the generation gap is vast. Geography today, which includes the study of culture, spatial distribution, spatial interaction, and functional regions, is quite different from the geography of

---

[13]Neville Scarfe, "Our Educational Environment and Geography Teaching," The Journal of Geography, LIX (March, 1960), 108.

> yesterday which emphasized the physical environment and the descriptive study of world regions.[14]

Commenting on the High School Geography Project, Patton assumes a discrepancy between discipline and school geography and says, "The course seeks to 'narrow the gap' between the discipline as it is taught in the high school classroom and the frontiers of current research and professional thinking in geography."[15] One of the strongest assertions of the "gap" hypothesis comes from Gould, who states flatly:

> In a country where many planning decisions depend on people's ability to judge and weigh various alternatives, and in an age where knowledge and understanding of our planetary home is crucial we are still turning out geographical illiterates.[16]

It may well be that the geographic knowledge held by students is less than adequate, but none of the individuals cited offers evidence to support his assertion. It appears that most know where they wish to go and assume that they know where they are. Some researchers have assessed the status of geographic knowledge, but the picture is far from

---

[14]Lorrin Kennamer, Jr., "Emerging Social Studies Curricula: Implications for Geography," in Focus on Geography: Key Concepts and Teaching Strategies, Fortieth Yearbook of the National Council for the Social Studies, ed. by Phillip Bacon (Washington, D.C.: National Council for the Social Studies, 1970), p. 391. (Emphasis mine.)

[15]Patton, "The Product," p. 3.

[16]Peter R. Gould, "The New Geography," Harper's Magazine, March, 1969, p. 100.

clear.[17] The paucity of data is especially noteworthy at the elementary level. This writer is not aware of any recent systematic attempt to assess geographic understandings of elementary students. It would seem that such data should be gathered and utilized by geographic educators, curriculum planners, and local school district officials who are interested in curriculum revision, as well as those interested in "closing the gap."

The need to determine what is actually known by students is not confined to geographic research. On a much broader scale, the National Assessment of Educational Progress is currently undertaking "a national census-like survey of the knowledge, skills, understandings, and attitudes of certain groups of young Americans."[18] The two major goals of the project are:

---

[17]See, for example, William John Hromyk, "An Evaluation of High School Senior's Knowledge of Anglo-American Geographic Concepts" (unpublished Ph.D. dissertation, University of Oregon, 1972); Randall Arthur Pelow, "Predicting Aptitude for Modern Geographic Concepts in Sixth Grade Pupils" (unpublished Ed.D. dissertation, The Pennsylvania State University, 1971); Bruce E. Adams, "Geographic Education in the Public and Parochial Schools of a Four County Sampling of Pennsylvania" (unpublished Ed.D. dissertation, The Pennsylvania State University, 1960); Maurice M. Zacur, "A Measure and Analysis of Geography Achievement in Twenty Junior-Senior High Schools of Pennsylvania" (unpublished Ed.D. dissertation, The Pennsylvania State University, 1970); and Francis Ann Slater, "An Inquiry Into the Kind of Geography Being Learned in Grades 3-12" (unpublished Ph.D. dissertation, University of Iowa, 1969).

[18]"Questions and Answers About National Assessment of Educational Progress," National Assessment of Educational Progress (Washington, D.C.: Government Printing Office, 1970), p. 1.

1. to make available the first census-like data on the educational attainments of young Americans, [and]
2. to measure any growth or decline which takes place in selected aspects of the educational attainments of young Americans in certain subject areas.[19]

The National Assessment study is deemed necessary because of the billions of dollars spent annually on education in the United States and the desire to know "how much good is the expenditure of so much money doing, in terms of what Americans know and can do."[20]

The determination of geographic learnings by National Assessment is included as part of the social studies assessment and does not receive the detailed treatment proposed in this study. For example, only one "geographic objective" for elementary students has been included among the many social studies objectives formulated. It reads:

> [The student should] know something of the location of major cities, states, nations, continents, and oceans; [he should] understand some of the influences the physical environment has on the way people live.[21]

One other objective indicates that students should be able to "obtain information from a variety of sources"[22] and

---

[19] Ibid. [20] Ibid.

[21] "Social Studies Objectives," National Assessment of Educational Progress (Washington, D.C.: Government Printing Office, 1970), p. 17.

[22] Ibid., p. 10.

lists, among other sources, the simple map, globe, and picture. These objectives are hardly representative of the structure of geography today and would not seem to provide an adequate base for the construction of valid test items.

In addition to National Assessment, several states have undertaken state-wide assessment programs which parallel closely the goals and organization of National Assessment. Michigan's assessment program in social studies is currently in the objective construction stage. Examination of these preliminary objectives indicates a greater emphasis on assessment of geographic understandings than those of National Assessment.[23] To illustrate, intermediate-level objectives include the topic, man's use and misuse of his natural environment. An example of an objective for this topic is:

> Given an illustration or a verbal description of man's use or manipulation of the natural environment, the student will identify possible consequences, both positive and negative, of that use or manipulation, as measured by minimum criteria on an objective referenced test.[24]

The following objectives have been proposed for map and globe reading skills:

---

[23]See, for example, "Social Studies Performance Objectives: Elementary Level, Working Draft" (Lansing, Mich.: Michigan Department of Education, 1973). (Mimeographed.)

[24]Ibid., p. 94.

1. Given a map and/or a globe, the student will interpret it by use of the key, as measured by minimum criteria on an objective-referenced test. [And]
2. Given a map and/or a globe the student will make observations and judgments as to how the geographic conditions of a certain area might affect the lives of the people who live there, as measured by minimum criteria on an objective-referenced test.[25]

These objectives seem to reflect more adequately some current geographic understandings than those of National Assessment.

The goals of the assessment programs and this study are similar, but this research will provide geographic educators with more specific information about geographic learnings in Michigan and perhaps serve as an example of how individual social science disciplines can supplement the work of national and state assessment programs to determine more exactly the state of school students' understandings in their particular disciplines.

Statement and Justification of the Problem

A first major need in the development of school curriculum reflecting the discipline is knowledge of just what geographic facts, concepts, and principles are known by children. It is necessary that this question be answered to determine if indeed a gap does exist between school geography and the academic discipline and, if so,

---

[25] Ibid., pp. 96-97.

the extent of such differences prior to any attempts to "close the gap." A major deterrent of attempts to determine the existence and extent of a "gap" is the paucity of achievement tests meeting established criteria. Wood emphasizes this point in discussing the geography test review entries in Buros' Sixth Mental Measurement Yearbook. She states, "Only about 10 percent of the authors of the geography achievement tests and the geography sections of social studies tests are members of the Association of American Geographers or the National Council for Geographic Education."[26] This observation clearly indicates that geographic assessment in the past has been carried out largely by non-geographers.

Most achievement tests carrying the geography label are outdated, inaccurate, or lacking in adequate supporting data. An examination of Buros' Tests in Print reveals that only three of the twelve geography tests listed have been revised as recently as 1961.[27] Except for the Geography Achievement Test for Beginning High School Students, sponsored by the National Council for Geographic Education and revised in 1968, and the Graduate Record Exam Advanced

---

[26]Susan Wood, "An Evaluation of Achievement Tests in Geography," in Evaluation in Geographic Education, The 1971 Yearbook of the National Council for Geographic Education, ed. by Dana G. Kurfman (Belmont, Calif.: Fearon Publishers, 1970), p. 134.

[27]Oscar K. Buros, ed., Tests in Print (Highland Park, N.J.: Gryphon Press, 1961), pp. 239-40.

Geography Test (GRE), only two of the tests reviewed in Buros' Sixth Mental Measurement Yearbook and Seventh Mental Measurement Yearbook are as current as 1964.[28]

Several social studies tests include questions to test geographic understandings. Unfortunately, some of these tests contain elementary errors in scale or political boundaries. Wood cites an example in discussing a "geography" test distributed by a major test publishing house.

> In drawing the political boundary of Alaska, the "staff artist" has committed what geographers call "cartographic aggression." This was accomplished by the apparent awarding of the territory at the mouth of the MacKenzie River to the United States and likewise awarding the complete panhandle and other southeastern Alaskan territory to Canada.[29]

Perhaps the tests that do the best job of measuring geographic knowledge are the achievement test survey batteries such as the Sequential Tests of Educational Progress (STEP) and the Iowa Test of Basic Skills (ITBS). These tests usually utilize subject matter specialists as well as evaluation specialists, but often lack breadth and depth in geography.[30]

---

[28]Oscar K. Buros, ed., The Sixth Mental Measurement Yearbook (Highland Park, N.J.: Gryphon Press, 1965), pp. 1231-32; and Oscar K. Buros, ed., The Seventh Mental Measurement Yearbook (Highland Park, N.J.: Gryphon Press, 1972), pp. 1308-10.

[29]Wood, "An Evaluation of Achievement Tests," pp. 134-35.

[30]Ibid., pp. 141-42.

To this writer's knowledge no current standardized geography achievement instruments exist to measure the concepts and principles currently advocated for school geography. As Wood points out:

> Many tests have been produced and marketed, after thoughtful deliberation, by geographers with excellent credentials and good intentions. But because these tests lack standardization their usefulness is limited and their value cannot be measured objectively.[31]

Wood goes on to comment on the status of achievement testing in geography and to cite some of the needs:

> Comments thus far have accented the negative aspects of geography achievement testing because, until recently, little that is noteworthy has been done in the field of geography testing. Trained geographers are needed to construct really sound geographic test items, but a geographic background alone is not enough. We have had a history of geographers validating their tests almost solely on the basis of approval by other geographers. This convoluted situation has meant that there are few geography achievement tests able to withstand critical review in Buros' Yearbooks. What is needed to help make the geographer's knowledge useful on a test is the contribution of the evaluation specialist.[32]

It is evident from Wood's comments that geographic educators must acquire expertise outside geography if they are to be effective in measuring acquired student learnings from the discipline. Yet, few geographers appear to have apprised themselves of acceptable methods of test

---

[31] Ibid., p. 131.

[32] Ibid., p. 135.

construction.[33] Geographic education literature on testing contains few references to educational testing and measurement literature. The needs are clearly evident and have been recognized, for example, in publication of results of the Greyston Conference of the National Council for Geographic Education. Participants of this conference identify measurement in geographic education--specifically "preparation of standardized tests to evaluate students' understandings of major portions of the world" and "evaluation of tests to discover those that best measure regional understanding, skill development, etc."--among the most pressing research needs of geographic education.[34]

An immediate need exists for an acceptable standardized measuring instrument focused on accepted geographic concepts and principles. The general lack of such instruments in geography exists not because of an inability to construct them; educational testing and measurement specialists have outlined the procedure in considerable detail.[35] It exists because an insufficient number of

---

[33]Dana Kurfman and Neville Scarfe are notable exceptions to this generalization.

[34]Phillip Bacon and Lorrin Kennamer, Jr., Research Needs in Geographic Education, Geographic Education Series No. 7 (Normal, Ill.: National Council for Geographic Education, 1967), p. 45.

[35]See, for example, Robert L. Ebel, Essentials of Educational Measurement (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972), Part II and Part V; William A. Mehrens and Irvin J. Lehmann, Standardized Tests in Education (New

geographers have examined and made use of such literature in their test construction efforts. It would seem to be in the best tradition of geography for geographic educators to acquire a working knowledge of the principles of educational testing and measurement for the purpose of constructing valid and reliable standardized measuring instruments, thus enabling them to verify or refute their claims about school geography.

If a wide discrepancy does exist between discipline and school geography, the first step which should be taken to close this gap is an effective evaluation of which geographic learnings have been acquired. It seems likely that some current geographic understandings have filtered into the schools, but little descriptive data exist to substantiate this assertion. The most direct and logical way of obtaining such data is through the use of a valid and reliable measuring instrument. Once important geographical concepts and principles have been identified by the discipline, an acceptable standardized achievement test constructed and testing carried out, and a determination made as to the extent and variation of geographic understandings possessed by school students, geographic

---

York: Holt, Rinehart and Winston, Inc., 1969), Chaps. I and III; David Magnusson, Test Theory (Reading, Massa.: Addison-Wesley Publishing Company, 1967); and Robert L. Thorndike, ed., Educational Measurement (Washington, D.C.: American Council on Education, 1971).

educators should be able to function much more effectively in the task of aligning professional and school geography.

The charge for geographic educators seems clear: (1) they must determine what discrepancies between geography in the discipline and geography in the schools exist, and the extent and variations of such discrepancies; (2) why such discrepancies exist; and (3) what actions can be taken to insure that young people will receive geographic understandings of the real world that will prove most useful at present and in future years.

This study is concerned primarily with a determination of the geographic understandings of fifth grade students in Michigan. Geographic understandings are defined from the perspective of the current trends in the discipline. Test items used in the assessment will reflect the discipline also, thus making possible a determination of the discrepancies between discipline and school geography.

Purposes of the Thesis

The purposes of this research are:

1. To state fundamental concepts, principles, information and skills important for elementary school pupils to learn.
2. To construct an elementary multiple-choice geography achievement test utilizing accepted principles and procedures.
3. To standardize the elementary geography achievement test in Michigan.
4. To assess the extent of current geographic understandings among fifth grade students in Michigan.
5. To report, interpret, and analyze the results of the test.

Answers to several questions will be provided by this research:

1. What specific concepts and principles considered important by the discipline have been acquired by students? For example: Can students, when observing a simplified stream drainage system, identify the direction of flow? Can students predict the most likely land use for farmland located on the periphery of an urban metropolitan area or select an acceptable explanation for the daily flow of people from a satellite community into a central city?
2. Are there significant variations in geographic understandings by type of school district? For example, do suburban children know more geography than rural children? Do inner city children perform less well than children in other types of districts?
3. Do students have a workable knowledge of basic geographic skills such as map and graph reading, and interpretation? For example: Can students determine direction on maps? Can students identify the location of the north pole on a polar projection? Can students locate a point on a map by using a number-letter coordinate system? Are students capable of interpreting a climograph by

giving precipitation amounts for a selected season or for the year?

4. Has geographic learning occurred only at the lowest level, or are students capable of comprehension and application of geographic information? For example: Can students use a city street map and select the best location for a service station? Can students select from among four different vegetation-covered landscapes the one that would have the least protection in a heavy rain?
5. Do students possess differential levels of understanding of physical geography, human geography, and geographic skills? For example, is student performance on the human geography items better than on the physical items? Are geographic skills understandings better developed than physical or human understandings?

Answers to these questions will provide a firm base from which geographic educators can operate in constructing curriculum materials which insure that young people will receive those geographic understandings of the real world that will prove most useful at present and in future years.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

Two types of literature are relevant to this research and will be reviewed here. First, an examination of geography achievement tests, geography items on social studies achievement tests, geography-related sections of prominent achievement test batteries, and reviews of geography testings will be undertaken to establish the status of testing and measurement within the discipline. Second, the research findings relative to intermediate-level student knowledge of facts, concepts, and principles from physical and human geography, as well as geographic skills, will be reviewed and set forth.

The Status of Geography Achievement Testing in the Intermediate Grades

The geography achievement tests considered in this review are grouped into three categories--(1) those developed by individual classroom teachers of geography

which have been published in educational journals;[1] (2) those developed for geography which have been published separately and reviewed by competent professional geographers and test specialists; and (3) those standardized test batteries and social studies tests which have been reviewed and contain pertinent geography sections.[2] Only tests appropriate for the intermediate grades 4 through 6 will be examined.

---

[1]For bibliographic references to these tests, see J. R. Gerberich, Specimen Objective Test Items: A Guide to Achievement Test Construction (New York: Longmans Green, 1956), pp. 343-45; John M. Ball, A Bibliography for Geographic Education (Athens: University of Georgia, Research and Development Center in Educational Simulation, 1968), pp. 55-56; and "Geography-Examinations, Questions, etc., and Tests and Scales," Education Index (New York: The H. W. Wilson Company, 1932-1974).

[2]For bibliographic references to and reviews of the last two groups of tests, see Oscar K. Buros, ed., The Nineteen Forty Mental Measurements Yearbook (Highland Park, N.J.: The Mental Measurements Yearbook, 1941), pp. 413-15; idem, ed., Third Mental Measurements Yearbook (New Brunswick, N.J.: Rutgers University Press, 1949), pp. 611-15; idem, ed., Fourth Mental Measurements Yearbook (Highland Park, N.J.: The Gryphon Press, 1953), pp. 661-63; idem, ed., Fifth Mental Measurements Yearbook (Highland Park, N.J.: The Gryphon Press, 1959), p. 850; idem, ed., Sixth Mental Measurements Yearbook (Highland Park, N.J.: The Gryphon Press, 1965), pp. 1231-32; idem, ed., Seventh Mental Measurements Yearbook, Vol. II (Highland Park, N.J.: The Gryphon Press, 1972), pp. 1308-10; Barbara A. Peace, "Bibliography of Social Studies Tests," in Evaluation in Social Studies, Thirty-Fifth Yearbook of the National Council for the Social Studies, ed. by Harry D. Berg (Washington, D.C.: National Council for the Social Studies, 1965), pp. 230-47; Susan Wood, "An Evaluation of Achievement Tests in Geography," in Evaluation in Geographic Education, The 1971 Yearbook of the National Council for Geographic Education, ed. by Dana G. Kurfman (Belmont, Calif.: Fearon Publishers, 1970), pp. 139-51; and Marion J. Rice, "Improving Elementary Geography: Tests in Geography," Elementary School Journal, LXVI (December, 1965), 134-35.

Teacher-made Tests

Tests produced by classroom teachers are marked by a number of common characteristics. They contain many items with outdated content, they emphasize recall of factual material, and they ignore many accepted guidelines for the construction of quality test items. But perhaps the most serious limitation of these tests is the lack of information on validity and reliability. Indeed, many tests in this group have been published with no indication of ever having been administered, much less standardized. The absence of performance data on these tests precludes any meaningful interpretation, comparison, or explanation of the results. Consequently, their utility in measuring geographic achievement is doubtful.

The two items listed below have been selected from different tests and are fairly typical of test items found on teacher-made tests. Carpenter suggests the following item as a measure of attainment of knowledge about the southern states of the United States.

12. _____, Alabama is often called the "Pittsburgh of the South."[3]

Zwicker uses the following item in a test on Italy and Sicily:

---

[3]Anna E. Carpenter, "A Test on the Southern States for Intermediate Grades," Grade Teacher, LXI (March, 1944), 55, 59, 71.

20. The mountains which extend north and south through the peninsula are the (Alps, Auvergnes, Apennines).[4]

These strictly recall-type questions can be criticized on two points--emphasis of trivial content and lack of accuracy. For instance, a student could receive credit on the Alabama item by recalling a simple memorized fact and never gain an understanding of the industrial similarity of the two cities to which the question refers. And if a student succeeds in remembering that the Apennines are in Italy, he will be misinformed by a question which tells him that these northwest-southeast trending mountains lie north and south.

Some of the tests in this group do display certain desirable characteristics. Baker developed a diagnostic, short-answer, map-reading interpretation test at the sixth grade level which measures attainment of such skills as the ability to determine direction, use scale, locate features on maps, identify different kinds of maps (e.g., rainfall and population, and make use of the legend to ascertain the meaning of map colors. This test has been used with sixth grade classes as a pretest, but has no data to verify its worth in assessing achievement.[5]

---

[4]Alma Zwicker, "A Test on Italy and Sicily for Upper Intermediate and Grammar Grades," Grade Teacher, LXX (September, 1952), 88, 90.

[5]E. V. Baker, "Diagnosing Children's Ability to Use Maps," Journal of Geography, XXXVII (September, 1938), 227-31.

Fuller constructed a test using various item types that makes use of a north polar projection. Questions require student understanding of locations, including relative locations, great circle routes, latitude and longitude, time concepts, and directions. No student performance data are given. While interpretation is occasionally required, most of the items can be answered by simple recall.[6]

Branom has formulated one of the most promising tests in this group. He offers a brief set of achievement test construction procedures from which a twenty-four item multiple-choice test on the St. Louis area has been constructed. The test contains graphs, pictures, maps, and reading passages that must be used by students to answer test items. Typical test items require students to locate certain physical and cultural features, select desirable transportation routes between points, determine maximum precipitation and temperature from a climograph, identify positions of rock layers and soil types, interpret reading passages, define special vocabulary terms, and make certain factual associations. Question number 7 is an example of an item requiring student use of a climograph:

---

[6]Kenneth A. Fuller, "Developing Map Reading Skills for Global Emphasis," Journal of Geography, XLII (September, 1943), 216-20; see also Elaine Forsyth, "Map Reading: Testing Skills," Journal of Geography, XLIII (May, 1944), 168-74.

7. As shown by the accompanying graph, the city of St. Louis normally has the heaviest rainfall during the month of (1) March, (2) April, (3) May, (4) June, (5) September.

Question number 23 is typical of items used to assess student interpretation of cultural facts.

23. St. Louis has been helped to become an important city (1) because of the keen competition of East St. Louis, (2) because it is located on the Mississippi River, (3) because flood plains provide much level land, (4) because it is the capital of the state of Missouri, (5) because the underlying rock is close to the surface.[7]

Here seems to be a valid test with quality items which require more than recall. However, since no technical data are provided, it is difficult to assess test quality. Reliability probably could be improved with addition of similar items, but further analysis would be necessary to provide the technical data required for meaningful assessment of item performance.

Brockmyer used pictures to assess fifth graders' retention of factual information about the United States. Her technique, for example, was to show students a picture of smudge pots burning in an orange grove and have them select, from a list of three, the state where the picture was taken. The exercise, however, requires only knowledge-level thinking since the necessary information needed to answer questions was provided in previous instruction;

---

[7]Mendel E. Branom, "Objective Diagnostic Testing Illustrations from the St. Louis Area," Journal of Geography, XLIV (September, 1945), 239-45.

students were not required to apply previously learned information to new situations. As with other tests reviewed from group one, no data are provided.[8]

In general, teacher-made tests are in greatest need of improvement. High-quality classroom tests serve a vital role in the school. Mehrens and Lehmann elucidate that role in the following statement:

> Classroom tests, because they can be tailored to fit a teacher's particular instructional objectives, are essential if we wish to provide for optimal learning on the part of the pupil and optimal teaching on the part of the teacher. Without classroom tests, those objectives that are unique to a particular school or teacher might never be evaluated.[9]

Further, these authors contend that teacher-made tests are deficient and are in need of improvement.[10] Thorndike and Hagen, too, have examined hundreds of teacher-made tests and have found them to be inadequate as evaluation devices.[11] Common faults of such tests include:

1. A failure to cover the range of teacher objectives. While teachers espouse a wide variety of

---

[8] Irene Brockmyer, "Testing with Pictures," Journal of Geography, L (February, 1951), 54-57.

[9] William A. Mehrens and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology (New York: Holt, Rinehart, and Winston, Inc., 1973), p. 169.

[10] Ibid., pp. 170-72.

[11] R. L. Thorndike and E. P. Hagen, Measurement and Evaluation in Psychology and Education (3rd ed.; New York: John Wiley and Sons, Inc., 1969), pp. 33-34.

objectives, their tests were found to require rote memory recall behavior on the vast majority of items.

2. Coverage of trivial content.

3. A type of test exercise not well adapted to measuring the content it is supposed to represent.

4. Ambiguous, unclear, test items.

5. An inappropriateness of the test for the teacher's purposes.

Commercially Available Geography Tests

There are seven geography tests in the second category which have been published and reviewed in Buros' Mental Measurements Yearbooks and elsewhere since 1931. The reviews have been done by individuals with expertise in testing and measurement, social studies, and, in some cases, geography. Their comments will serve here as the major source of information relative to these tests.

The Wiedefeld-Walther Geography Test for grades 4-8 was published in 1931. Reviews indicate a number of weaknesses, including speededness, failure to relate the source used to establish validity, overemphasis on factual information, confusing exercises, a difficulty level inappropriate for fourth and fifth graders, "correct" answers that, in some cases, are erroneous, and failure to keep maps and related questions on the same page. While the test authors purport to measure more than simple geographic information,

only one subtest--the one testing ability to read and interpret maps--received favorable comments to that effect from the reviewers.[12]

The Analytical Scales of Attainment in Geography was published in 1933 for use in grades 6-7. It is a 45-minute multiple-choice test that covers geographic vocabulary, human geography, industries, and products. To do well on this test, students would have to recall geographic information that ranges over the entire world and understand many terms that have little geographic significance. Raw scores, C-scores, and "geography ages" are reported, but no information is given on validity or reliability.[13]

In 1937, the Emporia Geography Test was developed for grades 4-7. Reviewer comments are quite critical, as revealed by the following:

> One has the feeling that the testmakers were apparently not clear in their own minds about the nature of the subject or how to test competence in it. . . . The test will, therefore, be helpful to the teacher who wishes to know whether his pupils are acquiring a wide knowledge of place locations and geographical information. . . . If one wishes to know whether children are attaining the aim of geography which consists in becoming geographical thinkers and reasoners, the test will be of little assistance.[14]

---

[12]Buros, The Nineteen Forty Mental Measurements Yearbook, op. cit., pp. 414-15; and Buros, Third Mental Measurements Yearbook, op. cit., pp. 614-15.

[13]Buros, The Nineteen Forty Mental Measurements Yearbook, op. cit., pp. 413-14.

[14]Buros, Third Mental Measurements Yearbook, op. cit., pp. 611-12.

At the time of review, 1949, the test had not been revised and was still on the market in its original form. Several items were out of date and therefore misleading to students. Percentiles for mid-year and end-of-year testing are reported for each grade in grades 4 through 7. Reliability levels seem to be adequate for groups but are too low for individual diagnosis.

The Geography Test: National Achievement Tests was designed for grades 6-8 and was first published in 1938. Reviews of this test are unfavorable, indicating that it is merely another instrument for measuring factual recall. It lacks comprehensiveness since no part of the test is designed to assess geographic skill attainment and no items require use of maps, pictures, graphs or reading material. The test authors are severely rebuked for the inclusion of incorrect information in several of the items. For example, one question asks: "Which country in North America is part of Great Britain?" Grade norms are provided, but no information on their derivation is given. Information on reliability and discriminating power of the items is lacking.[15]

Among the tests constructed for the intermediate grades, the Fourth Grade Geography Test of the National Council of Geography Teachers, published in 1940, has

---

[15] Ibid., p. 613; Peace, op. cit., p. 239; and Wood, op. cit., pp. 146-47.

received the most favorable reviews. The general consensus is that this test does require more than just recall from students. One reviewer comments:

> In contrast to the more usual questions on isolated facts, this test attempts to measure the child's understanding of relationships in the field of natural environmental conditions and in the area of man's reactions to these conditions. It also attempts to measure the child's ability to read maps, to read pictures and text, and to draw reasonable inferences from them.[16]

Validity of the test is based on the fact that questions were derived from content common to nine of the most widely used textbooks of the time. Reliability and discriminating power have been established and the range of item difficulties reported. Item development and final selection were effected through use of item analysis. The test has been criticized for its lack of published percentile norms, for not having comparable editions for retesting, and for not adhering strictly to test construction techniques.[17]

Another test receiving moderately favorable reviews is the Modern Geography and Allied Social Studies test for grades 6-8, published in 1938. Although the primary emphasis of the test was on Western civilization and almost totally neglected Eastern culture, this test had a number of merits. Perhaps the most outstanding

---

[16]Buros, Third Mental Measurements Yearbook, op. cit., p. 612.

[17]Ibid., pp. 612-13.

characteristic was that not all its various parts were devoted to recall of factual information; several parts required reasoning and judgment. Validity is usually claimed on the basis of extensive consultation of textbook material then in use, but no evidence is given to indicate test content review by experts and no use of external criteria was reported. Reliability was given as .92, computed on the basis of odd versus even scores; however, the population used for calculating norms was not specified. One can only assume that the group in grades 6-10 to whom the test was administered was also the norm group. In that event reliability coefficients probably would be lower for any one grade. The manual reports raw scores, raw score means, scaled scores, and standard deviations. Objectionable elements of the test include a suggested correction for guessing in the manual which is incorrect, absence of norms on the individual parts, emphasizing trivia on some of the items, questionable geographic accuracy of some correct options used, and violation of certain test construction principles. For example, on many multiple-choice items the correct choice is the longest one. In addition, several of the context-dependent items have cluttered and indistinct maps.[18]

The Hollingsworth-Sanders Geography Test, first published in 1962-63 for grades 5-7, is perhaps the most

---

[18]Buros, Fourth Mental Measurements Yearbook, op. cit., pp. 662-63; and Rice, op. cit., pp. 134-35.

recent intermediate level achievement test in geography. Unfortunately, since the test relies almost exclusively on recall of descriptive, definitional, and locational facts, it is little different from most of those published twenty to thirty years earlier. No emphasis is placed on geographic skills or relationships, and no maps or graphs have been used. While the authors claim equivalence of content for the two forms of the test, this is clearly not the case. For example, Form A has seventeen items on Africa and Form B has no African items. Reliabilities from .88 to .93 have been achieved, but it is questionable whether content validity has been established.[19]

Standardized Tests with Geographic Content

Tests in the third group achieve the highest technical quality of any tests reviewed in this chapter. This is not unexpected since these tests have been developed by people who are cognizant of the accepted principles of test construction; furthermore, these tests have been subjected to extensive preparation and standardization. Principal deficiencies of these tests as measures of geographic achievement are that: (1) since they are designed to measure social studies content, they do not provide detailed coverage of any one social science discipline; (2) in many

---

[19]Buros, Seventh Mental Measurements Yearbook, op. cit., pp. 906-7; and Wood, op. cit., p. 148.

cases the level of thinking does not rise above recalling factual information; and (3) much of the geographic content is inaccurate, out-of-date, or insignificant.

The Metropolitan Achievement Tests: Social Studies, first published in 1932, has an intermediate test designed for grades 5-6. If one considers both the factual knowledge and skill items together, geographic content is relatively well represented on this test. Sixty items, twenty of which relate to geography, are in the factual-recall part of the test, while the second part is devoted to map, graph, table, and chart reading and interpretation. Unfavorable comments from reviewers stem from the use of outdated information and material that is incorrect. Geographic determinism is all too evident in some of the questions, and one question lists as the correct option a figure which is inaccurate as China's approximate proportion of total world population.[20]

Geography is very poorly represented on the Social Studies Test: National Achievement Tests, grades 4-6, first made available in 1937. Less than 10 percent of the content is devoted to geography, and it is factual-recall material that hints of its pre-World War II origin. Unlike most tests in this group, this social studies test suffers from technical inadequacies also. No manual and

---

[20]Buros, Sixth Mental Measurements Yearbook, op. cit., pp. 1220-24; and Peace, op. cit., p. 234.

consequently no information on validity, reliability, or purpose are provided.[21]

Another intermediate-level test, the Geography Test: Municipal Tests: National Achievement Tests, first appeared in 1938. This test lacks important technical data on reliability, validity, and norms as well as descriptive information on the norming population. Reviews indicate that a period of twenty minutes is provided for the geography subtest and question whether a comprehensive assessment can be achieved in such a short time. Some items use trivial content, and inaccuracies characterize others. For example, knowledge of how houses are numbered is required on one item and students are informed that Canada is part of Great Britain on another. One reviewer concludes, "It is doubtful that valid conclusions could be drawn from the results with respect to the quality of instruction or learning in the field of geography."[22]

The Stanford Achievement Test: Intermediate Social Studies, first published in 1940, devotes approximately one-third of its items to geography. Most of the geography questions relate to knowledge of other countries including their natural resources, agriculture, place location and physical features. The content centers on traditional and

---

[21]Buros, Third Mental Measurements Yearbook, op. cit., pp. 608-9; and Peace, op. cit., p. 236.

[22]Buros, Fourth Mental Measurements Yearbook, op. cit., p. 662; and Peace, op. cit., p. 239.

factual geography, but it does not border on triviality. If, however, a prospective user wishes to assess student understanding, critical thinking abilities, and skills, he will not find this test to his liking. The construction, administration, and interpretation of this test have been done in a very professional manner and cause it to stand out among others in the social studies.[23]

According to Wood, the California Test in Social Studies and Related Sciences for grades 4-8, published in 1953, has little to recommend it. She criticizes the test for its lack of inspiration, emphasis on memorization of facts, outdatedness, and use of maps that are misleading and incorrect. Some items include stereotyped phraseology, like reference to torrid and temperate zones, while others refer to a map on which latitudinal-longitudinal distance is inconsistent with the linear scale distance given on the map.[24]

Two tests in this group that continue without equal in their effectiveness in measuring geographic skills are the Sequential Tests of Educational Progress: Social Studies (STEP) for grades 4-6, published from 1956-63, and the Iowa Test of Basic Skills: Work Study Skills, Map Reading (ITBS) for grades 3-9, published in 1964. Both

---

[23]Buros, Fifth Mental Measurements Yearbook, op. cit., p. 849; and Peace, op. cit., pp. 236-37.

[24]Wood, op. cit., pp. 143-44.

tests make valid use of pictures, maps, graphs, and charts to assess student ability to comprehend and interpret. The STEP intermediate test with thirty-three items focused on geography has perhaps the stronger representation of the discipline; however, since emphasis is on skills, other important geographic understandings are not adequately assessed. Technical quality and usability in general is quite good; nevertheless, one reviewer feels that content validity on the STEP could be more fully established.[25] The ITBS (Form 4) has thirty-two, thirty-six, and forty map-reading items for students in grades 4, 5, and 6 respectively, as well as a number of items for each grade designed to assess graph and table-reading skills. But, since emphasis is entirely on skills, assessment of student ability to apply and use geographic concepts and principles is lacking.[26] Reliability coefficients for both tests are very adequate and indicate a high level of internal consistency.

Commercial and standardized tests that have included items designed to assess geographic knowledge have fallen

---

[25] Buros, Sixth Mental Measurements Yearbook, op. cit., pp. 1224-26; Peace, op. cit., pp. 234-35; and Wood, op. cit., pp. 141-42, 145.

[26] E. F. Lindquest and A. N. Hieronymus, Iowa Test of Basic Skills: Multi-Level Edition for Grades 3-9 (Form 4) (Boston: Houghton Mifflin Company, 1964), pp. 52-64.

short of that goal. A number of probable explanations are suggested for this shortcoming.

1. Geographers and social scientists with professional zeal, wishing to make a needed and worthwhile contribution to the discipline, have constructed tests with little regard for proper procedures of test construction.

2. Many tests have inadequate technical data such as reliability; consequently, the bases for judging test performance and quality are absent.

3. Emphasis largely has been given to assessment of geographic skills to the neglect of concepts and principles from physical and human geography.

4. Too little attention has been given to the assessment of children's abilities to think critically about geographic phenomena.

Previous Findings Relative to Intermediate School Children's Knowledge of Geography

The research studies examined in this section are separated into two groups. In the first group are those studies which attempt to ascertain elementary students' knowledge of physical and human geography. Only a few studies were found for this group in the literature. The second group contains those studies aimed at determining student competence in the area of map reading and other skills important in geographic learning. Most of the research on geographic learning among elementary students has dealt with skills attainment; consequently, these studies will receive greater attention.

Research on Student Knowledge of Physical and Human Geography

Brown investigated sixth grade children's knowledge of important principles of physical geography. After identifying thirty-five important physical geography principles from textbooks and articles and employing the judgements of five authorities to validate those principles, an assessment instrument was constructed. After administering the test to a stratified proportional sample of 487 students in Kansas, he concluded that the curriculum in the schools sampled was not successful in teaching the thirty-five physical geography principles identified.[27]

Ellis worked with elementary students in an effort to determine their knowledge of topographical terms. He used multiple-choice questions and two different test presentation forms--a pictorial and a written form. His sample was 496 fourth and sixth grade pupils, randomly chosen, from twenty-classrooms in six school districts in Oregon. The results indicate that terms found to be difficult for fourth graders were also difficult for sixth graders, but that sixth graders generally achieved higher scores than fourth graders. Fourth grade boys scored

---

[27] Daniel H. Brown, "Knowledge of Important Principles of Physical Geography Possessed by Selected Sixth Grade Children," Dissertation Abstracts, XXIV (1963), 5072-73.

significantly higher than girls; at the sixth grade level the same differential was observed but it was not significant. Scores for both grades were higher on the pictorial form of the test. Many commonly used terms were missed, e.g., the term "cape" was the one most frequently missed in both grades. Other terms of considerable difficulty for fourth graders include "sand bar," "plateau," "cove," "channel," "peninsula," "bay," and "tributary." Sixth graders evidenced only moderate understanding of the terms "channel," "peninsula," "river," and "cove."[28]

Carnie has summarized research relative to development of views of other countries held by children ages ten and eleven. He found that, although there are differential amounts of knowledge known by different children, there is a set of ideas about countries and their inhabitants that contribute to the child's image of his world. He feels these images of places and people are derived primarily from emotional or affective attitudes, rather than from any objective assessment on the part of the child. Cognitive learning is inhibited by the development of affective viewpoints.[29]

---

[28]Arthur K. Ellis, "Fourth and Sixth Grade Students' Knowledge of Topographical Terms," Journal of Geography, LXVIII (September, 1969), 344-46.

[29]J. Carnie, "The Development of National Concepts in Junior School Children," in Perspectives in Geographical Education, ed. by John Bale, Norman Graves, and Rex Walford (Edinburgh, Great Britain: Oliver and Boyd, 1973), pp. 101-16.

Research on Student Knowledge of Geographic Skills

The research conducted on the attainment of geographic skills will be reviewed in three groups formed roughly on the bases of conclusions of the researchers. The first group contains studies which indicate that students have an acceptable command of a particular geographic skill or skills. The second group of studies, in general, found students deficient in geographic skills. Finally, those studies that point out that certain skills can be learned if properly taught will be reviewed.

Edwards tested fourth, fifth, and sixth grade students to determine their knowledge of cardinal points, altitude, area, and horizontal distance. Test items used in the assessment were based on commonly used textbooks. Examination of textbooks for each grade level allowed him to determine the maximum number of points on his test for which a particular grade could be held accountable. His results indicate that over 80 percent of the fourth and fifth grade students achieved a pre-established expected maximum, while 81 percent of the sixth graders achieved or exceeded the fifth grade expected maximum.[30]

Towler and Nelson experimented with elementary school children to ascertain the level at which the spatial

---

[30]John H. Edwards, "How Well Are Intermediate Children Oriented in Space?" Journal of Geography, LII (April, 1953), 133-43.

concept of scale develops. They used a stratified random sample of 120 students from three Edmonton, Alberta elementary schools. Ability to use scale, they conclude, is most strongly correlated with intelligence, chronological age, and grade level. Grade 5 and 6 students were able to use scale proficiently, but earlier grades had not yet developed the skill.[31]

Towler has also worked with elementary children to determine their ability to understand and use a system of coordinates to read and interpret maps. He designed a test which involved placement of objects on cards at different degrees of rotation and administered it to 120 children, age six to eleven years. As with the concept of scale, intelligence, chronological age, and grade level were the strongest correlates. He concludes that conventional reference systems can be used successfully by fifth grade, but with little or no success before this crucial age.[32]

Blaut and Stea conducted research to discover when children are capable of reading and interpreting aerial photographs. Their test population included lower and lower-middle income children from Massachusetts, children from Puerto Rico, and an isolated peasant community in the

---

[31]John O. Towler and L. D. Nelson, "The Elementary School Child's Concept of Scale," Journal of Geography, LXVII (January, 1968), 24-28.

[32]John O. Towler, "The Elementary School Child's Concept of Reference Systems," Journal of Geography, LXIX (February, 1970), 89-93.

West Indies. Their research indicates, at least for the population tested, that ability to read and interpret aerial photographs is fully developed by age nine and that this skill is learned informally, as a part of the child's normal developmental process.[33]

Several research studies point out that students lack necessary geographic skills. For instance, Howe conducted research with elementary students in New England to discover their knowledge of directions in "real" space and on a map. He concludes that (1) directions are not as well known as originally believed; (2) knowledge of directions is learned primarily through study of geography in school and that very little is acquired from home or other places; (3) intermediate level children have a better understanding of direction than do primary, but students at both levels still performed inadequately; (4) children frequently make wrong associations in their efforts to determine direction; and (5) boys excelled in knowledge of directions.[34]

Thomas reports the results of research designed to measure the ability of students in grades 4 through 7 to interpret several types of graphs. She tested 355

---

[33]J. M. Blaut and D. Stea, "Studies in Geographical Learning," Annals of the Association of American Geographers, LXI (1971), 387-93.

[34]George F. Howe, "A Study of Children's Knowledge of Directions," Journal of Geography, XXX (October, 1931), 298-304.

children selected from slow, average, and superior achieving students in each grade. Results revealed that fourth graders were the least adept at graph interpretation, and that students above this level could interpret simple types of graphs if they were adapted to their attainments. Most students experienced problems interpreting line graphs. The circle graph, two-dimensional diagram, and picture graph proved to be the easiest to interpret.[35]

Thorp investigated the success of children in using certain necessary geographic tools to learn geography. She tested a sample of 556 pupils from rural-consolidated, village, small city, and large city districts in grades 4 through 8. Several findings resulted. First, students in all grades could locate major parallels and the poles on the globe, select correct directions of other cultures from their own, and depict the life style of an Eskimo family. There was, however, a nearly complete lack of understanding of latitude and longitude. Students in grades 4 to 6 were only moderately successful in recognizing continental land masses and larger bodies of water. Second, concerning map-reading skills, students were deficient in use of scale to determine distance between cities until sixth grade, when moderate success was achieved.

---

[35]Katheryne Colvin Thomas, "The Ability of Children to Interpret Graphs," The Teaching of Geography, Thirty-Second Yearbook of the National Society for the Study of Education (Bloomington, Ill.: Public School Publishing Company, 1933), pp. 492-94.

Children were able to use the legend but did not know what name to apply to it. Third, all grades performed poorly on understanding of climatic elements. Finally, the ability to use indices, appendices, and graphs was lacking, with upper grades performing somewhat better than the lower grades.[36]

Anderson used questions from the Iowa Test of Basic Skills to measure sixth, seventh, and eighth grade student's knowledge of map reading, understanding of geographic terminology, and ability to read graphs and charts. He discovered a general inability of the students tested to perform adequately on items related to the areas above. Poor performance was particularly characteristic of the sixth grade, where average raw score achievement was around 50 percent.[37]

Zimmer developed a test with established reliability and validity to evaluate six aspects of map-reading skills among fifth, sixth, and seventh graders in Indianapolis,

---

[36]Mary T. Thorp, "Studies of the Abilities of Pupils in Grades Four to Eight to Use Geographic Tools," The Teaching of Geography, Thirty-Second Yearbook of the National Society for the Study of Education (Bloomington, Ill.: Public School Publishing Company, 1933), pp. 494-506.

[37]Howard R. Anderson, "Testing Basic Skills in the Social Studies," Elementary School Journal, XXXVI (February, 1936), 424-35. Similar findings are reported by Louise D. Wagner, "Measuring the Map-Reading Ability of Sixth-Grade Children," Elementary School Journal, LIII (February, 1953), 338-44.

Indiana. The skills include using appropriate terms, recognizing and using symbols, ascertaining directions, measuring distance, using latitude, and making inferences. Three final test forms were administered to 1,167 pupils in grades 5, 6, and 7 in an effort to reveal common difficulties in the use of map skills. Forty-six "core items" were common to all forms of the test. Results indicate that students had difficulty with vocabulary, such as "river mouth" versus "river source;" common map symbols, such as those used to identify city size; directions, such as reversing east and west; map scale to determine distance; coordinate systems; and correlating of information from several maps. Several possible sources of pupil error were suggested, e.g., the belief that north is up and south is down, that weather becomes increasingly colder northward and increasingly warmer southward, and that degrees of latitude equate with degrees of temperature. In most cases the errors commonly made by fifth graders were made with approximately equal frequency in grades six and seven. However, results on the "core items" alone show progressive improvement from grades five through seven.[38]

Bartz attempted to determine elementary and middle-school children's ability to use certain map skills

---

[38]Sister Stephanie Zimmer, "A Diagnosis of Map Skill Deficiencies in Elementary School Children," Dissertation Abstracts, XXVIII (1967), 2107A.

also. Unlike Zimmer, she concluded that by seventh grade most students can use scale to determine distance between points on a map, but admits that many forms of scale are not well understood by children. Relating two different map scales, as on a map with an inset, seems to constitute a major problem for most students. She reports that children have even greater problems using scale to determine area. Children are able to use a two-way coordinate location system like that found on a road map, but transfer of this understanding to a spherical surface for location leads to "grave perceptual and conceptual problems."[39]

Several individuals have carried out research studies designed to determine if elementary school pupils can learn geographic skills if they are properly taught. Thorp used an experimental group in her study and concludes that sixth graders can master geographic skills involving use of the globe, reading graphs and maps, and using the index and appendix.[40] Howe obtained similar results with fourth, fifth, and sixth graders when they were provided remedial instruction on map-reading skills.[41] Rushdoony,

---

[39]Barbara S. Bartz, "Maps in the Classroom," Journal of Geography, LXIX (January, 1970), 18-24.

[40]Thorp, op. cit., pp. 505-6.

[41]George F. Howe, "A Study of the Ability of Elementary School Pupils to Read Maps," The Teaching of Geography, Thirty-Second Yearbook of the National Society for the Study of Education (Bloomington, Ill.: Public School Publishing Company, 1933), pp. 487-92.

working with map-reading skills, has shown that those skills formally taught at grades 4 and 5 can be learned by third graders and has recommended a downward shift in the curriculum for the teaching of certain map-reading skills.[42] Carswell investigated the topographic map-reading and interpretation abilities of students in grades 4, 5, and 6 and observed changes in these abilities following a period of planned instruction. He concludes that students at these grade levels can learn to use large-scale topographic maps effectively and that scale, symbols, direction, the grid system, and elevation can be mastered.[43]

Conclusions Relative to Research on Intermediate Pupil's Knowledge of Geography

The review of research involving what elementary children know about geography would seem to warrant the following conclusions:

1. Findings of various researchers are conflicting, with some studies indicating, for example, that certain geographic skills are known, others indicating they are not known.

2. Many studies have not utilized tests meeting minimal criteria such as established validity and reliability.

---

[42]Haig A. Rushdoony, "Achievement in Map-Reading: An Experimental Study," Elementary School Journal, LXIV (November, 1963), 70-75.

[43]Ronald John Berry Carswell, "Topographic Map Reading Abilities of Learners in Grades Four, Five, and Six," Dissertation Abstracts, XXX (1968), 200A.

3. There appears to be an insufficient number of quality investigations, particularly in the area of understanding and applying geographic concepts and principles.

Douglass called attention to the lack of adequate research on map reading several years ago. He states:

> Learning to read maps continues to be one of the major emphases in geographic instruction. Map-reading, however, involves a wide range of skills and understandings. We need to know, therefore, what skills and understandings are needed in reading maps and the most propitious times at which these should be introduced. At the present time, no conclusive evidence is available.[44]

Barton agrees that research on geographic learning has been incomplete. He points out:

> It is impossible to design a precise numbered sequence for developing geography skills because as yet insufficient evidence has been collected by research in the behavioral sciences, experimental teaching, diagnostic testing, and other kinds of research. All that is available today are some suppositions as to what is or what is not too difficult for various groupings, and some suggestions concerning a sequence of techniques based on limited observations, logic, and judgment.[45]

In his summary of the pre-1960's research on children's ability to read maps, Rushdoony concludes that we cannot draw any generalizations from the research relative to specific map-reading skills of children. There has

---

[44]Malcolm P. Douglass, "Interrelationships Between Man and the Natural Environment for Use in the Geographic Strand of the Social Studies Curriculum" (unpublished Ph.D. dissertation, Stanford University, 1954), pp. 211-12.

[45]Thomas Frank Barton, "Geography Skills and Techniques," Curriculum Guide for Geographic Education, ed. by Wilhelmina Hill (Oklahoma City: Harlow Company, 1964), p. 69.

been an increase in map-reading studies in the sixties, but extensive, longitudinal research involving several classes engaged in study for extended periods of time is still lacking.[46]

Bacon and Kennamer also have identified the general inadequacy of research in geographic education. They state emphatically,

> At almost every turn, one is confronted by an ever-increasing concern that geographers have not and are not addressing themselves in significant numbers to problems related to geographic education. Geography in our schools is most strongly centered in the elementary grades; yet there is general dissatisfaction expressed by both geographers and educators about the nature of the offerings and the methods of instruction. Improvement in program and instruction are likely to come only as research in geographic education shows the way.[47]

---

[46]Haig A. Rushdoony, "A Child's Ability to Read Maps: Summary of the Research," Journal of Geography, LXVII (April, 1968), 213-22.

[47]Phillip Bacon and Lorrin Kennamer, Jr., "Introduction," Research Needs in Geographic Education, Geographic Education Series No. 7 (Normal, Ill.: National Council for Geographic Education, 1967), p. 11.

CHAPTER III

RESEARCH METHODOLOGY

The methods of research used to carry out this study will be detailed under two major headings in this chapter. First, the procedures used in development of a standardized geography assessment instrument will be set forth. These procedures involve: (1) formulation of purposes, assumptions and requirements of the test; (2) devising an appropriate test blueprint to guide test development; (3) stating the learning objectives to be assessed; (4) elaborating the methods used in writing test items and construction of the test instrument; and (5) explaining the manner in which test validity and reliability were established. Production of a suitable standardized test was crucial to the study since no adequate instrument was available.

Second, the techniques utilized in selection of a proportional stratified sample of Michigan's fifth grade student population will be explained. A proportional stratified sampling technique was used to obtain a representative sample. Use of this sampling procedure is

considered appropriate where the size of the total population strata is known.[1] In addition, the manner in which the data were collected from the selected sample of Michigan fifth grade students will be explained. The chapter concludes with a statement of the research hypotheses established for the study.

Developing the Test Instrument

Purpose, Requirements, and Assumptions

The Michigan Intermediate Geography Test (MIGT) has been constructed to serve one principal purpose--measurement of what fifth grade students know about the geography of Michigan and the United States. Specific dimensions of achievement selected for assessment include geographic content and cognitive behavior. Geographic content has been divided into physical geography facts, concepts, and understandings; human geography facts, concepts, and understandings; and geographic skills. Cognitive behavior categories include recalling important factual material and interpreting and applying geographic content in a problematic situation.

---

[1]Hubert M. Blalock, Jr., Social Statistics (New York: McGraw-Hill Book Company, 1960), pp. 399-400.

Certain assumptions have been made about the construction and use of the MIGT:

1. It was assumed that different cognitive behaviors would be required to answer different types of questions and that test items should assess the student's ability to perform at various cognitive behavior levels.

2. It was assumed that students should possess a basic knowledge of physical and human geography, as well as an understanding of basic geographic skills.

3. It was assumed that both individual student and school district achievement would vary from very low to very high and that sampling across the entire range would permit generalizing about the achievement levels of Michigan's fifth grade population.

The development of any standardized test of achievement begins with a rationale which lends guidance and direction to the test constructor in formulation of the test. MIGT is based on the following rationale:

1. Geography taught in the elementary schools in Michigan should reflect contemporary work in the discipline.[2]

---

[2]Neville V. Scarfe, "Designing the Curriculum to Develop Geographic Concepts," The Journal of Geography, LII (March, 1953), 100; Zoa A. Thralls, The Teaching of Geography (New York: Appleton-Century-Crofts, Inc., 1958), p. 16; Gertrude Whipple, "Geography in the Elementary Social Studies Program: Concepts, Generalizations, and Skills to Be Developed," in New Viewpoints in Geography, Twenty-ninth Yearbook of the National Council for the Social Studies, ed. by Preston E. James (Washington, D.C.: National Council for the Social Studies, 1959), pp. 112-13; and Theodore Kaltsounis, "An Analysis of Teaching Strategies in Emerging Geography Curricula," in Focus on Geography: Key Concepts and Teaching Strategies, Fortieth Yearbook of the National Council for the Social Studies, ed. by Phillip Bacon (Washington, D.C.: National Council for the Social Studies, 1970), p. 409.

2. Geography taught in the elementary schools in Michigan should focus on application of concepts and principles as well as the acquisition of important facts, interpretation, and skills abilities.

3. Geography taught in the elementary schools in Michigan should teach worthwhile, usable information that will serve as the basis for future learning.

The Test Blueprint

Test blueprints define the scope and emphasis of achievement tests, and guide the selection of appropriate content and cognitive behaviors on which to base test items. Most blueprints are constructed by placing content categories on the one axis of a matrix and cognitive behavior categories on the other axis. The content axis indicates the nature and relative importance of each content area to be assessed while the cognitive behavior axis states the nature and relative importance of each type of intellectual activity required to answer particular questions.

Test constructors give attention to the cognitive behavior dimension of test blueprints because extending the range and complexity of thought processes is a primary goal of education. One scheme for analyzing learning outcomes from the perspective of cognitive behavior is the Taxonomy of Educational Objectives: Cognitive Domain.[3]

---

[3]Benjamin S. Bloom, ed., Taxonomy of Educational Objectives; Handbook I: Cognitive Domain (New York: Longmans, 1956).

The Taxonomy establishes six behavioral domains for which test items could be written:

1. Knowledge--This behavioral domain requires that the student remember or recall ideas, material or phenomena he has acquired and stored in his mind in a previous learning situation. Information stored could relate to specifics such as terminology, particular places or events or principles and generalizations germane to a discipline.

2. Comprehension--This behavioral domain requires that students, when confronted with an oral, written, or symbolic communication, know what is being communicated and be able to make some use of the information communicated. Comprehension behaviors include translation, interpretation, and extrapolation.

3. Application--This behavioral domain requires students to apply known and comprehended information in a correct manner in a new situation where no previous modus operandi and/or solution has been specified. Use of an abstraction comprehended from a verbal or pictorial communication to derive a correct solution in a previously unencountered problematic situation is an example of application behavior.

4. Analysis--This behavioral domain requires that students break down information into its constituent parts and detect relationships and organizational patterns inherent in the information. Analysis behavior aids in fuller comprehension and serves as a prelude to evaluation of information. Such behavior could involve students in noting how ideas relate to one another, in the identification of implicit assumptions in a communication, or in recognizing organizational principles of a communication.

5. Synthesis--This behavioral domain involves students in piecing together elements and parts of a communication so as to form a pattern or structure not clearly evident before. Synthesis permits the use of creative behavior to develop something unique and original from a diversity of informational sources.

6. Evaluation--This behavioral domain involves students in the making of judgments relative to the worth or value of ideas, methods, or materials in a particular situation. It requires students to use self-derived or established criteria and standards for appraising the extent to which particulars are accurate, effective, economical, or satisfying and may involve both qualitative and quantitative decisions.

The content dimension is an important part of any test blueprint. "Content is the vehicle through which behavioral objectives are taught and through which, once learned, they are demonstrated."[4] The content categories listed on the content axis of the test blueprint determine the type and specificity of the items which a test constructor writes. The degree of specificity is dependent on the purposes of the test. Wood suggests that subdividing content categories in the blueprint helps define more exactly the areas where questions can be written and can provide valuable assistance to a test constructor in deciding the number of items to construct for each content domain in a test blueprint.[5]

The test items ultimately constructed for each cell in the blueprint, and therefore the final test

---

[4]Sherman N. Tinkelman, "Planning the Objective Test" in Educational Measurement, ed. by Robert L. Thorndike (Washington, D.C.: American Council on Education, 1971), p. 52.

[5]D. A. Wood, Test Construction: Development and Interpretation of Achievement Tests (Columbus, Ohio: C. E. Merrill, 1960), p. 31.

itself, represent only a sample of the total number of items that could be constructed. The test blueprint insures that a representative sample is drawn from the universe of tasks constituting achievement in a given cell of the blueprint. Such representativeness is essential if one wishes to draw valid inferences about student achievement in a course and to achieve content validity.[6]

The MIGT Test Blueprint

There is no standard scheme that should be used for every test blueprint. Each is designed for the special purposes of each test. The blueprint is not a rigid instrument that remains unalterable during item construction. Adjustments based on logic and reason are made whenever they will increase the probability of achieving originally stated purposes. Writing test items often helps clarify the test objectives, and the blueprint is often altered as items are written. Tinkelman says:

> In general therefore, it is undesirable to view the original test outline as frozen; rather, the test author should deliberately strive to improve the test outline as the test is being built. Probably, the test outline cannot be considered final until after the items have been written and tried out, and the final test assembled.[7]

---

[6] Tinkelman, op. cit., p. 49.

[7] Ibid., p. 55.

The two-by-three test blueprint for MIGT is given in Table 1. It provided a framework for the development and classification of test items and insured that approximately equal representation of the cognitive behavior X geographic content categories was achieved. Each cell in the blueprint represents a domain of knowledge that was sampled by test items on the MIGT. However, its primary function was enabling development of specific objectives to be assessed by the MIGT.

Learning Objectives

Specialists in educational testing and measurement usually agree on the role and characteristics of objectives in the test construction process. Marshall and Hales define achievement testing as the determination of the extent to which the objectives of instruction have been achieved and argue that objectives suitable for achievement testing must be stated in terms of expected student behavior.[8] Gerberich holds that construction of achievement tests depends "on the objectives of instruction and the behavioral changes, or outcomes, developed in pupils as a result of instruction."[9] Mehrens and Lehmann emphasize

---

[8]Jon Clark Marshall and Loyde Wesley Hales, Classroom Test Construction (Reading, Mass.: Addison-Wesley Publishing Company, 1971), p. 12.

[9]J. R. Gerberich, Specimen Objective Test Items: A Guide to Achievement Test Construction (New York: Longmans Green, 1956), p. 13.

TABLE 1.--Test blueprint for the Michigan Intermediate Geography Test

Cognitive Behaviors	Geographic Content		
	Physical Geography	Human Geography	Geographic Skills
Knowledge or Recall	23,* 26, 28, 29 32, 41, 42	6, 9, 11, 34	18, 20, 21, 22, 25, 30, 31, 33, 37
Understanding	12, 16, 24, 43, 45, 46, 47, 48, 49	2, 4, 5, 10, 13, 27, 44	1, 3, 7, 8, 14, 15, 17, 19, 35, 36, 38, 39, 40

*Numbers refer to the classification of the test items on MIGT.

the role objectives play in prescribing direction for our educational efforts as well as evaluation of student learning:

> They tell us in which way to head, a decision that is necessary before taking the first step on an educational journey. Specifically, objectives help an instructor in instructional planning, guide student learning, and provide a criterion for evaluating student outcomes.[10]

Even Anderson, who is critical of the traditional test construction procedure, concedes that there possibly is some merit in the use of objectives stated in behavioral terms. He states: "I am moderately enthusiastic about behavioral objectives. It surely makes good sense to be as explicit as you can about what you want to teach toward, and test for."[11]

But how are these objectives established? The traditional approach to the construction of an achievement test involves the derivation of objectives from the textbooks, instructional materials, and curriculum documents currently in use. This approach was not feasible for this study since the geography curriculum in the elementary schools in Michigan was found to be highly diversified,

---

[10]William A. Mehrens and Irvin J. Lehmann, Measurement and Evaluation in Education and Psychology (New York: Holt, Rinehart, and Winston, Inc., 1973), pp. 18-19.

[11]Richard C. Anderson, "How to Construct Achievement Tests to Assess Comprehension," Review of Educational Research, XLII, No. 2 (1972), 160.

and in some cases nonexistent.[12] Therefore, objectives on which to base test items could not be derived in the traditional manner. Instead, MIGT is based on geographic curriculum elements thought to be desirable for Michigan's intermediate students.

MIGT rests on two assumptions about teaching and learning geography in the elementary schools:

1. Elementary school geography should be consistent with the structure of the parent discipline--human geography, physical geography, and geographic methodology.

2. Elementary school geography in Michigan should focus on understanding concepts and principles as well as acquiring important facts.

These two assumptions in effect constitute basic goals for geographic education in elementary schools, and it is from those goals that the objectives assessed by MIGT were derived.

Thorndike and Hagen provide a list of desirable attributes for instructional objectives[13]:

1. They should be stated in terms of some observable student behavior.

2. They should begin with an active verb.

---

[12]An initial survey of the geography textbooks available for use in the intermediate grade classrooms of Michigan's elementary schools was conducted by members of the Board of Directors of the Michigan Council for Geographic Education in 1971-72.

[13]R. L. Thorndike and E. P. Hagen, Measurement and Evaluation in Psychology and Education (3rd ed.; New York: John Wiley and Sons, Inc., 1969), pp. 36-37.

3. They should be stated precisely using terms of uniform meaning.

4. They should focus on only one intellectual process.

5. They should be stated at an appropriate level of generality.

6. They should represent the intended outcomes of a series of planned learning experiences.

7. They should be realistic, i.e., they should take into consideration teaching time and student characteristics.

MIGT Learning Objectives

Given these guidelines for instructional objectives and the two basic goals for geographic education discussed above, the following objectives were formulated for MIGT.

Geographic Skills Objectives

Given appropriately constructed maps the student will demonstrate understanding by answering correctly questions involving:

1. the use of map scale to determine size and the distance between points;

2. the use of legend to determine numbers of people moving from place to place;

3. identification and location of commonly recognized physical and cultural features;

4. the use of coordinates to determine location;

5. spatial relationship of political units;

Given appropriately constructed graphs the student will demonstrate understanding by answering correctly questions involving:

6. the use of correct terminology;

7. ability to read, interpret, and extrapolate information displayed in the graphs.

Human Geography Objectives

Given a series of human geography questions with appropriate maps or pictures the student will demonstrate understanding by selecting correct answers on questions involving:

1. the most favorable location of a gasoline station within an urban settlement;

2. the regular movement of people from suburb to central city;

3. the use of land located along the peripheries of major cities;

4. knowledge of major institutions and industries in Michigan and the United States;

5. knowledge of urban terminology;

6. environmental factors associated with urban areas;

7. reference material for information on urban places;

8. recognition of likely combinations of crops and livestock on a Michigan farm.

9. regional terminology applied to the group of states including Michigan.

Physical Geography Objectives

Given a series of physical geography questions with appropriate maps, or pictures the student will demonstrate understanding by selecting correct responses to questions involving:

1. a determination of the direction of flow of water in a river;
2. knowledge of the origin of certain physical landscape features;
3. appropriate terminology for the landscape of Michigan and surrounding states;
4. the influence of Lake Michigan on the temperatures in western Michigan in fall;
5. climatic and vegetative patterns in various parts of Michigan and the United States;
6. the processes associated with erosion of the landscape;
7. the movement of the sun in relation to the earth.

The Test Item

Wesman identifies the mind of the item writer as the main source of ideas from which test items originate. He believes that the formulation of good ideas is the most difficult part of item writing and that a quality test depends largely upon the success of an item writer in originating ideas on which novel test items can be constructed.[14] However, two other sources of ideas for

---

[14]A. G. Wesman, "Writing the Test Item," in Educational Measurement, ed. by Robert L. Thorndike (Washington, D.C.: American Council on Education, 1971), p. 86.

items are frequently used. Often student difficulties on certain issues and problems are revealed in written and verbal communication. Sometimes an item writer can capitalize on the source of such difficulties and produce items that are highly discriminating. Another useful source for ideas is related to job analysis. The item writer might rely on descriptions of what individuals proficient in a certain job can or cannot do as a source of ideas for building useful items. Regardless of the source, those ideas for measuring levels of thinking higher than recall are the most difficult to develop. It is the formulation of these ideas that requires the greatest knowledge of one's subject matter and the most work in developing appropriate novel situations on which to base good test items.[15]

Because they are too simple or too difficult, some ideas might lead to test items that are nondiscriminating. In some cases, however, it may be desirable to include these items in a test. This would be appropriate if such items contribute to the apparent validity of a test or have a significant influence on learning. Inclusion of nondiscriminating items would be particularly desirable where a test might exert influence on the curriculum or the subject matter of an individual teacher or school district. This use of a test would be especially valid where the curriculum is in need of updating. As Wesman puts it, "In

---

[15] Ibid., pp. 86-87.

the areas where teaching practices are somewhat behind the recommendations of leaders in the field, tests can be of some help in leading the way toward improvement."[16]

It is generally agreed that test items should be expressed as clearly as possible.[17] Ambiguity can increase the difficulty and reduce the discriminating power of an item. Difficulty should be attributable to the problem presented and not to the language used. Ambiguity may arise from an inappropriate choice or awkward arrangement of words or from failure of the item writer to clarify his own thinking before writing the item. Ambiguity can be minimized by using words with precise meanings. Technical vocabulary should be used only if it has been taught as a part of a teacher's objectives or the test constructor has good reason to believe that such vocabulary is known to examinees. The objective for achievement measurement is assessment of content knowledge and not the student's reading ability, although it is impossible to do more than minimize this latter factor. Questions should be phrased as directly as possible. Unnecessarily specific and non-functional words should be avoided. Only meaningful and significant generalizations should be used, and accuracy should be maintained in all parts of the item; stereotyped, textbook phraseology should not be used.

---

[15] Ibid., pp. 86-87.

[16] Robert L. Ebel, Essentials of Educational Measurement (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972), p. 200; and Wesman, op. cit., p. 102.

There is general agreement that the multiple-choice item is the most popular objective item type in use today. It is free from many weaknesses inherent in other item forms and is adaptable to a wide range of item topics and objectives. Thorndike and Hagen identify two principal parts to a multiple-choice item--the stem and the options. The stem is an introductory question or incomplete statement that presents the problem.[18] Most experienced item writers prefer the incomplete statement stem since it provides for a smooth transition from the problem to the answers. However, item writers using the incomplete stem must be admonished against creating grammatical inconsistencies between stem and options that serve as clues for the test-wise student. In general, the incomplete statement is most appropriate for students above the sixth grade level. Younger students usually comprehend and perform better when the question form is used in the stem.[19]

There is no set number of options that should be provided with multiple-choice items. While four or five options are common, there is no clear and definitive answer concerning how many are best. In theory, the more options the better since each additional option reduces further the chance of guessing. In practice, however, seldom can more

---

[18]Thorndike and Hagen, op. cit., p. 103.

[19]Jum C. Nunnally, Educational Measurement and Evaluation (New York: McGraw-Hill Book Company, 1972), p. 169.

than two or three meaningful discriminating distractors be derived, so additional options, in most cases, do not improve item quality. The number of choices would seem to be less important than the effectiveness of those used. Other considerations include economy of reading time and the age of the examinees. The primary concern should be how well each option functions.[20]

The content or context-dependent item relies on an external source of information, such as a picture, map, graph, table, diagram, or reading, to present the problem. Two forms of the content or context-dependent item are common. The pictorial form is ideally suited for the elementary grades where reading is not highly developed, and where a picture can present a problem clearly. The interpretative form requires student interpretation of a reading passage, map, diagram, or graph and is well suited to measuring higher levels of cognitive behavior. Use of this type of item has been increasing in recent years and seems to have inherent advantages for assessment of geographic understandings. Such items can be used to assess a student's ability to read, interpret, and draw inferences from maps and other representational forms.[21]

---

[20] Wesman, op. cit., pp. 99-102.

[21] Ibid., pp. 120-28; and Mehrens and Lehmann, op. cit., pp. 291-306.

Content-dependent items provide a more realistic basis for item construction than does straight verbal reasoning. They require less complex language, particularly when the pictorial form is used, and provide for a minimum of contamination from reading comprehension and irrelevant factual material. They lend structure to a problem and provide a similar frame of reference for both teacher and student. In addition, they permit asking of several questions from a single map, picture, complex passage, etc.

Item difficulty.--The difficulty level of items can have a major impact on validity and reliability of a test. For this reason, defining the level and distribution of item difficulties is among the most important tasks for the test planner. Item difficulty has been defined by Tinkelman as the percentage of students selecting a correct answer.[22] Ebel, however, defines item difficulty as "the proportion of examinees in a group who do not answer the test item correctly."[23] Although both definitions are used, this study will employ the definition given by Ebel.

It seems reasonable to assume that a test should have items with a range of difficulty from very low to very

---

[22]Tinkelman, op. cit., p. 62.

[23]Ebel, Essentials of Educational Measurement, op. cit., p. 553.

high. This assumption ignores the fact that extremely easy and extremely difficult items contribute very little to the effectiveness of an achievement test. When achievement tests are given for the purpose of ranking students, as most classroom tests are, then any item that is answered either correctly or incorrectly by nearly all students cannot contribute to the purpose of the test. As Tinkelman explains,

> The test reliability and the variance of the test scores increase as the variance of the item difficulty distribution decreases. That is, it is generally preferable that the items in a test have a fairly narrow range of difficulty around the average difficulty level.[24]

An exception to this general rule would be when test content is very homogeneous or the students are an extremely heterogeneous group. Under these circumstances, one might want to strive for a wider spread of difficulty among the test items. In practice, few test constructors can achieve all items of medium difficulty; even when a conscious attempt is made, a certain amount of spread of item difficulty is to be expected.[25]

Optimal or medium difficulty is achieved when approximately half of the examinees answer each item correctly where there is no chance for guessing. Where guessing is possible, medium difficulty would be halfway

---

[24]Tinkelman, op. cit., pp. 63-64.

[25]Ibid.

between chance and perfect success. For four choice multiple-choice items there would be a 25 percent chance of success by guessing, so medium difficulty would be represented by an item answered correctly by 62 percent of the examinees. Such an item would have an index of difficulty of thirty-eight, i.e., 38 percent of the respondents missed the item. The literature seems to be quite clear regarding the level of item difficulty recommended for achievement tests designed to rank students.[26] Items of medium difficulty yield the most information about differences in student achievement and should compose, by far, the majority of the items on classroom achievement tests.[27]

---

[26]See, for example, Wood, op. cit., p. 37; Ebel, Essentials of Educational Measurement, op. cit., p. 116; Marshall and Hales, op. cit., pp. 223-24; Nunnally, op. cit., p. 187; Harold Gulliksen, "The Relation of Item Difficulty and Inter-Item Correlation to Test Variance and Reliability," Psychometrika, X (1945), 79-91; Marion W. Richardson, "The Relation Between the Difficulty and Differential Validity of a Test," Psychometrika, I (1936), 33-49; L. J. Cronbach and W. G. Warrington, "Efficiency of Multiple-Choice Tests as a Function of Spread of Item Difficulties," Psychometrika, XVII (1952), 127-47.

[27]Other test specialists, while in essential agreement with those above, advocate the use of items that are somewhat easier than medium difficulty. Tinkelman elaborates on the reason for this position on item difficulty. "The explanation for this shift toward relatively easier items lies in the fact that the error variance due to chance tends to be greater for more difficult items where more guessing occurs." Ebel's research findings are in agreement on the question of guessing on more difficult items. He found that blind guessing is more typical of poor students and that more guesses were made more frequently on difficult than on other items. See Tinkelman,

Developing test items for MIGT.--Several preliminary steps were taken in the development of test items for the final version of the MIGT. Efforts were made to devise novel items which constituted a fair sample of the content domains represented in the test blueprint. Such representation was a prerequisite to the establishment of content validity.

The initial step taken involved originating ideas and illustrative material on which to base appropriate problem situations. Ideas and illustrations were devised that were in keeping with the purpose of the test and which conformed to the test blueprint. Item ideas, in some cases, were based on situations which were intuitively surmised to be areas of commonly held student misconceptions. For example, the ideas that the north pole is always at the top of a map and that north is "up" and south is "down" were used to originate certain test items. Ideas applicable to real-life situations were used also to construct test items. In one instance, the concept "journey to work" was used to formulate such an item for the human geography X understanding cell of the test blueprint.

---

op. cit., p. 63; Thorndike and Hagen, op. cit., pp. 48-50; F. M. Lord, "The Relation of the Reliability of Multiple-Choice Tests to the Distribution of Item Difficulties," Psychometrika, XVII (1952), 181-94; and Robert L. Ebel, "Blind Guessing on Objective Achievement Tests," Journal of Educational Measurement, V (1968), 321-25.

Multiple-choice stems were written to conform with the original ideas and/or problem situations given in the illustrations. A concerted effort was made to eliminate possible sources of ambiguity in the item stems. Questions were phrased in as direct, clear, and concise a manner as possible and at a level deemed to be appropriate for the reading ability of students in the intermediate grades. Technical vocabulary was used only where its meaning could be ascertained from the illustration or where it was considered important geographic content that should be part of the student's knowledge. Caution was exercised to maintain item independence, so that content of one question did not provide clues to answers in other questions.

Item options were constructed in accordance with the established guidelines presented earlier in this chapter. In each case, the correct option was clearly correct or appeared as the best choice for those students who possessed sufficient knowledge of the geographic fact or concept measured. Specific determiners and clues to the correct choice were avoided. For example, the correct choice did not "stand out" because it was consistently more lengthy than the foils. Option identifiers for the correct choice were selected in approximately equal numbers and on a random basis to avoid spuriously high scores due to student response sets. Foils were constructed that would

seem plausible to the student with incomplete geographic knowledge. Grammatical consistency between stems and options was maintained in all questions with incomplete stems to avoid unnecessary clues to correct options. All numerical option sets were arranged from low to high, and no overlapping responses were used.

When assembled in booklet form, the test items were subjected to a series of field testings and item analyses to provide information relative to their performance prior to selection for inclusion in the final test. Drafts of the test were administered to intermediate students in Okemos, Grand Rapids, Big Rapids, and Haslett, Michigan respectively. In each preliminary testing, sufficiently large numbers of answer sheets were item analyzed to provide information on possible technical weaknesses and deficiencies among the items. Several test items were either revised, rejected, or replaced following each field testing. Illustrations were revised and clarified as sources of ambiguity, and confusion emanating from them were identified. The final version of the test contains forty-nine items and is presented as Appendix A.

Content validity.--Testing experts generally agree that quality in achievement tests is dependent upon certain essential interrelated characteristics.[28] Validity is

---

[28]See, for example, Gerberich, op. cit., pp. 24-27; Ebel, Essentials of Educational Measurement, op. cit.,

of prime importance among these characteristics. A valid test measures accurately and consistently the content and behavior the test constructor sets out to measure; obviously a test that does not perform this task is worthless as an indicator of attainment of desired learnings. There are several types of validity presented in the educational testing and measurement literature. Ebel suggests that the different varieties of validity can be grouped in two major categories: primary, or direct, validity and secondary, or derived, validity. Content, curricular, intrinsic, and face validity are examples of direct validity, which is established by rational analysis and professional judgment of the test constructor. Concurrent, predictive, empirical, and construct validity are examples of derived validity, which is established through empirical and statistical procedures.[29]

"Content validity is concerned with the adequacy of sampling of a specified universe of content."[30] If experts, knowledgeable of the content X behavioral domains specified for assessment in the test blueprint, can reach

---

pp. 359-82; Marshall and Hales, op. cit., pp. 8-11; and John A. Green, Teacher-Made Tests (New York: Harper and Row Publishers, 1963), pp. 83-100.

[29]Ebel, Essentials of Educational Measurement, op. cit., pp. 435-44.

[30]Ibid., p. 437.

a reasonable degree of agreement that an item does in fact sample a particular content X behavioral domain, then that item possesses content validity. It is of principal interest where assessment of achievement is concerned and is the type that was developed for MIGT. Content validity was established for MIGT through use of a panel of six judges with expertise in geographic education, elementary social science education, and educational testing and measurement (Appendix B). Preliminary training sessions were conducted with the judges to provide specific information on the validation process to be used, establish common understandings of the terminology used, and administer twelve practice items for classification and discussion. Use of the practice items served to insure that all judges understood what tasks they were being asked to perform and had a precise knowledge of the meanings of the content X behavioral domains into which test items were to be categorized.

Each judge was given a copy of Draft 5 of the test, and a copy of the validation instrument (Appendix C) used in the practice session, and asked to categorize each test item into one of the six cells of the test blueprint. Initial categorization of the test items by the judges utilized a blueprint with two geographic content areas--physical geography and human geography--and three levels of cognitive behavior--knowledge,

comprehension, and application (Table 2). Many of the items on Draft 5 of the test were clearly geographic skills items and could logically be placed in either one of the geographic content areas. Examination of the independent classifications on the content axis of the matrix revealed considerable inconsistencies among the categorizations of the judges. No such problem was experienced for the nonskills items or for categorization on the cognitive behavior axis. Therefore, the geographic content axis of the matrix was modified so as to achieve maximum utilization of the classification decisions.

TABLE 2.--Classification of items by geographic content and cognitive behavior

Cognitive Behavior Domains	Geographic Content Domains	
	Physical Geography	Human Geography
Knowledge		
Comprehension		
Application		

A new blueprint was devised into which the original categorizations of the judges could be placed without jeopardizing the content validation process (Table 1). One additional geographic content area was added--geographic skills--and the comprehension and application categories

on the cognitive behavior axis were collapsed in a single category--understanding. Any item requiring the student to use scale to determine distance or area, identify map features, determine location on maps through use of co-ordinates, or otherwise, or read graphs was placed in the geographic skills column of the classification blueprint given in Table 1.

Content validity was established for forty-nine of the original fifty-nine items on Draft 5 by using the judge's classifications in the modified blueprint. Validity was defined as agreement by four of the six judges as to which content domain was being sampled by a particular item; items on which at least four judges could not agree were discarded. The item classifications are listed in the appropriate content domains in Table 1.

Interobserver reliability provides an index of agreement between judges on both individual items and the test as a whole. It is determined by calculating the percentage of agreement between judges and is represented by the following equation[31]:

$$r = \frac{\text{total number of agreements}}{\text{total number of opportunities to agree}}$$

An index of 1.0 would indicate perfect agreement between opportunities to agree and actual agreement. Any item with an r of .67 was considered valid. A similar threshold was

---

[31]Gilbert Sax, Empirical Foundations of Educational Research (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968), pp. 190-91.

used by Zacur.[32] Interobserver reliability on MIGT was .82.

Reliability.--Reliability is the ability of a test to measure consistently. It is achieved when two or more measures of the same achievement are obtained and found to be positively correlated. Several methods are available for determining reliability on objective achievement tests. They include the test-retest, equivalent forms, split-halves, and Kuder-Richardson methods. The test-retest method is perhaps the simplest version. It involves two administrations of the same test to the same group of examinees and correlating student scores to determine the consistency of performance between the separate administrations. Equivalent forms of the same test can be used to obtain independent measures of achievement for use in a similar manner. The split-halves method involves correlating scores on two subtests created from a single test--one formed by the odd-numbered items and one formed by the even-numbered items.

The Kuder-Richardson Formula #20 also obtains an estimate of reliability from a single administration of a test. It calculates the proportion of correct responses for each item and multiplies by the proportion of incorrect responses for each item. These products are added for all

---

[32]Maurice M. Zacur, "A Measure and Analysis of Geography Achievement in Twenty Junior-Senior High Schools of Pennsylvania" (unpublished Ed.D. dissertation, The Pennsylvania State University, 1970).

items. The resulting sum is divided by the variance, subtracted from 1, and multiplied by the product of the number of items in the test divided by 1 minus the number of items to obtain the estimate of reliability.[33] Thus, Kuder-Richardson #20 provides an internally derived correlation coefficient which indicates the consistency of a test in measuring the achievement of similar groups of students.

The KR #20 formula,

$$r = \frac{k}{k-1}\left[1 - \frac{\Sigma pq}{\sigma^2}\right]$$

where k = number of items in the test

p = proportion of the responses correct on one item

q = proportion of the responses incorrect on one item

σ^2 = variance

was used to obtain reliability coefficients for the several drafts of MIGT. KR #20 yields a more accurate estimate of the reliability coefficient when item difficulty levels vary, as they do in MIGT. Test reliability was increased on each successive draft of the test by revising certain items to make them more appropriate and definitive, and

---

[33] Ebel, Essentials of Educational Measurement, op. cit., pp. 410-16.

through the addition of items similar in difficulty to those already on the test. The KR #20 reliability coefficient obtained for MIGT was .80.

Sampling Methodology

Selection of a Representative Sample

The population of interest in this study consists of all fifth grade students enrolled in the public schools in Michigan and a proportional stratified 1 percent sample of that population provided the data for this research. Two factors received consideration in the selection of students, classrooms, and districts for testing: type of community served by the school and student achievement. Each public school district in Michigan has been categorized by the Michigan Educational Assessment Program into one of five types of districts according to the type of community served[34]:

1. Metropolitan Core Districts
2. City Districts
3. Town Districts
4. Urban Fringe Districts
5. Rural Districts

---

[34]Michigan Educational Assessment Program, Local District and School Report: Explanatory Materials (Lansing, Mich.: Michigan Department of Education, 1973), pp. 31-32. Appendix D details the qualifying criteria for each type of district.

The number of students enrolled in each school district type is available in the Fourth Report of the Michigan Educational Assessment Program.[35] The number of fifth grade students in Michigan, approximately 170,000, was obtained from the Fourth Friday Statistics of the Michigan Department of Education.[36] The percentage of students in each district type in the state was calculated from these statistics (Table 3). This percentage figure was then multiplied by the number of fifth grade students, needed for a 1 percent sample, i.e., 1,700. This resulted in the number of students needed for testing from each type of district. The number of classrooms needed from each type of district was then estimated on the basis of twenty-seven students per class (Table 4).

Three tricounty areas containing all five community types were selected for data collection: (1) the Detroit area consisting of Macomb, Oakland and Wayne counties; (2) the Jackson area consisting of Jackson, Lenawee and Calhoun counties; and (3) the Grand Rapids area consisting of Kent, Ottawa and Allegan counties. Finally, the number

---

[35]Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972).

[36]See Michigan Educational Statistics: A Statistical Summary of Michigan Education (Lansing, Mich.: Michigan Department of Education, 1972), p. 5, for a listing of the numbers of students, K-12, in Michigan's public schools.

TABLE 3.--School enrollment by type of district--1972

Type of District	Number of Students in Michigan	Percentage of State Total in Each Type of District
Metropolitan Core	521,754	24
City	232,179	11
Town	281,166	13
Urban Fringe	759,377	36
Rural	344,257	16
Totals	2,138,733	100

Source: Derived from Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972).

TABLE 4.--Number of fifth grade classrooms needed for testing in each type of district

Type of District	Percentage of State Total in Each Type of District	Sample Size	Number of Students Needed From Each Type of District for Testing	Number of Classrooms Needed From Each Type of District for Testing
Metropolitan Core	24	x 1700 =	408	15
City	11	x 1700 =	187	7
Town	13	x 1700 =	221	8
Urban Fringe	36	x 1700 =	612	23
Rural	16	x 1700 =	272	10
Totals	100		1,700	63

Source: Derived from Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972).

of classrooms needed for testing from each type of district in each of the three tricounty areas was calculated (Table 5). These selection procedures insured proportionality of the sample across the type of district.

Basic skills composite achievement scores[37] were computed by Michigan Assessment for each fourth grade student in Michigan; fourth grade scores were used because fifth grade scores were not available. It was assumed that basic skill scores for students in fourth grade would be approximately the same as basic skill scores for the fifth grade in each type of district. Mean student scores were computed for each district and percentile ranks assigned. For purposes of this investigation, the percentile rankings were assigned to a high (66.6-100%), medium (33.3-66.5%), or low (0.33-2%) category. This ranking made possible a proportional sampling across basic skills composite achievement that would be similar to the selection made across the types of districts.

The number of school districts in each of the basic skills composite achievement categories was calculated for each of the types of districts in the state, and then the number of classrooms required for the proportional sampling was determined (Table 6). Since all but two--Ann Arbor and Bay City--of the Metropolitan Core districts in the

---

[37]These scores were derived by averaging the individual's standard score on reading, the mechanics of written English, and the mathematics tests.

TABLE 5.--Number of fifth grade classrooms needed for testing by type of district from the three selected geographical areas of lower Michigan

Type of District	Student Population	Percent of Total in Each Type of District in Each Geog. Core Area	Number of Students	Number of Classrooms
Metropolitan Core				
Detroit	273,161	86	351	13
Grand Rapids	32,599	10	40	2
Jackson	13,285	4	17	1
City				
Detroit	103,051	83	155	5
Grand Rapids	11,114	9	17	1
Jackson	9,884	8	15	1
Town				
Detroit	70,893	70	155	5
Grand Rapids	19,305	19	42	2
Jackson	10,463	10	24	1
Urban Fringe				
Detroit	482,605	86	526	20
Grand Rapids	52,502	9	55	2
Jackson	26,885	5	31	1
Rural				
Detroit	12,249	22	60	2
Grand Rapids	23,675	43	117	5
Jackson	19,617	35	95	3

Source: Derived from Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972).

TABLE 6.--Basic skills composite achievement rankings of Michigan's school districts by type of district.

Basic Skills Composite Achievement Ranking	Number of Districts	Percent in Each Category	Number of Students in Each Needed for Testing	Number of Classrooms Needed for Testing
Metropolitan Core Districts				
High	1	7	29	1
Medium	1	7	29	1
Low	13	86	350	13
City Districts				
High	12	44	82	3
Medium	9	34	62	2
Low	6	22	41	2
Town Districts				
High	36	37	82	3
Medium	36	37	82	3
Low	25	26	57	2
Urban Fringe Districts				
High	55	41	251	10
Medium	43	32	196	7
Low	36	27	165	6
Rural Districts				
High	74	22	60	2
Medium	101	30	82	3
Low	160	48	131	5

Source: Derived from Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972).

state were in the low category on basic skills composite achievement, selection of geographical areas which excluded them seemed justifiable.

Individual districts conforming to the specifications established above were selected from the three geographical areas. Several districts from each type of district for each basic skills composite achievement category qualified for selection. The method used in selecting the city districts serves as an example of how all of the districts were selected for testing (Table 7).

TABLE 7.--City districts selected for sampling

Basic Skills Composite Achievement	Geographical Areas		
	Detroit	Jackson	Grand Rapids
High	Birmingham (1 class)	Plymouth (1 class)	Grand Haven (1 class)
Medium	Wyandotte (2 classes)		
Low	Monroe (2 classes)		

Source: Derived from Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972).

Districts were contacted to request permission for testing in the desired number of classrooms. Initial contacts were made with the assistance of the Social Studies Supervisor from the Michigan Department of Education

(Appendix E). If refusals were given,[38] other qualified school districts were selected and contacted until the desired number of districts in the appropriate categories was achieved. The only exception to this procedure involved one Metropolitan Core District--Jackson--where testing arrangements could not be made and no substitute was available. Exclusion of this one classroom was not a serious liability since additional classrooms were tested in both the Detroit and Grand Rapids metropolitan cores. Table 8 lists the districts and corresponding data where permission for testing was granted.

Where individual districts had more than the desired number of fifth grade classrooms, administrative personnel were asked to select those classes most representative of their district. In most cases, selection of classrooms was dependent upon the individual classroom teacher's willingness to cooperate.

Collection of the Test Data

Once permission to test was obtained from the twenty school districts meeting the established criteria, administrative personnel in each school district were contacted to arrange for data collection. Meetings were held with cooperating teachers and administrators from each district to explain the nature of the research; pass out directions for administration (Appendix F), test booklets, scori scoring keys, and answer sheets; and provide an opportunity

[38]The refusal rate for schools contacted was 38%.

TABLE 8.--Districts selected for testing and corresponding data

Cooperating Districts	Basic Skills Composite Achievement Rank	Number of Students Needed for Testing	Number of Students Tested	Number of Classrooms Needed for Testing
Metropolitan Core Districts				
Detroit	Low	351	382	13
Grand Rapids	Low	40	96	2
Totals		408	478	15
City Districts				
Plymouth	High	27	27	1
Wyandotte	Medium	54	54	2
Grand Haven	High	27	19	1
Monroe	Low	54	59	2
Birmingham	High	27	26	1
Totals		189	185	7
Town Districts				
Lowell	Low	54	46	2
Oxford	Medium	81	79	3
Rochester	High	54	44	2
Marshall	High	27	19	1
Totals		216	188	8
Urban Fringe Districts				
E. Grand Rapids	High	54	46	2
Western	Medium	27	28	1
Southfield	High	216	170	8
Fraser	Medium	81	74	3
Madison Heights	Medium	81	80	3
Taylor	Low	162	165	6
Totals		621	563	23
Rural Districts				
Clarkston	High	54	63	2
Clinton	Medium	81	71	3
Coopersville	Low	135	141	5
Totals		270	275	10

Source: Derived from Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972).

for cooperating personnel to examine and discuss the test instrument.

Cooperating teachers were asked to administer the test as soon as possible to their fifth grade classes during their regular social studies period. Teachers rather than researchers administered the test in an effort to avoid the intrusion of an unfamiliar figure in the classroom and keep the test administration in as near normal an environment as possible. Most testing was carried out within one week of the meetings with individual district personnel. Testing began in late October, 1973 and was completed by late February, 1974.

Statement of Hypotheses

Hypothesis 1:

There is a significant difference in mean student achievement by type of district.

Hypothesis 2:

There is a significant difference in student achievement between the knowledge and understanding cognitive behavior subtests of the MIGT.

Hypothesis 3:

There are significant differences in mean student achievement among the geographic content subtests of the MIGT.

Hypothesis 4:

There are significant differences among the means of the geographic content X cognitive behavior subtests of the MIGT.

CHAPTER IV

TEST RESULTS AND ANALYSIS

Introduction

Results obtained on MIGT are reported and analyzed according to the following format. First, raw score data are reported and examined to elucidate overall performance of the test group and secure general impressions of the student's understandings of geography as it is defined on MIGT. Second, item analysis data are examined to obtain information on the operation of the test as a whole; this procedure substantiates that the MIGT is functioning appropriately as a measuring device and lends credence to the results reported. Third, student achievement for the various subtests of MIGT is reported and analyzed. This analysis reveals significant differentials in student performance among the specific subtests and identifies geographic content and cognitive behavior areas where understanding of geography is relatively strong or weak. Finally, student achievement on MIGT is reported and analyzed by type of school district. Analysis performed on the data in this form identifies district types where

student's understandings of geographic facts, concepts and principles differ significantly from other district types. Statistical tests utilized for the subtest and type of district analysis include simple analysis of variance, two-way repeated measures analysis of variance and a test for multiple comparison of means.

Scoring and analysis were performed on 1,689 of the 1,810 answer sheets collected. One hundred and twenty-one student answer sheets were eliminated from the scoring and analysis because they had less than 85 percent of the items answered. These "incompletes" presumably represent the slower students and/or students whose reading abilities are not at the intermediate level. Their elimination probably caused sample means to be slightly higher than might be otherwise expected.

Raw Score Results

Raw scores ranged from 4 to 49 items correct with a mean of 20.66, a standard deviation of 7.17, and a variance of 51.47 (Table 9). Nine hundred and fifteen students, or 54 percent, answered fewer than 21 items correctly; 774, or 46 percent, answered 21 or more items correctly. Only 42 students answered more than three-fourths of the test items correctly. Two-thirds of the scores fell between 13 and 28 items correct, and all but 62 of the students sampled were in the 7 to 35 range of items correct.

TABLE 9.--Scores achieved on MIGT

Raw Score	Frequency	Cumulative Frequency	Percentile Rank	Standard Score*
49	1	1	99	89.5
43	2	3	99	81.1
42	2	5	99	79.7
41	2	7	99	78.3
40	3	10	99	76.9
39	10	20	99	75.5
38	12	32	98	74.1
37	10	42	97	72.7
36	17	59	97	71.3
35	18	77	95	70.0
34	20	97	94	68.6
33	21	118	93	67.2
32	39	157	91	65.8
31	28	185	89	64.4
30	41	226	87	63.0
29	42	268	85	61.6
28	45	313	82	60.2
27	41	354	80	58.8
26	50	404	77	57.4
25	57	461	74	56.0
24	79	540	70	54.6
23	67	607	66	53.2
22	75	682	61	51.8
21	92	774	56	50.4
20	80	854	51	49.0
19	98	952	46	47.6
18	87	1039	41	46.2
17	84	1123	35	44.8
16	104	1227	30	43.5
15	105	1332	24	42.1
14	94	1426	18	40.7
13	71	1497	13	39.3
12	75	1572	9	37.9
11	43	1615	5	36.5
10	38	1653	3	35.1
9	18	1671	1	33.7
8	10	1681	0	32.3
7	5	1686	0	30.9
6	2	1688	0	29.5
4	1	1689	0	26.7

Mean = 20.66 S.D. = 7.17 Variance = 51.47

*A standard score has a mean of 50 and S.D. of 10.

These raw score distributions indicate that understandings of the geographic facts, concepts, and principles on MIGT are not well developed in Michigan's fifth grade population. Approximately one-eighth of the test group scored 12 or less items correctly. These scores could easily be accounted for by guessing alone.

Item Analysis Results

Item analysis provides useful indicators of the quality of an achievement test, particularly the indices of difficulty and discrimination. The index of difficulty is computed in the item analysis for each test item and is expressed simply as the proportion of the examinees that miss any particular item. Indices of difficulty for each item on MIGT are given in Table 10. The indices ranged from 27 to 86, with a mean index of 58. Ninety percent of the items have indices within the 25-75 difficulty range. Items in this difficulty range maintain at least three-fourths of their maximum discriminating potential and function adequately as items of medium difficulty.[1] The remaining 10 percent of the items have indices higher than 75. This large percentage of items in the accepted difficulty range lends credence to the adequacy of MIGT in discriminating between students with adequate and inadequate understandings of geography. However, more than 86 percent of the items on MIGT have indices above the

---

[1]Robert L. Ebel, Essentials of Educational Measurement (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972), p. 390.

TABLE 10.--Indices of difficulty and discrimination for test items on MIGT

Test Item	Index of Difficulty	Index of Discrimination
1	53	37
2	47	55
3	75	16
4	38	48
5	65	33
6	47	51
7	34	47
8	27	42
9	48	48
10	28	47
11	57	23
12	84	12
13	62	25
14	65	36
15	46	44
16	61	33
17	47	53
18	36	55
19	61	52
20	65	30
21	59	43
22	68	44
23	64	20
24	83	7
25	62	47
26	75	- 1
27	71	35
28	48	50
29	63	32
30	43	51
31	86	0
32	49	52
33	40	49
34	50	59
35	45	67
36	82	0
37	57	48
38	56	40
39	52	46
40	68	18
41	68	24
42	62	39
43	63	26
44	37	49
45	79	17
46	55	40
47	66	3
48	62	50
49	71	16

ideal index of 38. The high mean index and large percentage of the items above the ideal index indicate that geography, as defined on MIGT, is not well understood by Michigan's fifth graders. Greater knowledge of the facts, concepts, and principles embodied in MIGT by students should cause a mean index nearer the ideal of 38 to be achieved.

The index of discrimination provides information which is basic to evaluation of an item's potential for distinguishing students who have certain knowledge from those who do not. It is expressed as the difference between the number of correct and incorrect discriminations an item makes. Optimal item discrimination is achieved when the test group is divided into upper (27%), middle (46%), and lower (27%) groups.[2] An item makes a "correct" discrimination when it is answered correctly by a student in the upper group and incorrectly by a student in the lower group. A correct answer by a student in the lower group and an incorrect answer by a student in the upper group on an item constitute an "incorrect" discrimination. "Correct" discriminations provide bits of differential information about test items and allow the making of judgments relative to their potential for separating knowledgeable from non-knowledgeable students. The index of discrimination is a

---

[2]See Ibid., pp. 388-92 for a statistical justification of this grouping procedure.

measure of the amount of differential information that is generated by a test item and is obtained in item analysis by subtracting the proportion of the lower group responding correctly on an item from the proportion of the upper group responding correctly. Items with discrimination indices above 30 are quite acceptable and those over 40 are considered very good.[3] If students in both the upper and lower groups lack adequate knowledge on certain items, then indices of discrimination are likely to be low. Where the content of such items is considered important for the curriculum, it may be desirable to retain such items in a test.

Indices of discrimination are given for MIGT in Table 10. Mean item discrimination achieved was 35.[4] Nearly 70 percent of the items had indices of discrimination of 30 or better, and over 50 percent had indices of 40 or higher. In general, the indices of discrimination achieved on MIGT are adequate, particularly if the very difficult items are ignored. Difficulty and discrimination have a direct relationship; as item difficulty levels decrease, the percentage of correct response for the upper group will

---

[3]Ebel, op. cit., p. 399.

[4]The low discrimination of several of the very difficult items caused the mean discrimination to be lower than it would have been otherwise. However, they were retained because their content was considered important for fifth graders.

rise more rapidly than it will for the lower group, yielding higher levels of discrimination. Given the MIGT, relatively high mean index of difficulty obtained on it seems appropriate to conclude that indices of discrimination would be improved substantially if student knowledge of geography were improved. Viewed in this way, indices of discrimination on the MIGT are as high as could be expected.

Subtest Analysis

Eleven different subtests are identified in MIGT (Table 11). Analysis of the scores was conducted to determine if there were significant differences among the various subtests of MIGT.

A Fortran computer program was written to compute mean achievement for each student on each of the six geographic content X cognitive behavior subtests of MIGT from punched data provided by the Dylakor DYL-260 Routine for the IBM System 370, Model 155 computer in the Data Processing Department, Michigan State University. The proportion of the items scored correctly on each geographic content X cognitive behavior subtest by each student was punched on IBM cards for use in the subtest analysis.[5] Then mean proportions of items correct

---

[5]The unequal number of test items appearing in each geographic content X cognitive behavior subtest necessitated the use of calculated proportions in the subtest analysis.

TABLE 11.--Subtests of the Michigan Intermediate Geography Test (MIGT)

Cognitive Behavior	Geographic Content Categories			
	Physical Geography	Human Geography	Geographic Skills	
Knowledge	Physical Geography/ Knowledge Subtest	Human Geography/ Knowledge Subtest	Geographic Skills/ Knowledge Subtest	Knowledge Subtest
Understanding	Physical Geography/ Understanding Subtest	Human Geography/ Understanding Subtest	Geographic Skills/ Understanding Subtest	Understanding Subtest
	Physical Geography Subtest	Human Geography Subtest	Geographic Skills Subtest	

were calculated from the data for each of the subtests of MIGT (Table 12).

The calculated mean proportions of items correct on the knowledge and understanding subtests of MIGT are .435 and .421 respectively. A two-way repeated measures analysis of variance (ANOVA)[6] computer program was employed to test for the significance of differences in these cognitive behavior subtest means. The following null hypothesis was tested at the .01 level of significance:

Null Hypothesis 1:

There is no significant difference between achievement scores obtained on the knowledge and understanding cognitive behavior subtests of MIGT.

Table 13 presents the summary data of the two-way repeated measures ANOVA. The computed F statistic of 16.82 indicates that a significant difference in mean student achievement does exist between the cognitive behavior subtests of MIGT. The null hypothesis was rejected; apparently students were more likely to answer recall or knowledge questions than the questions requiring understanding.

The calculated mean proportions of items correct on the physical geography, human geography, and geographic

---

[6]For an explanation of this analysis of variance, see B. J. Winer, Statistical Principles in Experimental Design (New York: McGraw-Hill Book Company, Inc., 1962), chapter 7; or Fred N. Kerlinger, Foundations of Behavioral Research (2nd ed.; New York: Holt, Rinehart and Winston, Inc., 1973), chapter 21.

TABLE 12.--Mean proportions of items correct for the subtests of MIGT

Cognitive Behavior	Geographic Content Categories			
	Physical Geography	Human Geography	Geographic Skills	
Knowledge	.386	.493	.425	.435
Understanding	.306	.503	.453	.421
	.346	.498	.439	

TABLE 13.--Two-way repeated measures ANOVA for the subtests of MIGT

	Sum of Squares	Degrees of Freedom	Mean Square	Computed F Statistic	Tabled F Statistic @ .01
Grand Mean	1,854.59	1	1,854.593	----	--
Subjects	230.36	1,688	.136	----	--
Cognitive Behavior	.53	1	.527	16.82*	6.64
Geographic Content	39.80	2	19.900	587.90*	4.60
Subject by Cognitive Behavior	52.85	1,688	.031	--	--
Subject by Geographic Content	114.27	3,376	.034	--	--
Cognitive Behavior by Geographic Content	5.69	2	2.845	102.18*	4.60
Subjects by Behavior by Content	93.99	3,376	.028	--	--

*Significant at the .01 level

skills subtests of MIGT are .346, .498, and .439 respectively. Two-way repeated measures ANOVA was performed to test for the significance of differences in these geographic content subtest means. The following null hypothesis was tested at the .01 level of significance:

Null Hypothesis 2:

There are no significant differences among achievement scores obtained on the geographic content subtests of MIGT.

The computed F statistic of 587.90 for the geographic content subtests shown in Table 13 is significant at the .01 level and indicates that rejection of the null hypothesis is appropriate.

Student performances on the geographic content subtests of MIGT do differ, but which means differ significantly from the others? To answer this question, Scheffé's test for multiple comparison of means was applied and provided information relative to the significance of each geographic content mean's deviation from every other geographic content mean.[7] Table 14 provides summary information for all two-way comparisons made. Results of the comparisons indicate that achievement on the physical geography subtest differed significantly from

---

[7]For an explanation of this multiple comparison of means test, see Allen L. Edwards, Experimental Design in Psychological Research (3rd ed.; New York: Holt, Rinehart and Winston, Inc., 1968), pp. 150-53.

the means on both the human geography subtest and the geographic skills subtest. Significant also is the difference between the human geography subtest and the geographic skills subtest.

TABLE 14.--Summary of geographic content subtest comparisons for significance of differences among the means of MIGT

	Physical Geography	Human Geography	Geographic Skills
Physical Geography	--		
Human Geography	XX	--	
Geographic Skills	XX	XX	--

Since each geographic content mean was significantly different from every other geographic content mean, it is possible to make the following generalizations relative to student performance on MIGT.

1. Knowledge and understanding of human geography, as defined on MIGT, is better developed in Michigan's fifth grade population than either geographic skills or physical geography knowledge and understandings.

2. Geographic skills, as defined on MIGT, though not as well developed as student knowledge and understanding of human geography, are better developed in Michigan's fifth graders than physical geography knowledge and understanding.

3. Student knowledge and understanding of physical geography, as defined on MIGT, is the least well developed of the three geographic content areas among Michigan's fifth grade pupils.

Means were calculated for each of the geographic content X cognitive behavior subtests of MIGT and are listed in Table 12. The following null hypothesis was tested through use of the two-way repeated measures ANOVA to determine if significant differences existed among the six subtest means. Testing of the null hypothesis was carried out at the .01 level of significance.

Null Hypothesis 3:

There are no significant differences among achievement scores obtained on the geographic content X cognitive behavior subtests of MIGT.

The computed F statistic of 102.18 (Table 13) indicates that a significant difference does exist among the means of the geographic content X cognitive behavior subtests, and the null hypothesis was rejected.

It was necessary to seek answers to the following questions in order to determine which interaction contrasts among the geographic content X cognitive behavior subtest means were significant and to ascertain which differences in the subtest means contributed to the statistically significant difference revealed by the two-way repeated measures ANOVA test (Figure 1).

1. Is the difference in mean student performance between the physical geography X knowledge and physical geography X understanding subtests significantly different from the difference in mean student performance between the human geography X knowledge and human geography X understanding subtests?

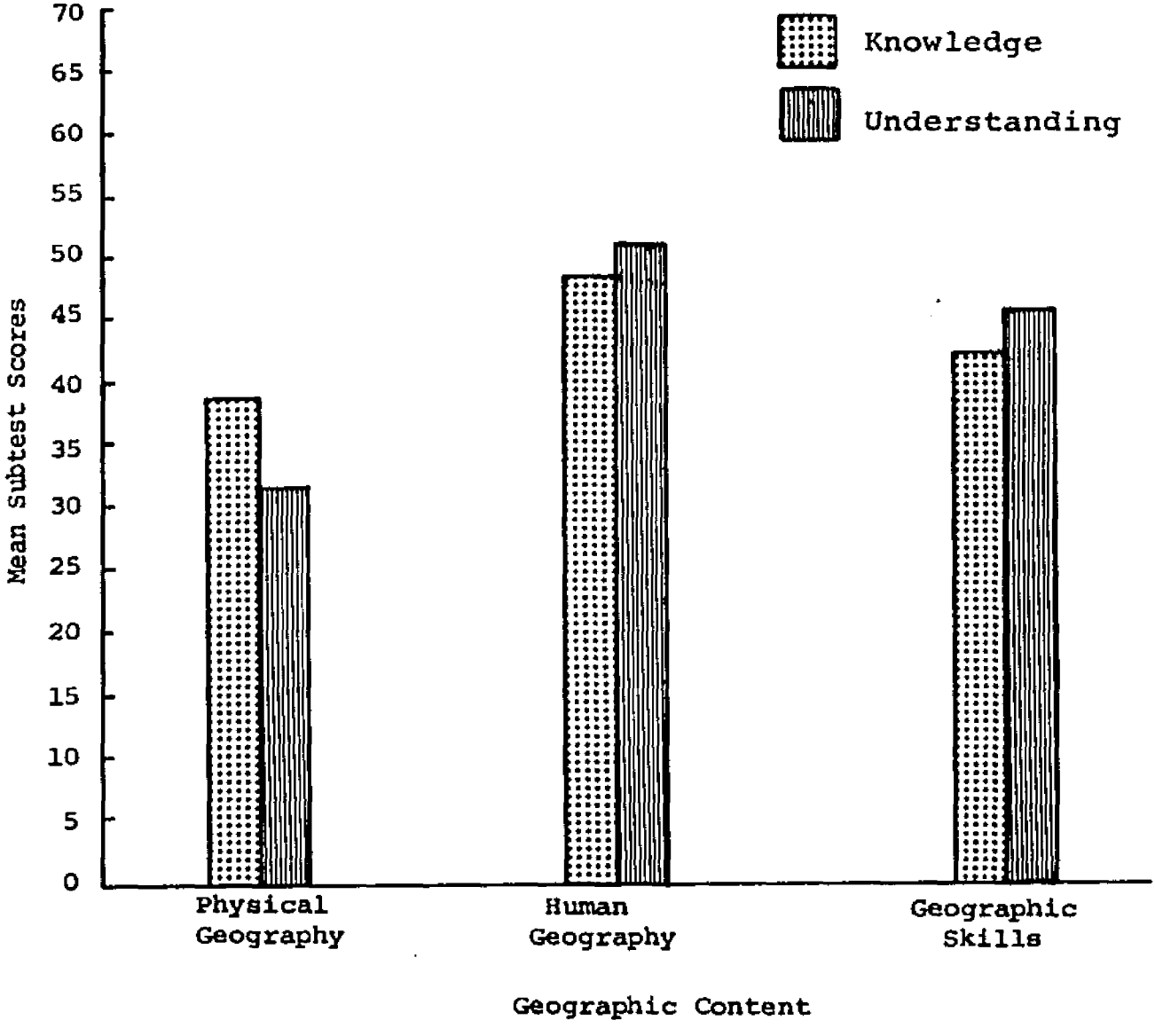


Fig. 1.--Mean scores for each geographic content X cognitive behavior subtest of MIGT.

2. Is the difference in mean student performance between the physical geography X knowledge and physical geography X understanding subtests significantly different from the difference in mean student performance between the geographic skills X knowledge and geographic skills X understanding subtests?

3. Is the difference in mean student performance between the human geography X knowledge and human geography X understanding subtests significantly different from the difference in mean student performance between the geographic skills X knowledge and geographic skills X understanding subtests?

The differences between the subtest means of each cognitive behavior category for each geographic content area was determined, e.g., the physical geography X knowledge subtest mean minus the physical geography X understanding subtest mean. Scheffé's test for multiple comparison of means was then employed to determine if the differences in these derived means were significant (Table 15).

TABLE 15.--Summary of the geographic content X cognitive behavior subtest comparisons of MIGT

	PGXK - PGXU	HGXK - HGXU	GSXK - GSXU
PGXK - PGXU	--		
HGXK - HGXU	XX	--	
GSXK - GSXU	XX	00	--

XX = Significant at the .05 level
00 = Not significant at the .05 level
PG = Physical geography
HG = Human geography
GS = Geographic skills
K = Knowledge
U = Understanding

All possible two-way comparisons of the derived means differences were made. Results revealed that the PGXK - PGXU derived mean difference is significantly different at the .05 level from the HGXK - HGXU derived mean difference. Significant also at the .05 level is the difference between the PGXK - PGXU and the GSXK - GSXU derived mean difference. One comparison, the HGXK - HGXU and the GSXK - GSXU derived mean difference, was not significant at the .05 level.

These statistical comparisons reveal that mean student performance in physical geography both at the knowledge and understanding cognitive behavior levels caused the significant difference among the six geographic content X cognitive behavior subtest means obtained in the two-way repeated measures ANOVA test.

Type of District Analysis

Analysis of the test data by type of district was performed to answer two important questions: (1) Do differences in achievement of geographic learning exist among the five types of districts in Michigan? (2) Assuming differences do exist, in what types of school districts in Michigan is learning of geography, as defined on MIGT, best developed?

Student responses for each item on the test were converted to a one for a correct response or a zero for an incorrect response through use of the DYLAKOR DYL-260

Routine for the IBM System 370, Model 155 Computer in the Data Processing Department of Michigan State University. This information for each student was then punched on IBM cards for use in analysis; a code number was used to identify cards by type of district. The mean number of items correct was calculated for each district type (Table 16). The means ranged from a high of 22.6516 for the Urban Fringe districts to a low of 17.5390 for the Metropolitan Core districts. Means for the City, Town, and Rural districts clustered around 21 items correct.

A simple one-way ANOVA program for unequal numbers of category observations was employed to test for the significance of differences in the means. The following null hypothesis was tested at the .01 level of significance:

Null Hypothesis 4:

There is no significant differences among MIGT achievement scores earned by students representing five different types of school districts.

Table 17 presents the summary data of the ANOVA. The computed F statistic of 38.433, when compared with the tabled value of F at the .01 level of significance, 3.32, indicates that a significant difference does exist among the means of the five types of districts; therefore, the null hypothesis was rejected.

Scheffé's test for multiple comparison of the means was employed to determine which means differed from

TABLE 16.--Mean numbers of items correct by type of district

District Type	Number of Correct Responses	Number of Students	Mean Number Correct
Metropolitan Core	8,524	486	17.5390
City	4,018	185	21.7189
Town	4,020	188	21.3830
Urban Fringe	12,549	554	22.6516
Rural	5,794	276	20.9928

TABLE 17.--Simple one-way ANOVA for type of districts

	Sum of Squares	Degrees of Freedom	Mean Square	Computed F Statistic	Tabled F Statistic @ .01
Total	86,873.666	1,688	--	--	--
Between	7,267.350	4	1,816.838	--	--
Within	79,606.316	1,684	47.272	38.433*	3.32

*Significant at the .01 level

the others. All possible two-way comparisons among the means were performed. The results of the comparisons are shown in Table 18.

TABLE 18.--Summary of type of district comparisons

	Metropolitan Core	City	Town	Urban Fringe	Rural
Metropolitan Core	--				
City	XX	--			
Town	XX	00	--		
Urban Fringe	XX	00	00	--	
Rural	XX	00	00	XX	--

XX = Significant at the .05 level
00 = Not significantly different at the .05 level

Student performance in geography in the Metropolitan Core districts is significantly different from performance in any of the other types of districts. Only one other two-way comparison--Urban Fringe and Rural--proved to be significant. In all other possible comparisons, no significant difference in student mean performance in geography was revealed.

Knowledge of geography as defined on MIGT is best developed among Michigan's fifth graders in the Urban Fringe school districts. Although improvements in the geography curricula of all the types of districts sampled in Michigan would probably be desirable, it is particularly

the lower achieving district types--Metropolitan Core and Rural--that stand in greatest need of assistance from geographic educators and other curriculum specialists.

CHAPTER V

CONCLUSIONS, INTERPRETATIONS, RECOMMENDATIONS, AND IMPLICATIONS

In this chapter conclusions and judgments relative to the purposes and findings of this study are given. Major findings are summarized and compared with those of previous research. Some possible interpretations and explanations of the findings are offered, and several limitations of the study are cited. In addition, suggestions for further research are made. This chapter also takes up some of the applied, or practical, implications of the study. Consideration is given to the place of geography in the elementary school curriculum as well as the responsibilities of teachers, curriculum planners and geographers for geographic education in the elementary schools of Michigan.

The Research Dimension

The major goal established for this study has been a determination of the geographic knowledge and understandings of Michigan's fifth grade students. Several important tasks were completed to achieve this goal. First, fundamental concepts, principles, skills and information thought to be important for elementary geography curricula were

established and converted into learning objectives. Second, an assessment instrument was constructed in accord with recommended and accepted principles and procedures set forth in the literature on educational testing and measurement. Third, a representative sample of Michigan's fifth grade population was selected and tested to provide the data used in the study. Finally, the test results were analyzed to determine the status of geographic learning among fifth grade students in Michigan.

Based on the findings of this research, the following generalizations seem appropriate:

The Content Generalization

> Students' knowledge and understandings are best developed in human geography, less developed in geographic skills and least developed in physical geography.

The Cognitive Behavior Generalization

> Michigan's fifth graders are more adept at answering questions depending on recall of specific information than questions requiring understanding of concepts and principles.

The Nature of the Community Generalization

> Achievement in geographic learning is highest among urban fringe students; metropolitan core students possess the least knowledge and understanding of geography.

The Sex Difference Generalization

> Boys achieved a higher mean score on the MIGT than did the girls.

The Content Generalization

Michigan students were deficient in their understanding of many aspects of physical geography. For

example, those physical geography principles associated with the directional flow of a river, the effects of differential heating and cooling of land and water bodies, and the erosional potential of running water were very difficult. Identification and understanding of certain physical terms, i.e., "plain" and "continent," were missed frequently, and large numbers failed to recognize certain land masses and larger bodies of water. Similar findings relative to physical geography have been reported by Ellis, Zimmer, and Thorp.[1]

Student performance on questions related to human geography exceeded that on physical geography questions; nonetheless, it was inadequate. For example, students evidenced little understanding of land use along the periphery of expanding cities or crowding characterizing the central portions of urban settlements. Few students associated city size and converging transportation routes with probable air, water, and noise pollution in major cities. Large numbers of fifth graders could not locate cities, institutions, or industries in Michigan. Only a few students knew the name for the group of states which includes Michigan.

---

[1]Arthur K. Ellis, "Fourth and Sixth Grade Students' Knowledge of Topographical Terms," Journal of Geography, LXVIII (September, 1969), 344-46; Sister Stephanie Zimmer, "A Diagnosis of Map Skill Deficiencies in Elementary School Children," Dissertation Abstracts, XXVIII (1967), 2107A; and Mary T. Thorp, "Studies of the Abilities of Pupils in Grades Four to Eight to Use Geographic Tools," The Teaching of Geography, Thirty-Second Yearbook of the National Society for the Study of Education (Bloomington, Ill.: Public School Publishing Company, 1933), pp. 494-506.

The research findings relative to student proficiencies in certain geographic skill areas are mixed. Some researchers have indicated that intermediate age students have the abilities to determine area, distance, and direction on maps; use scale and coordinates; and read and interpret maps and graphs.[2] Others have reported that students are deficient in their abilities to perform these geographic skills.[3] With two exceptions--use of coordinates and identification of certain map symbols--Michigan intermediate students found the geographic skill items on MIGT too difficult; hence, this research supports the latter group of studies. As a whole, Michigan fifth graders are not proficient in calculating distance or area, determining direction, using the legend, or interpreting graphs.

The Cognitive Behavior Generalization

Results of this study indicate that Michigan students performed lower level cognitive behavior tasks better than those at higher levels. A plausible explanation for this finding involves the level of difficulty of the tasks students were asked to perform. Abilities to perform

---

[2]See reviews of studies conducted by Towler and Nelson, Towler, Thorp, and Edwards in Chapter II relative to their findings on geographic skills attainment.

[3]See reviews of studies conducted by Thomas, Anderson, Thorp, and Zimmer in Chapter II relative to their findings on geographic skills attainment.

different tasks--particularly remembering versus understanding tasks--develop at different rates and times for different children.[4] Fifth grade students may be too immature to perform understanding level tasks, particularly when the task is to be performed on unfamiliar subject matter. Other variables such as differences in I.Q., aptitudes, and sex may also contribute to the explanation of these differences and could account for the relatively lower mean performance achieved on the understanding subtest of MIGT.

A more likely explanation of this differential in achievement between knowledge and understanding involves the students' exposure to questions and thinking requiring higher levels of cognition. Textbooks and teachers are two major sources of questions to which students are exposed. A number of recent investigations of the type of questions asked in classrooms revealed that the cognitive level of the majority of questions presented in textbooks is very low and that such questions "are limited generally to those that require little more than a memorized response."[5] Davis and Hunkins have examined the questions

---

[4]Millie Almy, "The Psychologist Looks at Spatial Concept Formation: Children's Concepts of Space and Time," in Research Needs in Geographic Education, ed. by Phillip Bacon and Lorrin Kennamer, Jr., Geographic Education Series No. 7 (Normal, Ill.: National Council for Geographic Education, 1967), pp. 23-40.

[5]U. S. Chaudhari, "The Role of Questions in Thinking and Learning from Text: A Research Perspective," Educational Technology, XIV (January, 1974), 7-11.

from three fifth grade social studies textbooks--one emphasizing geography--and found an overwhelming concentration on knowledge or recall to the neglect of questions requiring higher mental operations.[6] A recent review of studies of the kinds of questions asked by teachers in the classroom revealed a dominant emphasis on knowledge, or recall, type questions and a general lack of questions involving the higher mental processes.[7]

The Nature of the Community Generalization

The highest mean achievement was attained by students in the urban fringe districts; the lowest by metropolitan core students. These findings may have occurred because of the relationship between the socioeconomic characteristics of communities and educational achievement. It is well known that urban fringe students in the United States generally come from learning environments with high socioeconomic status (SES) and exhibit high educational expectations and achievement. Coleman found that achievement is

---

[6]O. L. Davis and F. P. Hunkins, "Textbook Questions: What Thinking Processes Do They Foster?" Peabody Journal of Education, XLIV (March, 1966), 285-92.

[7]Gary A. Manson, "The Effects of Immediate and Post-Poned Observer Feedback on The Acquisition of Higher Order Questioning Skills by Prospective Teachers," (unpublished Ph.D. dissertation, University of Washington, 1970), p. 7.

strongly correlated with SES and the educational aspirations of other students in a school.[8] Other studies also indicate that "there is substantial evidence leading to the conclusion that a strong connection exists in the United States between the level of educational achievement attained by students within a particular school and the socioeconomic backgrounds of their families."[9]

The results obtained on MIGT are consistent with these previous findings. Table 19 reveals, in general, a positive relationship between mean scores achieved by Michigan fifth graders on MIGT and a rank ordering based on estimates of SES given by Michigan Educational Assessment.[10]

It seems reasonable to conclude that learning has a spatial configuration. Yeates has acknowledged the existence of the spatial component of educational phenomena,[11]

---

[8]James S. Coleman, Equality of Educational Opportunity, Summary Report (Washington, D.C.: Government Printing Office, Department of Health, Education, and Welfare, 1966), pp. 21-22.

[9]See, for example, the reviews in Wilbur B. Brookover, et al., Elementary School Social Environment and School Achievement, (East Lansing, Mich.: College of Urban Development, Michigan State University, 1973), Chapter II.

[10]Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972), Section IV, Presentation of the Data.

[11]Maurice H. Yeates, "Hinterland Delimitation: A Distance Minimizing Approach," The Professional Geographer, XVI (1963), 7-10.

TABLE 19.--Mean student achievement and socioeconomic status rank by type of district

District Type	Mean Achievement Score	Rank Order Based on Socioeconomic Status
Urban Fringe	22.6516	I
City	21.7189	II
Town	21.3830	III
Rural	20.9928	IV
Metropolitan Core	17.5390	V

Source: Derived from Michigan Educational Assessment Program, Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program (Lansing, Mich.: Michigan Department of Education, 1972).

but few systematic examinations of such phenomena have been conducted by geographers. Hones and Ryba noted the absence of spatial studies of educational phenomena and called for the establishment of a "Geography of Education" within the area of social and cultural geography. They point out that a "geographical background is deemed relevant" in many studies of comparative education, but that "rarely . . . are such considerations made by a geographer and the results are of limited value or even, at times, patently misleading."[12]

[12]Gerald H. Hones and Raymond H. Ryba, "Why Not A Geography of Education?" Journal of Geography, LXXI (March, 1972), 135-39.

The Sex Difference Differentiation

While the purpose of this thesis has not been to account for individual or subgroup test score differences, the data do permit a comparison of the test scores earned by boys with the test scores earned by girls. Most research relative to differences in the achievement of boys and girls has led to the conclusion that girls excel in overall school achievement and in such subjects as English, spelling, writing and art, while boys excel in such subjects as mathematics, history, geography and science. For instance, after reviewing numerous studies of sex differences, Tyler concluded that "males are clearly superior on tests of mathematical reasoning, spatial relationships, and science. Females are superior in verbal fluency, rote memory, perceptual speed, and dexterity."[13] With respect to achievement in geography, Ellis found that boys scored significantly higher than girls on a test of topographical terms[14] and Howe found boys superior in their knowledge of directions.[15] Towler's research, however, does not support the male superiority findings of other researchers; he

---

[13]Leona E. Tyler, The Psychology of Human Differences (3rd ed.; New York: Appleton-Century-Crofts, 1965), p. 247.

[14]Ellis, op. cit., p. 345.

[15]George F. Howe, "A Study of Children's Knowledge of Directions," Journal of Geography, XXX (October, 1931), 298-304.

reported no significant correlation between sex, socio-economic status, and the development of the concept of a reference system.[16] Towler and Nelson also failed to find significant differences between boys' and girls' understanding of the concept "scale."[17]

The results relative to differences in achievement by sex obtained in this study are reported in Tables 20 and 21 which reveal that boys performed significantly better than girls on MIGT. While these differences are statistically

TABLE 20.--Comparison of student achievement by sex

	Sex	
	Male	Female
Mean score	22.00*	19.89
Standard deviation	7.72	6.44
Variance	59.63	41.56

*Significantly different at the .01 level using a two-tailed "t" distribution.

[16]John O. Towler, "The Elementary School Child's Concept of Reference Systems," Journal of Geography, LXIX (February, 1970), 91.

[17]John O. Towler and D. Nelson, "The Elementary School Child's Concept of Scale," Journal of Geography, LXVII (January, 1968), 26.

TABLE 21.--Comparison of student achievement by item and sex

Item	Percentage scoring items correct	
	Males	Females
1	53	44
2	61	46
3	27	23
4	65	60
5	36	35
6	57	51
7	68	67
8	72	75
9	56	49
10	73	72
11	47	40
12	19	12
13	41	36
14	38	33
15	56	53
16	45	34
17	57	51
18	68	61
19	46	34
20	39	31
21	47	36
22	36	29
23	36	36
24	19	15
25	43	36
26	26	25
27	34	25
28	54	52
29	40	34
30	62	53
31	16	11
32	53	49
33	63	59
34	50	50
35	56	56
36	19	16
37	46	42
38	45	44
39	51	48
40	34	32
41	33	31
42	43	36
43	35	39
44	64	66
45	23	18
46	47	44
47	32	38
48	41	37
49	32	26

significant, and theoretically interesting, the practical implications of this finding are unclear. One could question whether a difference of two points between the scores of the boys and the scores of the girls is educationally significant. The fact that the mean score on MIGT is only nine points above chance would seem to be of greater educational significance than the difference in achievement between the sexes, and suggests that greater emphasis should be placed on instruction and curriculum materials that promote learning of geography. Greater emphasis on geographic instruction of both sexes, using appropriate geographic curriculum materials, in the public schools should cause a higher mean score on MIGT.

If sex differences in achievement scores persist after appropriate instruction and materials are provided, then answers to the following questions about the significance of this finding would be warranted:

1. Are there developmental differences in males and females which promote or retard different kinds of learning, including geographic learning?

2. Does society, and therefore the school, have different expectations for boys and girls and are these expectations reflected in geographic achievement?

3. Are girls provided with geographic learning experiences, outside as well as within the school, which are different from those provided for boys? If so, does this signify that schools should provide different amounts and kinds of instruction for boys and girls?

The Applied Dimension

Several implications for geographic education in Michigan's elementary schools emerge from this study. Perhaps the most obvious implication is that students are not learning much geography, at least as it is defined on MIGT. Numerous studies have shown that students in the intermediate grades can master geographic concepts such as location, density and scale, and skills such as the use of the globe and the reading and interpretation of maps and graphs.[18] Yet the students who took MIGT experienced considerable difficulty and achieved low scores when asked to perform tasks using such concepts and skills. Why?

The nature of the curriculum in Michigan's public schools probably does not promote the learning of geography as defined on MIGT.[19] In the absence of appropriate

---

[18]See reviews of studies by Towler and Nelson, Blaut and Stea, Towler, Thorp, Howe, Rushdoony, and Carswell in Chapter II.

[19]This would seem to be a logical conclusion stemming from a survey conducted by the Board of Directors of the Michigan Council for Geographic Education in 1971-72, which revealed that geography plays only a minimal role in the curricula and textbooks found in Michigan's elementary schools.

geographic courses, textbooks and instruction, it is unlikely that Michigan's fifth graders would learn the kind of geography represented on MIGT. Those individuals in control of elementary curricula in Michigan may consider geography to be insignificant and may attach greater importance to other outcomes. The result of such emphasis may be that children are learning many things in school but they are not learning geography as it is defined in this study.

The nature of previous extra-school experiences of students may also influence the amount and kind of geography learned. For example, the amount of travel, the number and location of previous residences, and the number of exposures to "foreign" places probably contribute to geographic achievement. Conversely, students who have spent the vast majority of their lives in the immediate environs are less likely to perform well on tests such as MIGT.

The training received by Michigan's elementary social studies teachers may be another factor influencing the geographic learning of fifth grade students. Several studies have revealed the poor preparation of teachers of geography and the social studies.[20] One researcher has

---

[20]See J. R. Skretting and J. E. Sundeen, "Social Studies Education," Encyclopedia of Educational Research, ed. by Robert L. Ebel (New York: The MacMillan Co., 1969), p. 1232; James J. Veltkamp, "An Analysis of the Status of Geography Education in the Intermediate Grades in a Tri-State Regional Area," Dissertation Abstracts, XXVIII (1967), 1635A; and John H. Hansen, "The Social Studies Program of a

concluded that "the geographic preparation of teachers was significantly related to student learning with some materials."[21] Knowledge of the content in a particular discipline would seem to be one logical indicator of a teacher's potential for quality teaching in that discipline. It is doubtful whether a teacher, not well versed in the content of a subject, will devote much instructional time to the teaching of that subject. If this is the case with Michigan's elementary social studies teachers, then achievement in geography is likely to be adversely affected.

Further, if a teacher has had only limited exposure to the content of a discipline, it is not likely that he would possess a knowledge of the structural components of that discipline that would permit the teaching of concepts and principles requiring complex cognitive behaviors. And students could not be expected to perform effectively many of the tasks on MIGT.

If Michigan students' understandings of geography are being adversely affected by elementary social studies teachers inadequately trained in geography, then geographic

---

Representative Sample of Wisconsin Junior High Schools and the Preparation of Social Studies Teachers" (unpublished Ph.D. dissertation, University of Wisconsin, 1964).

[21]Dana G. Kurfman, "A Curriculum Revision Evaluation Model: Its Tryout and Revision," Dissertation Abstracts International, XXX (1969), 67A.

educators and teacher training institutions must be held partially accountable for a failure to provide adequate pre-service and in-service training in geography and curriculum materials reflecting the latest trends in learning and the discipline.

Geography can provide a broader perspective to the social studies curriculum, but as Anderson relates,

> As geography evolved from a descriptive to a physical to a social study it did not attain its rightful place [in the curriculum] because it did not possess an organizing or unifying concept or a unique body of social content. As an interpretative subject, geography has served too frequently as a basis for the study of other subjects.[22]

In an effort to remedy this situation, some geographic and social studies educators have proposed conceptual approaches for geography in the past decade.[23] However, according to Hanna, these efforts are too few and "unless a great deal more research is activated in this field, geography may continue to flounder."[24]

---

[22]Randall C. Anderson, "Geography in Secondary Education" (unpublished Ph.D. dissertation, University of Nebraska Teachers College, 1963).

[23]See Nicholas Helburn, "High School Geography and What Is Being Done About It," Social Education, XXX (1966), 631-32, 645; Charlotte Crabtree, "Inquiry Approaches to Learning Concepts and Generalizations in Social Studies," Social Education, XXX (1966), 407-11, 14; and Paul R. Hanna et al., Geography in the Teaching of Social Studies: Concepts and Skills (Boston: Houghton Mifflin Company, 1966).

[24]Ibid.

The differences in geographic learning detected by this research raise a number of interesting questions for geographic educators in Michigan.

1. What is the nature of the elementary curricula in Michigan? Are curriculum materials which promote geographic learning representative of the discipline and available for use in Michigan's elementary schools? What learnings are being promoted by existing curricula. Is there a common set of social studies goals and objectives in Michigan's elementary schools? Are any aspects of geography represented in the social studies curricula?

2. What is the extent of Michigan elementary teachers' knowledge and understanding of geography as it is defined by the discipline? How much formal training in geography has been acquired by Michigan's elementary social studies teachers? What teaching techniques and learning strategies are being employed by Michigan's elementary teachers.

3. What is the nature of extra-school experiences to which Michigan's elementary students are exposed? Do these experiences, or lack of such experiences, exert an influence on geographic learning? What factors, other than SES, can be identified to account for differences in achievement and learning in geography and other subjects?

The lack of adequate answers to these questions constitutes an obvious and major limitation of this study. Data should be collected and relationships established if the findings and conclusions of the present study are to be understood.

Irrespective of the reasons, Michigan's fifth graders do not have a well-developed knowledge and understanding of geography as it is defined on MIGT. This fact should provide a number of directives for those concerned

with the quality of elementary geography in Michigan. Additional efforts need to be made in Michigan to ascertain the extent of geographic understandings held by students at all grade levels. Geographic educators should become actively involved with teachers and curriculum specialists in the schools in the development and use of curriculum materials. Teacher training institutes, workshops, and methods courses focused on acquisition of current geographic content, teaching techniques, and learning strategies should be effected throughout the state. Perhaps the most pressing of the needs is construction, dissemination and use of curriculum materials developed with participation of geographers, learning specialists, and teachers. Future assessment studies utilizing valid and reliable test instruments should follow adoption and use of such curriculum materials. Only then will geographic educators be able to evaluate the effectiveness of the programs and materials provided to improve geographic education in the schools of Michigan.

APPENDICES

APPENDIX A

MICHIGAN INTERMEDIATE GEOGRAPHY TEST*

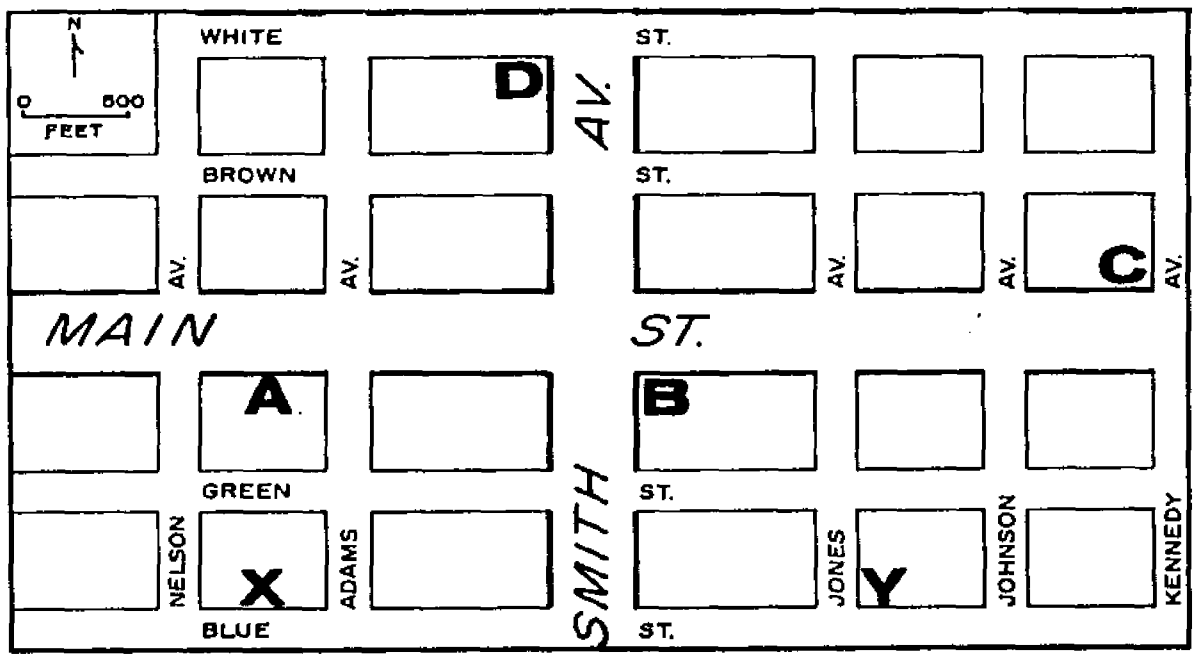
Prepared by

Gary A. Manson and Norman C. Bettis
Department of Geography
Michigan State University

*A similar version of this test has been published and copyrighted by the Michigan Council for Geographic Education.

MAP A

SMITH TOWN

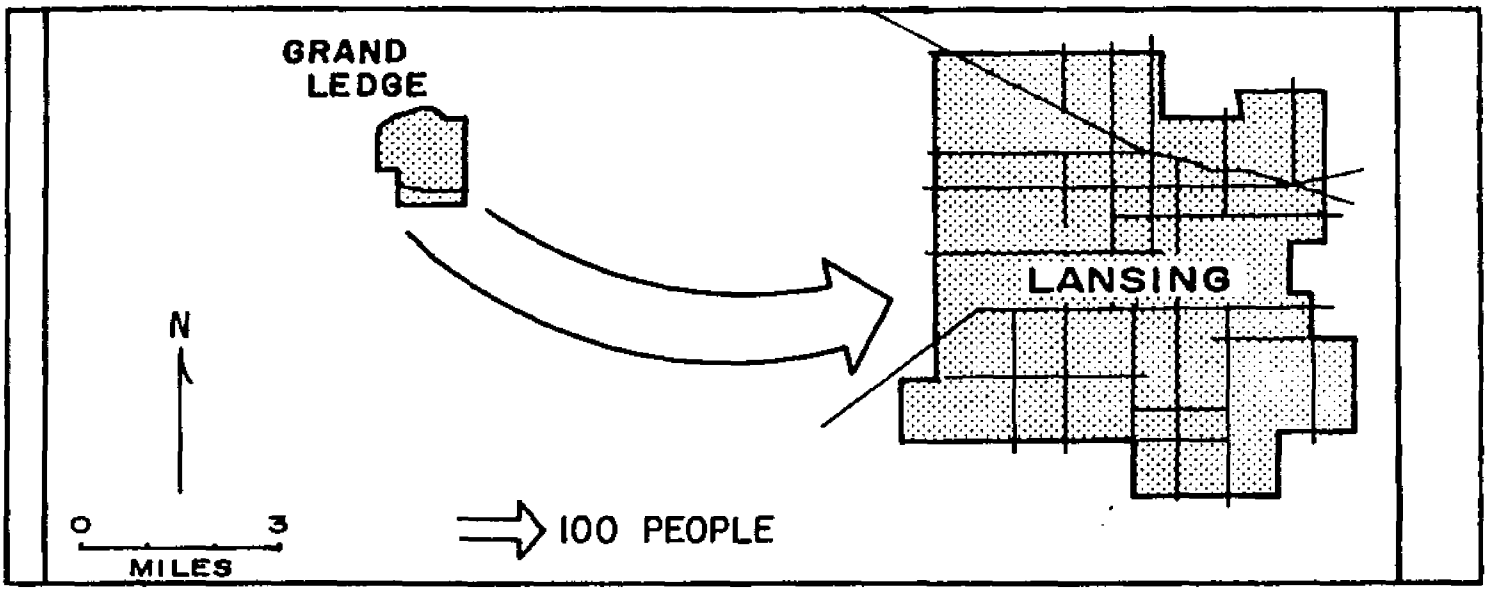


1. Look at Map A. John lives in Smith Town at Point X. He wishes to visit his friend, Bob. Bob lives at Point Y. How far must John travel to visit Bob?
   1. About 1000 feet
   2. About 3000 feet
   3. About 5000 feet
   4. About 7000 feet

2. Look at Map A. Mr. Brown wants to build a gas station. He wants to build it where many cars would drive by. Where should Mr. Brown build his gas station?
   1. Point A
   2. Point B
   3. Point C
   4. Point D

MAP B

LANSING-GRAND LEDGE



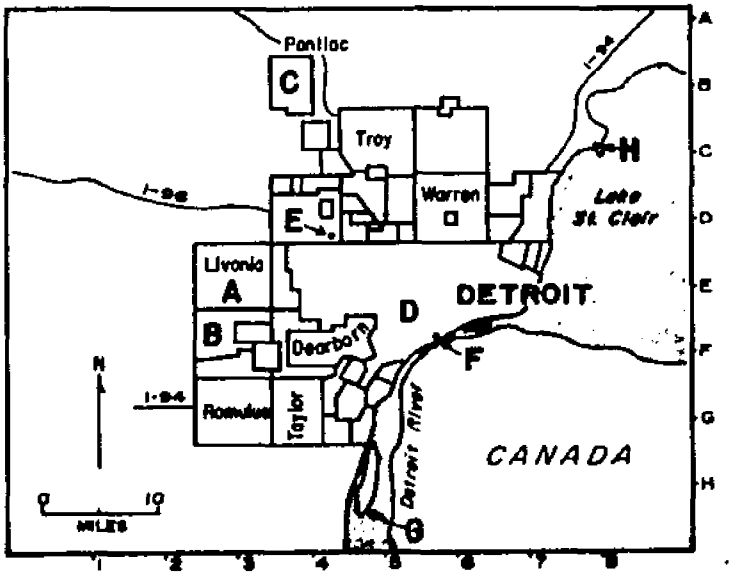
3. The arrow shows Grand Ledge people traveling to Lansing each day. About how many people travel to Lansing each day?
   1. About 100
   2. About 300
   3. About 600
   4. About 900

4. Why do so many Grand Ledge people travel to Lansing each day?
   1. They are visiting their friends and relatives in Lansing.
   2. They like to get away from Grand Ledge.
   3. They don't have any stores in Grand Ledge.
   4. They are going to their jobs in Lansing.

5. What is probably happening to the land between Grand Ledge and Lansing?
   1. It was farm land. Now people are building homes on it.
   2. People had homes on it. Now farmers are growing food on it.
   3. There were many factories there. Now they are gone and many parks are being built there.
   4. There were parks on it. Now many factories are being built there.

6. Which of these famous places is located in Lansing?
   1. Ford Motor Company
   2. State Capital
   3. Kellogg Cereals
   4. University of Michigan

MAP C
DETROIT METROPOLITAN AREA



7. Which arrow is pointing at an island?
   1. Arrow E
   2. Arrow F
   3. Arrow G
   4. Arrow H

8. What is located at 4F?
   1. Livonia
   2. Dearborn
   3. Taylor
   4. Troy

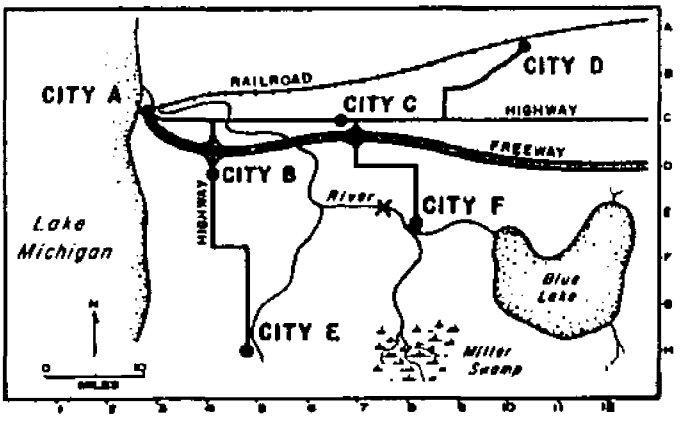
9. Warren, Dearborn, and Livonia are called_____.
   1. Suburbs
   2. Counties
   3. States
   4. Nations

10. Look at Map C. Where is pollution likely to be greatest?
   1. Area A
   2. Area B
   3. Area C
   4. Area D

11. Where on Map C are people most crowded together?
   1. Area A
   2. Area B
   3. Area C
   4. Area D

MAP D

LAKE MICHIGAN, TOWNS AND CITIES



12. John throws a stick into the River at Point X. It floats along with the river current. Which way will the stick float?
    1. Toward Blue Lake
    2. Toward City E
    3. Toward Miller Swamp
    4. Toward Lake Michigan

13. Which city probably has the greatest amount of air, water, and noise pollution?
    1. City A
    2. City B
    3. City C
    4. City D

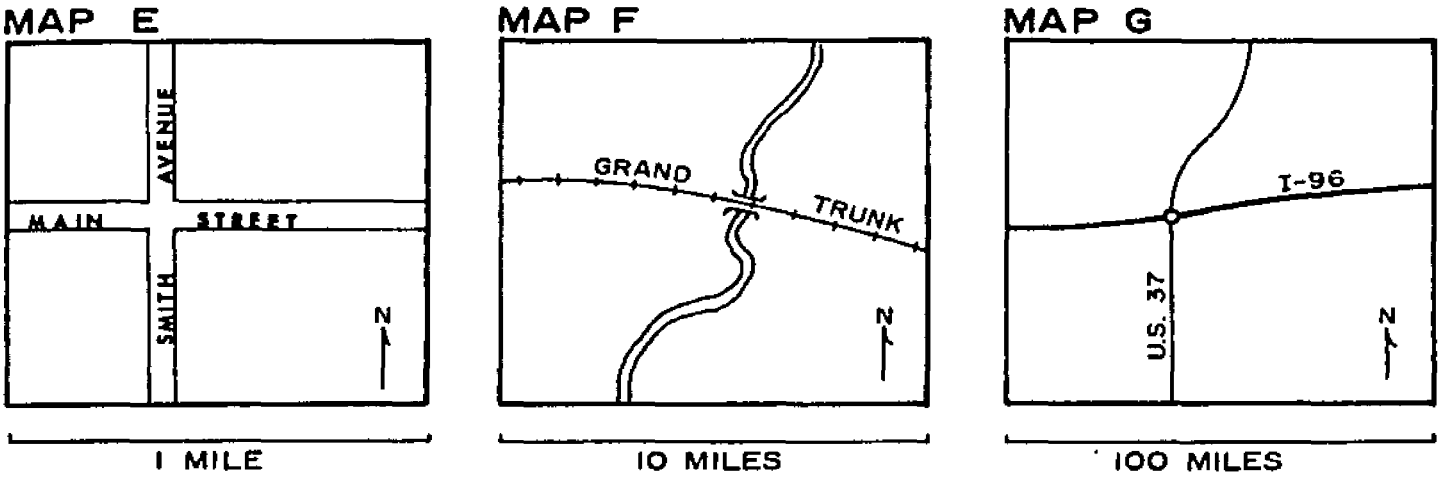
14. About how many miles is it from City D to City A by railroad?
    1. 30 miles
    2. 40 miles
    3. 50 miles
    4. 60 miles

15. City F is located nearest Point_____.
    1. 7E
    2. 7F
    3. 8E
    4. 8F

16. Many places like Miller Swamp are found in Michigan. Miller Swamp was probably formed by_____.
    1. Man
    2. Earthquake
    3. Wind
    4. Glacier

MAPS E-G

TRANSPORTATION ROUTES

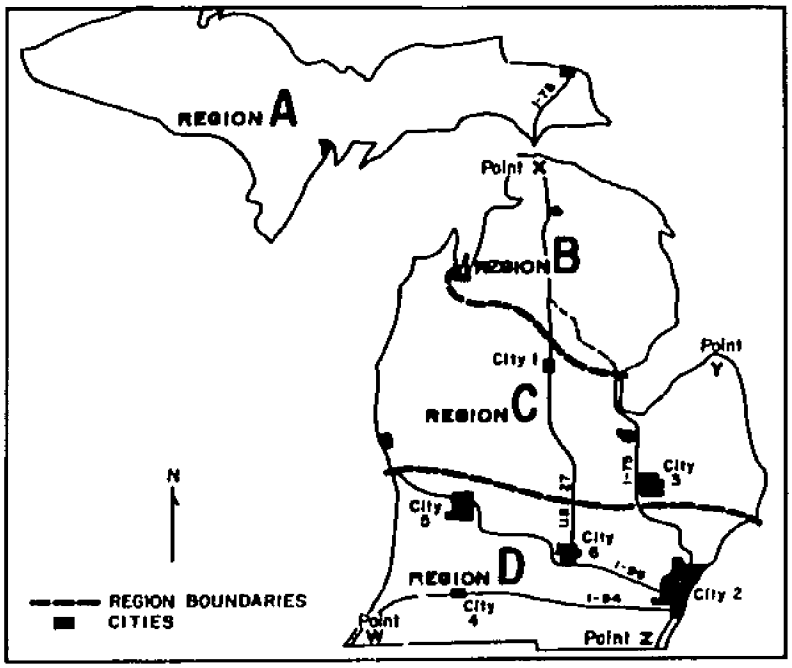


17. Here are three sketch maps. Which one shows the largest area?
    1. Map E
    2. Map F
    3. Map G
    4. All maps show the same area

18. On Map G, I-96 and U.S. 37 are_____.
    1. Pipelines
    2. Highways
    3. Rivers
    4. Streets

MAP H

MICHIGAN'S REGIONS, CITIES AND HIGHWAYS

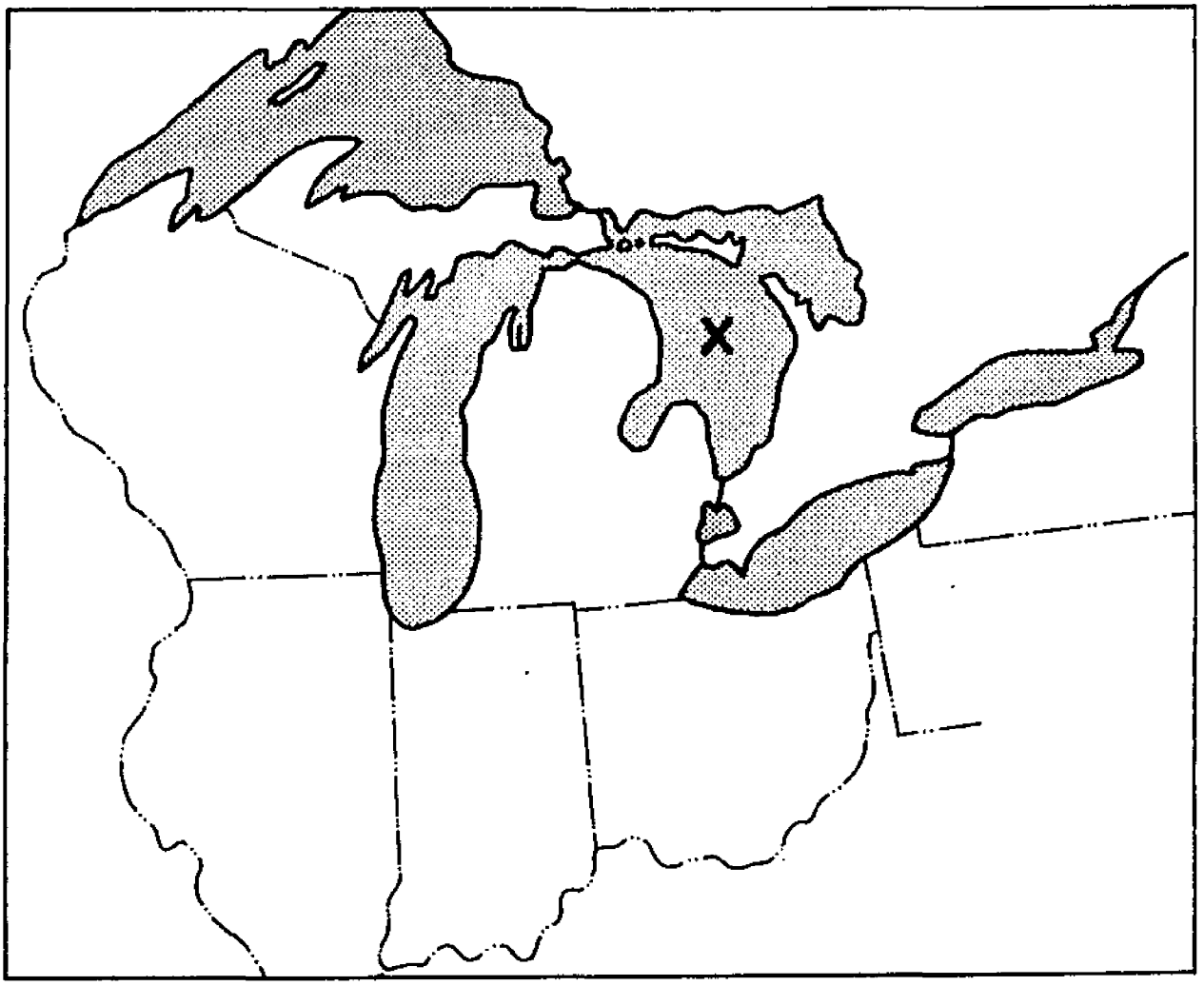


19. John lives in City 1. He wishes to travel by car to visit his grandparents in Cleveland, Ohio. If he takes the shortest route, he would leave the state at_____.
    1. Point W
    2. Point X
    3. Point Y
    4. Point Z

20. Which city is Kalamazoo?
    1. City 1
    2. City 2
    3. City 3
    4. City 4

MAP I

AROUND MICHIGAN



21. Which Great Lake is marked with an X?
    1. Lake Michigan
    2. Lake Superior
    3. Lake Huron
    4. Lake Erie

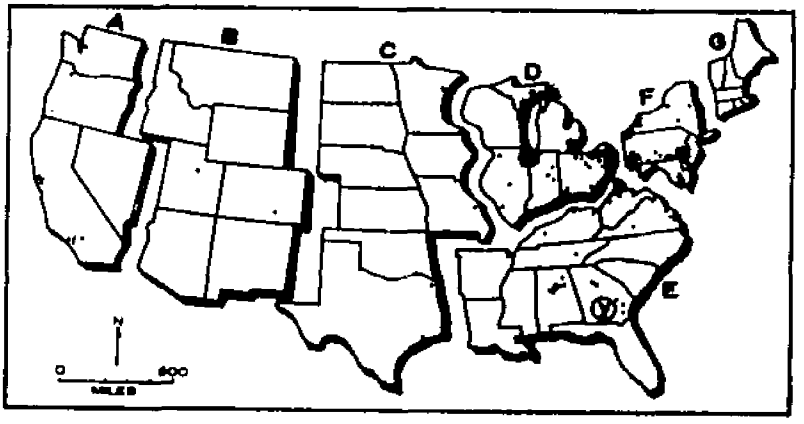
22. Which state does <u>not</u> touch Michigan?
    1. Wisconsin
    2. Illinois
    3. Indiana
    4. Ohio

23. The land surface of Michigan and surrounding states is best described as a_____.
    1. Plateau area
    2. Plain area
    3. Mountainous area
    4. Hilly area

24. How does Lake Michigan affect the temperature in Michigan in fall?
    1. Western Michigan is warmer than eastern Michigan in fall
    2. Eastern Michigan is warmer than western Michigan in fall
    3. The upper peninsula is warmer than the lower peninsula in fall
    4. Lake Michigan does not affect the temperature in Michigan in fall

MAP J

THE UNITED STATES



25. Where are the Rocky Mountains?
    1. Region A
    2. Region B
    3. Region C
    4. Region D

26. In which region is the state of Michigan?
    1. New England
    2. Northeast
    3. Midwest
    4. Southeast

27. Each dot on Map J shows a kind of activity. What kind of activity do you think it is?
    1. Tool manufacturing
    2. Cattle ranching
    3. Snow skiing
    4. Oil wells

28. Which of these ideas applies to the place marked with a Y?
    1. Long hot summers
    2. High mountains
    3. Little rain
    4. Heavy snow

29. Choose the best description of Region G.
    1. Plains; mild wet winters and hot wet summers; cotton growing
    2. High mountains; cold dry winters and cool dry summers; gold mining
    3. Hills and mountains; cold wet winters and cool wet summers; lobster fishing
    4. Plains; cold wet winters and hot wet summers; corn farming

MAP K

THE WORLD

30. Here is a map of part of the world. Where do you live?
    1. Place 1
    2. Place 2
    3. Place 3
    4. Place 4

31. Where is the North Pole?
    1. Point W
    2. Point X
    3. Point Y
    4. Point Z

32. Places 1, 2, 3, and 4 are_____.
    1. Continents
    2. Countries
    3. Cities
    4. Oceans

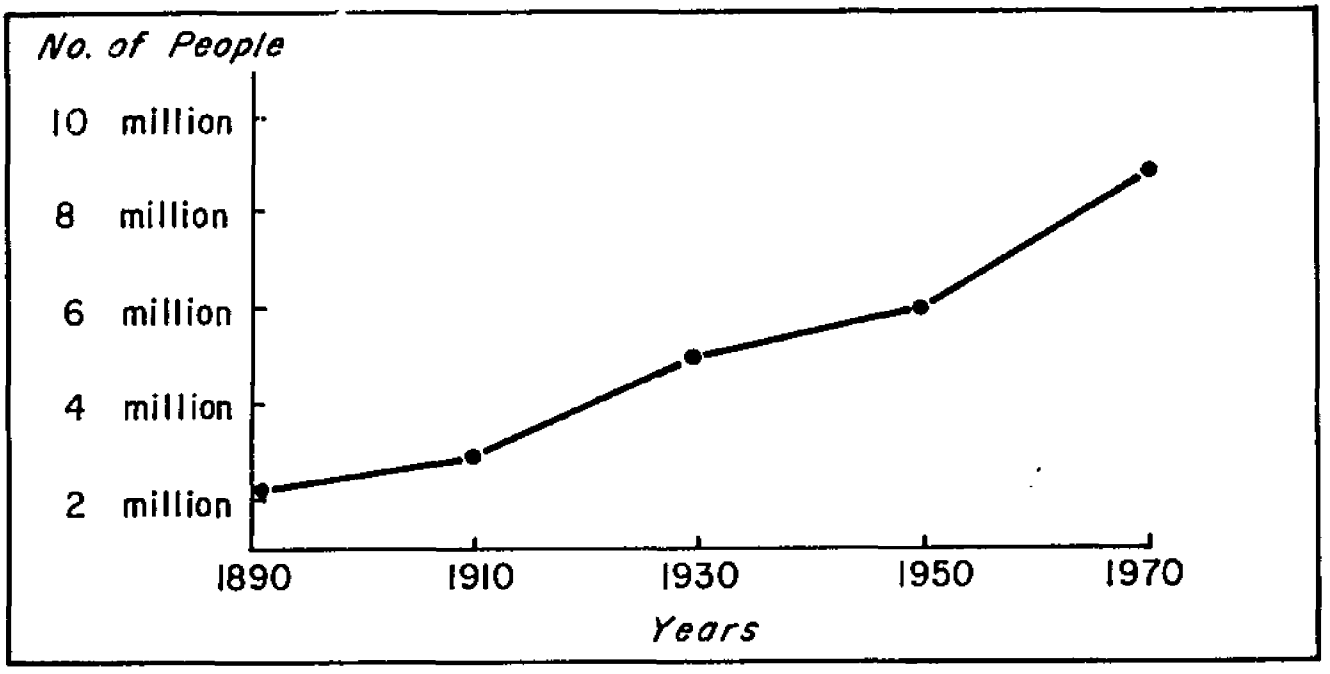
33. The Pacific Ocean is found at_____.
    1. Place 5
    2. Place 6
    3. Place 7
    4. Place 8

34. John's family is planning a trip around the world. They will visit London, Rome, Calcutta, and Tokyo. John would like to know where these cities are. Where would be the best place to look?
    1. An atlas
    2. An encyclopedia
    3. A dictionary
    4. A travel magazine

FIGURE 1

MICHIGAN'S POPULATION

GROWTH 1890-1970



35. The figure above shows how many people lived in Michigan in 1890, 1910, 1930, 1950, and 1970. How many people lived in Michigan in 1950?
    1. About 4 million
    2. About 6 million
    3. About 8 million
    4. About 10 million

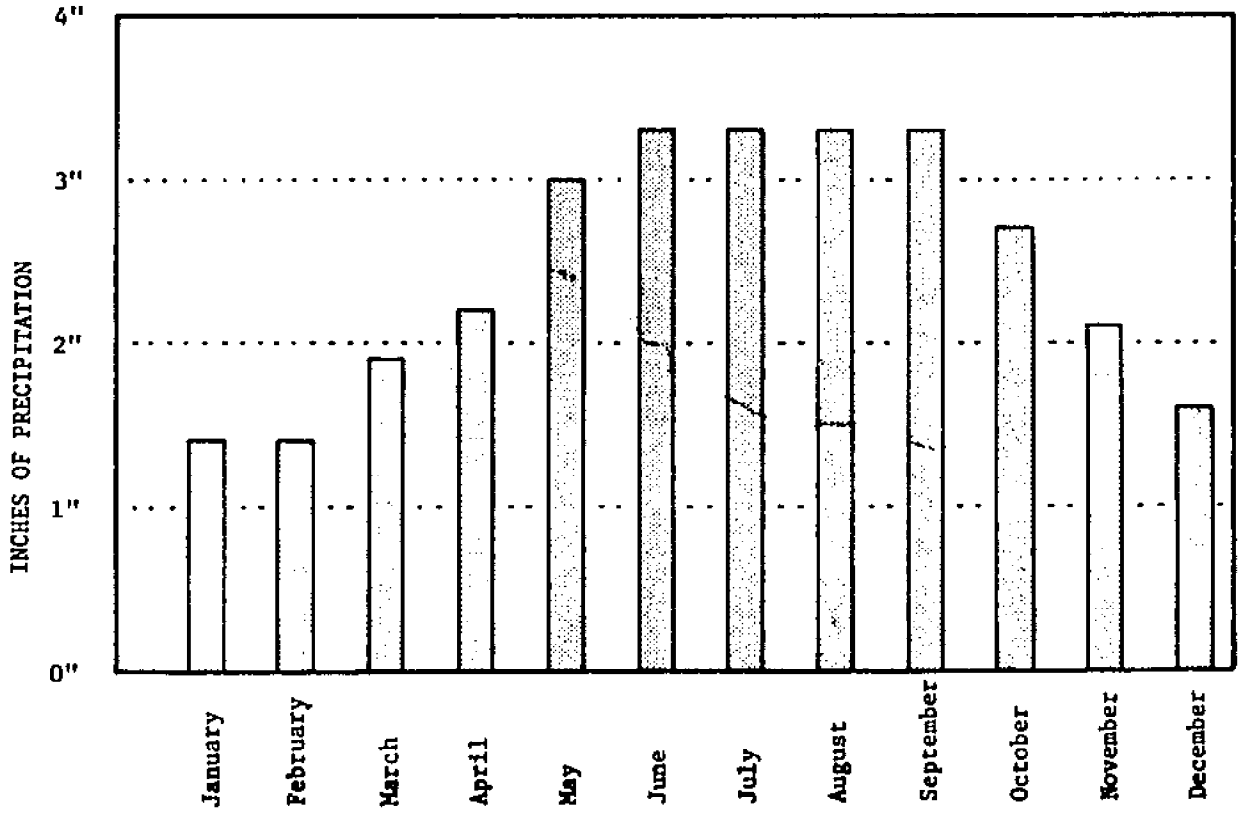
36. When did the population of Michigan grow by about 2 million people?
    1. Between 1890 and 1910
    2. Between 1910 and 1930
    3. Between 1930 and 1950
    4. Between 1950 and 1970

37. The figure above is called_____.
    1. A graph
    2. A drawing
    3. A diagram
    4. A table

38. What was probably true of Michigan's population in 1870?
    1. Fewer people lived in Michigan than in 1890
    2. More people lived in Michigan than in 1890
    3. The same number of people lived in Michigan as in 1890
    4. No people lived in Michigan in 1870

FIGURE 2

PRECIPITATION IN ESCANABA, MICHIGAN



39. Which season has the most precipitation?
    1. Summer
    2. Fall
    3. Winter
    4. Spring

40. About how many inches of precipitation falls in Escanaba in one year?
    1. About 10 inches
    2. About 20 inches
    3. About 30 inches
    4. About 40 inches

41. Which of these states has much less precipitation than Michigan?
    1. Mississippi
    2. South Carolina
    3. New Mexico
    4. Florida

FIGURE 3

PLACES IN THE U.S.

Picture 1 – Forest

Picture 2 – Wheat and Grassland

Picture 3 – Farmland

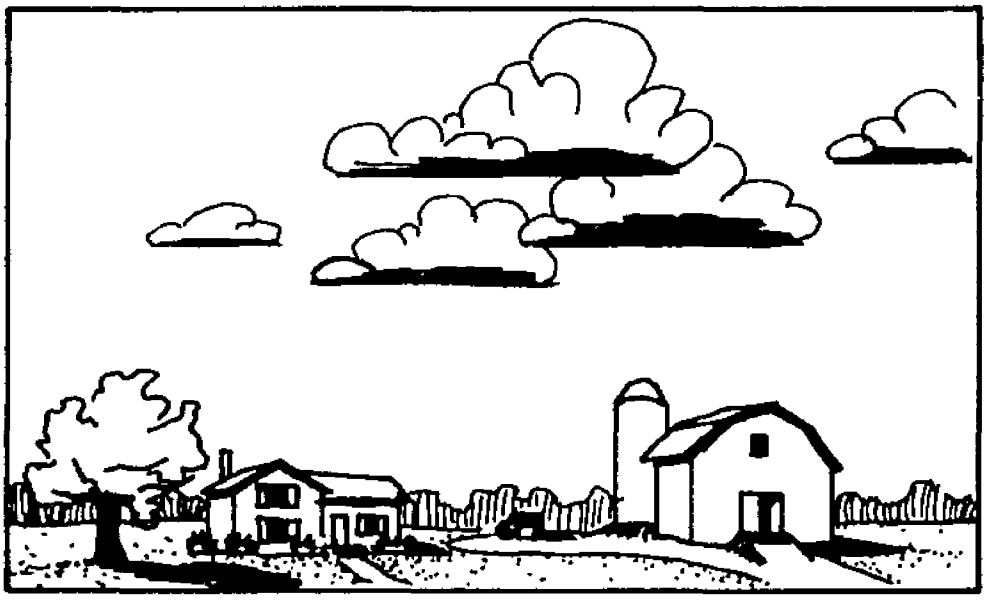
Picture 4 – House Construction Site

42. Mary took a vacation in northern Michigan and traveled all over the upper peninsula. Which picture shows what Mary probably saw most often as she traveled?
    1. Picture 1
    2. Picture 2
    3. Picture 3
    4. Picture 4

43. If it rained very hard, which land would have the least protection from the rain?
    1. Picture 1
    2. Picture 2
    3. Picture 3
    4. Picture 4

FIGURE 4

JOE'S FARM



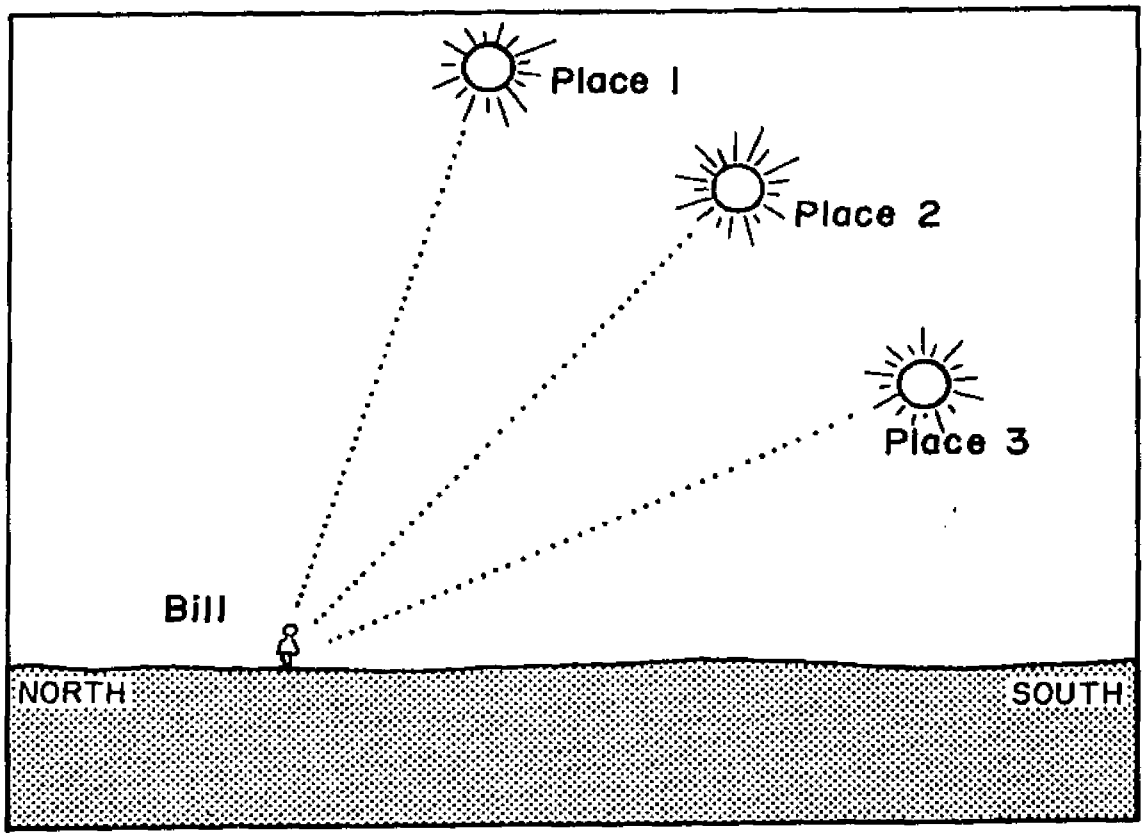
44. What crops and animals would Joe's father most likely have on his farm?
    1. Cotton, corn, and chickens
    2. Corn, rice, and sheep
    3. Wheat, tobacco, and hogs
    4. Corn, hay, and dairy cattle

45. Joe's house faces south. It is summer. When walking from the barn to the house the sun shines in his eyes. What time of day is it?
    1. Early morning
    2. Late morning
    3. Noontime
    4. Late afternoon

46. Look at the picture of Joe's farm. What kind of climate does Joe have throughout the year?
    1. High temperatures and little rainfall all year long
    2. Warm summers with rain; cold winters with snow
    3. High temperatures, much rainfall, and no snow all year long
    4. Very hot summers with much rain; cool winters with no snow

FIGURE 5

THE SUN AT NOON IN LANSING



47. Bill lives in Lansing. It is lunchtime. Where should the sun be if it is winter?
    1. Place 1
    2. Place 2
    3. Place 3

48. In springtime, if the sun shines through the east windows on your house in the early morning, which windows will it shine through in the late afternoon?
    1. North
    2. East
    3. South
    4. West

49. Bill watches the sun set from his Lansing home on Christmas Day. Where will it set?
    1. In the northeast
    2. In the southeast
    3. In the northwest
    4. In the southwest

APPENDIX B

PANEL OF JUDGES USED TO ESTABLISH CONTENT VALIDITY ON MIGT

Dr. Gordon Oosterman
Social Science Consultant and Professor of Geography
Calvin College and Grand Rapids Christian Schools
Grand Rapids, Michigan

Dr. Joseph P. Stoltman
Department of Geography
Western Michigan University
Kalamazoo, Michigan

Dr. George Vuicich
Department of Geography
Western Michigan University
Kalamazoo, Michigan

Dr. Janet Alleman
Department of Elementary and Special Education
Michigan State University
E. Lansing, Michigan

Dr. William A. Mehrens
Department of Counseling Personnel Services and Educational Psychology
Michigan State University
E. Lansing, Michigan

Dr. Robert L. Ebel
Department of Counseling Personnel Services and Educational Psychology
Michigan State University
E. Lansing, Michigan

APPENDIX C

VALIDATION OF THE TEST INSTRUMENT

> Content validity is concerned with the adequacy of sampling of a specified universe of content.[1]

Content validity is of principal interest when assessment of cognitive achievement is desired, and is the type of validity that will be established for the Michigan Intermediate Geography Test. The tasks to be undertaken involve the following: (1) defining the geographic content domain; (2) defining, in operational terms, the cognitive behavior a student must exhibit to answer test items; (3) asking experts to categorize each test item into one of the cells of the universe of content matrix; and (4) acceptance of those items on which a majority of the experts agree for use in the validated test instrument.

Two categories of geographic content are used for the Michigan Intermediate Geography Test. They are

---

[1]Robert L. Ebel, Essentials of Educational Measurement (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972), p. 437.

physical geography facts, concepts, and principles; and human geography facts, concepts, and principles (Table A.1).

TABLE A.1.--Matrix for classification of items by geographic content and cognitive behavior

Cognitive Behavior	Geographic Content	
	Physical Geography	Human Geography
Knowledge		
Comprehension		
Application		

An acceptable definition of physical geography is the study of "the physiographic-biotic system as determined by natural and cultural processes, with emphasis on the spatial distribution of the system."[2] The major criterion for classifying a question in the physical geography domain should be the geographic content of the question. If the major content emphasis has to do with facts, processes, or systems (of which man might be a part) involving (1) earth form, (2) direction, (3) area,

[2]Committee on Geography, The Science of Geography (Washington, D.C.: National Academy of Sciences, National Research Council, 1967), p. 2.

(4) distance between physical features, (5) location--or areal relationship of physical features, (6) earth-sun relations, (7) the atmosphere, (8) climate, (9) soils, (10) vegetation, or (11) landforms, then the question should be listed in the physical domain.[3]

Human geography is viewed as "a social and behavioral science." As such, it is concerned with "the principles which govern human spatial behavior . . ." and requires " . . . a comprehensive knowledge of the ways man perceives, values, and uses space and places." In addition an " . . . ability to explain and predict human spatial behavior and to modify human spatial organization . . ." is considered desirable.[4] As with physical geography, the major criterion for classifying a question in the human geography domain should be the geographic content of the question. If the major content emphasis deals with facts, processes, or systems (which may include physical elements of the earth) related to (1) direction or area within, or distance between human structures; (2) location or areal association of human structures or activities; (3) transportation and movement; (4) human

---

[3]This list of content topics is commonly covered in introductory college physical geography textbooks. See, for example, Arthur Strahler, Physical Geography (3rd ed.; New York: John Wiley and Sons, Inc., 1969).

[4]Ronald Abler, John S. Adams, and Peter Gould, Spatial Organization (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1971), p. xiii.

use or misuse of resources; (5) urban settlements; or (6) population, then the question should be classified in the human geography domain.

The first three levels from the Taxonomy of Educational Objectives: Cognitive Domain are used to categorize student cognitive behavior (Table 1) for the Michigan Intermediate Geography Test. They include (1) Knowledge, (2) Comprehension, and (3) Application. Of the first level Bloom says:

> Knowledge as defined here includes those behaviors and test situations which emphasize the remembering, either by recognition or recall, of ideas, material, or phenomena. The behavior expected of a student in the recall situation is very similar to the behavior he was expected to have during the original learning situation. In the learning situation the student is expected to store in his mind certain information, and the behavior expected later is the remembering of this information.[5]

More precisely, this level includes (1) knowledge of specifics such as terminology and facts; (2) knowledge of ways and means of dealing with specifics, such as ways of organizing, studying, judging, and criticizing ideas and phenomena; and (3) knowledge of the universals and abstractions in a field such as the major ideas, schemes, and patterns by which phenomena and ideas are organized.

---

[5]Benjamin S. Bloom, ed., Taxonomy of Educational Objectives. Handbook I: Cognitive Domain (New York: David McKay Company, Inc., 1956), Part II, p. 62.

The comprehension level contains the "largest general class of intellectual abilities and skills emphasized in schools and colleges." Bloom states:

> . . . when students are confronted with a communication, they are expected to know what is being communicated and to be able to make some use of the material or ideas contained in it. The communication may be in oral or written form, in verbal or symbolic form, or, if we allow a relatively broad use of the term "communication," it may refer to material in concrete form as well as to material embodied on paper.[6]

Here the emphasis is on the understanding and use of the communication. Three types of behavior are exhibited in comprehension: (1) translation, where an individual puts a communication into other language or terms, or another form of communication; (2) interpretation, which involves the student in an identification and comprehension of the major ideas in a communication as well as an understanding of their interrelationships; and (3) extrapolation, which requires a student to make estimates or predictions relative to an understanding of the trends, tendencies, or conditions described in the communication.

The application level necessitates a higher degree of abstract thinking from the student than either level discussed thus far. In fact, application subsumes the previous two levels. Application behavior then requires both recall and comprehension. Bloom explains:

---

[6] Ibid., p. 89

> A problem in the comprehension category requires the student to know an abstraction well enough that he can correctly demonstrate its use when specifically asked to do so. "Application," however, requires a step beyond this. Given a problem new to the student, he will apply the appropriate abstraction without having to be prompted as to which abstraction is correct or without having to be shown how to use it in that situation.[7]

So a student practicing application behavior will derive correct answers to new and unfamiliar situations without previous specification of how that answer is derived.

Following are practice test items, each of which can be categorized into one cell of the geographic content X cognitive behavior matrix (Table A.1). These items should be examined by the experts and categorized in light of the definitions given above. Discussion as to the correct categorization of each practice item will clarify the process and help eliminate misunderstandings prior to independent categorization of items on the Michigan Intermediate Geography Test.

1. From the three maps showing the migration of people in the United States, where do you think the greatest number of people will be living in the year 2000?

   (1) West
   (2) Midwest
   (3) Northeast
   (4) South

---

[7] Ibid., p. 120.

2. What is the earth's mean distance from the sun?

   (1) 8,000,000 miles
   (2) 25,000,000 miles
   (3) 63,000,000 miles
   (4) 93,000,000 miles

3. The Pacific Ocean clearly affects the climate of western Washington. If there were no such large body of water and all other climatic controls remained constant, the annual temperature range would be likely to_____ and the annual precipitation would be likely to_____.

   (1) increase, decrease
   (2) increase, increase
   (3) decrease, increase
   (4) decrease, decrease

4. Generally as the population of a city increases, so does the_____.

   (1) percent of people involved in retail trade occupations
   (2) frequency people go shopping
   (3) percent of income spent on food
   (4) number of goods and services it offers

5. The process by which particles of sand and gravel are pushed and rolled along the bottom of a stream is known as_____.

   (1) corrosion
   (2) abrasion
   (3) quarrying
   (4) traction

6. The landscape illustrated in the photograph has most likely resulted from

   (1) faulting and erosion
   (2) solution erosion
   (3) folding and erosion
   (4) vulcanism and erosion

7. In which of the following areas of the Holly Bay community would you be most likely to find new factories and warehouses?

   (1) Area A
   (2) Area B
   (3) Area C
   (4) Area D

8. Which of the following statements would be true of the climate of the place shown in the graph in Figure 2?

   (1) most of the rain falls during the summer months
   (2) summer days are likely to be sunny with few clouds
   (3) convectional thunderstorms are common during summer months
   (4) temperatures of winter days are often below freezing

9. Which of the following is an accurate description of how man uses his natural resources?

   (1) Man uses his resources in various ways according to his cultural interpretation of these resources.
   (2) Man cannot use certain resources until society expresses a demand for them.
   (3) The use of resources is usually closely tied to the economics of the particular time period.
   (4) The need for certain resources depends upon the creation of a felt need.

10. Study the chart showing the increase in the number of cars being used in Washington state. From this chart, which of the following is the most reasonable conclusion that one could make?

   (1) The number of cars in the state will level out after more people arrive in the state.
   (2) The number of cars cannot be allowed to increase without control.
   (3) The number of cars in the state will require higher gasoline taxes.
   (4) The number of cars in the state will mean additional revenue for the state.

11. Examine the topographic map in Figure 6. Which of the following enclosed areas contains the largest number of mountain peaks over 8,000 feet?

   (1) Area A
   (2) Area B
   (3) Area C
   (4) Area D

12. Mr. Jones is going to locate a new steel plant in the state of Michigan. Using the map of Michigan in Figure 3, in which location will he be able to produce steel most economically?

    (1) In or near City 1
    (2) In or near City 2
    (3) In or near City 3
    (4) Equal-distance between City 2 and City 3
    (5) Between City 1 and City 4 but closer to City 4

The validation process will require the experts to respond to a single question for each item on the Michigan Intermediate Geography Test.

Into which cell of the universe of content matrix should this test item be categorized?

Please use the following code in the left margin of the test booklet supplied to categorize your items.

Geographic Content Domain

H - Human Geography

P - Physical Geography

Cognitive Behavior Domain

1 - Knowledge level

2 - Comprehension level

3 - Application level

APPENDIX D

DEFINITIONS OF THE FIVE TYPES OF SCHOOL DISTRICTS IN MICHIGAN

1. Metropolitan Core Cities:

Communities are classified as Metropolitan Core Cities if they meet at least one of the following criteria:

(a) the community is the central city of a Michigan Standard Metropolitan Statistical Area; or

(b) the community is an enclave within the central city of a Michigan Standard Metropolitan Statistical Area.

(c) the community was previously classified as a Metropolitan Core City.

Note: The U.S. Census Bureau defines the central city of a Standard Metropolitan Statistical Area. (See U.S. Department of Commerce, Statistical Abstract of the United States [Washington: Bureau of the Census, 1968], p. 2).

2. Cities:

Communities are classified as Cities if they have a population of 10,000 or more and have not been classified as a Metropolitan Core City or Urban Fringe.

3. Towns:

Communities are classified as Towns if they have a population of 2,500 to 9,999. Rural communities impacted by large military installations nearby are also classified as Towns.

4. Urban Fringe:

Communities are classified as Urban Fringe, regardless of their size, if they meet at least one of the following criteria:

(a) the mailing address of the community is a Metropolitan Core City or a City unless it is on a RFD Route; or

(b) the community is within ten miles of the center of a Metropolitan Core City; or

(c) the community is within five miles of the center of a city.

5. Rural:

Communities are classified as Rural if they have a population of less than 2,500, or if their address is an RFD Route of a Town, City, Urban Fringe, or Metropolitan Core, and they lie outside the perimeter defined above under Urban Fringe.

NOTE: No communities in Wayne County are classified rural.

These definitions of community types were established in the fall of 1971. They have been developed to make the classification as objective and consistent as possible without altering the basic principles of classification. All classifications have been made using 1970 census data and the most recent address available for each district.

APPENDIX E

LETTER OF INTRODUCTION REQUESTING THE PARTICIPATION OF THE SCHOOL DISTRICTS IN THE RESEARCH

Dear _____:

Mr. Norman C. Bettis from the Geography Department of Michigan State University is participating in the development of a needs assessment instrument in geography. Data gathered via this instrument will be used for doctoral research designed to assess geographic understandings among fifth grade students in Michigan. Since the objectives from which his test has been developed are very similar to several of the Student Minimum Performance Objectives specified by the Michigan Department of Education, we have expressed an interest in his research.

The test and research study are not intended to reflect upon the quality of instructional procedures, but rather to determine level of achievement in geography among fifth graders. Your district is but one of twenty he has selected for sampling in Michigan. Tests from all districts will be pooled previous to scoring. Students, teachers, and districts will remain anonymous.

The purpose of this letter is to solicit your cooperation and to request permission for the administering of the test to eight representative fifth grade classes from your district. The instrument contains approximately 50 multiple-choice items and requires about one hour to administer. And Mr. Bettis would need to meet for about 30 minutes with teachers in whose classrooms testing would be carried out in order to answer questions and explain procedures previous to test administration.

Please complete the enclosed card and return it as soon as possible. Mr. Bettis will contact you concerning a convenient time for testing and other details if you indicate a willingness to cooperate.

Thank you for your assistance.

Professionally yours,

Robert L. Trezise
Social Studies Specialist
Michigan Department of Education

Encl.

POSTAL CARD

_____ My school district will assist with the research described.

_____ My school district will not assist with the research described.

_____
(Name of person to contact)

_____
(Address and phone no.)

_____
(School district)

APPENDIX F

DIRECTIONS FOR ADMINISTERING THE MICHIGAN INTERMEDIATE GEOGRAPHY TEST

The examiner should become thoroughly familiar with the test materials and directions before attempting to administer this test. Deviation from the prescribed procedure may invalidate the results.

Time and Space Requirements

A period of 45 to 55 minutes is usually required for administering the test to most students. However, time is not a major concern in this test and students should be allowed whatever time they need, within reason, for completing the test. Preliminary activities, distribution of materials, and reading specific instructions should require no more than 5 or 10 minutes.

The place for testing should be the student's regular classroom. Physical conditions such as spacing of student desks well apart, lighting, ventilation, and freedom from interruption should be satisfactory for optimum test results. Students should be asked to bring a book to the testing session or have some other quiet activity planned in the event they finish ahead of others.

Testing Materials Required

The test booklets have been designed so that students must mark their answers on a separate machine scored answer sheet. Students should not write on the test booklets. Your district may wish to reuse the test with other classes at a later time. Each student will need:

1. a test booklet
2. a Michigan State University machine scored answer sheet form no. 0-7928

3. a Michigan State University scoring pencil or a No. 2 lead pencil (pencils should not be sharpened to a fine point)

The examiner will need:

1. several extra copies of the test booklet for replacement of faulty booklets
2. additional answer sheets for students who may ruin their first one
3. a supply of spare pencils for students who may need a replacement
4. a complete set of directions for administering the test

Instructions for Marking Answer Sheets

Say to students:

> You are about to take a geography test. All answers to questions on this test must be marked on the separate answer sheet I am now going to pass out. You are not to make any stray marks on your answer sheet. Write only where I instruct you to write. Now clear your desks of all materials. Separate your desks so that they are well apart.

Now pass out answer sheets and Michigan State scoring pencils.

Say to students:

> Turn your answer sheet so that the cut-off corner is at upper left.

NOTE FOR EXAMINER:

If you want to know how your students have performed on this test, you will need to use student names or insert an identification number on the answer sheet at this point. This identification is not important to me for this research.

Say to students:

> Read silently the directions for marking the blanks that appear just beneath INSTRUCTORS NAME while I read them aloud.

Now read:

> "Use pencil only. Erase completely when necessary. Mark your answers firm and clear;" Use an up and down motion when marking the boxes. Note the way in which the example box is marked. You should attempt to make all your marks like this one. If your pencil breaks, raise your hand and I will give you another one. Do not bend or fold your answer sheet.

Now call student attention to the lower right hand corner of the answer sheet.

Say to students:

> In the lower right hand corner of the answer sheet you will find a box marked SEX. Please mark the M if you are a boy, the F if you are a girl. Do this now.

Now have students reorient the answer sheet.

Say to students:

> Turn your answer sheet so that the cut-off corner is at lower left. Notice that the answer boxes are numbered across the page from left to right. You must work across the page, not down the page, when you mark your answers. Be sure that the number of the set of answer boxes is the same as the question you are answering on the test.

Place the following practice question on the chalk board before class.

> Snow would be most likely to fall in Michigan in the month of _____?
>
> 1. June
> 2. August
> 3. September
> 4. January

Say to students:

> I have placed a practice test question on the chalk board. Find question number 168 on your answer sheet. Read this practice question silently while I read it aloud. What is the number of the correct answer?

Select one student to answer the question. When the correct answer has been determined have all students mark the correct box in the set numbered 168 on their answer sheets while you demonstrate correct marking procedure on the chalk board.

Now say:

> Is there anyone who does not understand how he or she is to mark the answer sheet?

Pause for questions.

Instructions for Using Test Booklets

When you are sure everyone understands how to mark the answer sheet, pass out the test booklets.

Say to students:

> Before you begin your test please remember you are not to write on the test booklet. Now open your booklet to page 1. Each page of the test contains a map, picture drawing, or figure on the top of the page. Questions about each map, picture drawing, or figure always appear on the same page below the map, picture drawing, or figure. You will have to use the maps, picture drawings, and figures on each page of the test to find the correct answer for almost every question on the test.

Now call student attention to the type of question and their arrangement.

Say to students:

> Turn to page 2 of your test. Notice that questions are printed on the front and back of each page. Be careful not to skip a page of the test. Notice too, that each question has four suggested answers. You are to read each question carefully and decide which of the answers is best. Then fill in the numbered box on your answer sheet that is the same as the one you have chosen in the test booklet. There is only one correct answer for each question. Do you have any questions?

Pause briefly for questions.

Now say to students:

> This test is not timed but you should work as fast as you can without making mistakes. Try to answer all questions, but do not spend too much time on any one question. Read each question carefully. If you have trouble with reading, raise your hand. Turn to page 1 of the test and begin.

General In-Session Instructions for Teachers

You should circulate among the students to answer questions and to see that answer sheets are being marked properly. Assist students with reading and word pronunciation as long as your assistance does not "give away" an answer. Do Not read or interpret maps, pictures or figures for them. Many of the skills, concepts, and principles for which we are testing require student interpretation of the maps, pictures, and figures. Many of the questions are "think" questions.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abler, Ronald; Adams, John S.; and Gould, Peter. Spatial Organization. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1971.

Adams, Bruce E. "Geographic Education in the Public and Parochial Schools of a Four County Sampling of Pennsylvania." Unpublished Ed.D. dissertation, The Pennsylvania State University, 1960.

Almy, Millie. "The Psychologist Looks at Spatial Concept Formation: Children's Concepts of Space and Time." Research Needs in Geographic Education. Edited by Phillip Bacon and Lorrin Kennamer, Jr. Geographic Education Series No. 7. Normal, Ill.: National Council for Geographic Education, 1967.

Anderson, Howard R. "Testing Basic Skills in the Social Studies." Elementary School Journal, XXXVI (February, 1936), 424-35.

Anderson, Randall C. "Geography in Secondary Education." Unpublished Ph.D. dissertation, University of Nebraska Teachers College, 1963.

Anderson, Richard C. "How to Construct Achievement Tests to Assess Comprehension." Review of Educational Research, XLII (1972), 145-70.

Bacon, Phillip, and Kennamer, Lorrin Jr. Research Needs in Geographic Education, Geographic Education Series No. 7. Normal, Ill.: National Council for Geographic Education, 1967

Baker, E. V. "Diagnosing Children's Ability to Use Maps." Journal of Geography, XXXVII (September, 1938), 227-31.

Ball, John M. A Bibliography for Geographic Education. Athens: University of Georgia, Research and Development Center in Education Simulation, 1968.

Barton, Thomas Frank. "Geography Skills and Techniques." Curriculum Guide for Geographic Education. Edited by Wilhelmina Hill. Oklahoma City: Harlow Co., 1964.

Bartz, Barbara S. "Maps in the Classroom." Journal of Geography, LXIX (January, 1970), 18-24.

Blaut, J. M., and Stea, D. "Studies in Geographic Learning." Annals of the Association of American Geographers, LXI (1971), 387-93.

Bisque, Ramon E., and Heller, Robert L. Investigating the Earth: Earth Science Curriculum Project. Boston: Houghton Mifflin Co., 1968.

Blalock, Hubert M., Jr. Social Statistics. New York: McGraw-Hill Book Company, 1960.

Bloom, Benjamin S., ed. Taxonomy of Educational Objectives. Handbook I: Cognitive Domain. New York: Longmans, 1956.

Branom, Mendel E. "Objective Diagnostic Testing Illustrations from the St. Louis Area." Journal of Geography, XLIX (September, 1945), 239-45.

Brockmyer, Irene. "Testing with Pictures." Journal of Geography, L (February, 1951), 54-57.

Brookover, Wilbur B. et al. Elementary School Social Environment and School Achievement. East Lansing, Mich.: College of Urban Development, Michigan State University, 1973.

Brown, Daniel H. "Knowledge of Important Principles of Physical Geography Possessed by Selected Sixth Grade Children." Dissertation Abstracts, XXIV (June, 1964), 5072.

Buros, Oscar K., ed. Tests in Print. Highland Park, N.J.: Gryphon Press, 1961.

_____, ed. The Fourth Mental Measurement Yearbook. Highland Park, N.J.: The Gryphon Press, 1953.

_____, ed. The Fifth Mental Measurements Yearbook. Highland Park, N.J.: The Gryphon Press, 1959.

_____, ed. The Nineteen-Forty Mental Measurements Yearbook. Highland Park, N.J.: The Mental Measurements Yearbook, 1941.

Buros, Oscar K., ed. The Seventh Mental Measurements Yearbook. Highland Park, N.J.: Gryphon Press, 1972.

_____, ed. The Sixth Mental Measurement Yearbook. Highland Park, N.J.: Gryphon Press, 1965.

_____, ed. The Third Mental Measurement Yearbook. New Brunswick: Rutgers University Press, 1949.

Carnie, J. "The Development of National Concepts in Junior School Children." Perspectives in Geographical Education. Edited by John Bale, Norman Graves, and Rex Wolford. Edinburgh, Great Britain: Oliver and Boyd, 1973.

Carpenter, Anna E. "A Test on the Southern States for Intermediate Grades." Grade Teacher, LXI (March, 1944), 55, 59, 71.

Carswell, Ronald John Berry. "Topographic Map Reading Abilities of Learners in Grades Four, Five, and Six." Dissertation Abstracts, XXX (1968), 200A.

Charles, Cheryl L. "Anthropological Curriculum Study Project: Patterns in Human History." Social Education [special issue], XXXVI, No. 7 (November, 1972), 724-26.

Chaudhari, U. S. "The Role of Questions in Thinking and Learning from Text: A Research Perspective." Educational Technology, XIV (January, 1974), 7-11.

Coleman, James S. Equality of Educational Opportunity, Summary Report. Washington, D.C.: Government Printing Office, Department of Health, Education, and Welfare, 1966.

Committee on Geography. The Science of Geography. Washington, D.C.: National Academy of Sciences, National Research Council, 1967.

Crabtree, Charlotte. "Inquiry Approaches to Learning Concepts and Generalizations in Social Studies." Social Education, XXX (1966), 407-11, 14.

Cronbach, L. J., and Warrington, W. G. "Efficiency of Multiple-Choice Tests as a Function of Spread of Item Difficulties." Psychometrika, XVII (1952), 127-47.

Davis, O. L., and Hunkins, F. P. "Textbook Questions: What Thinking Processes Do They Foster?" Peabody Journal of Education, XLIV (March, 1966), 285-92.

Davis, William Morris. Geographical Essays. Edited by Douglas Wilson Johnson. New York: Dover Publications, Inc., 1954.

Douglass, Malcolm P. "Interrelationships Between Man and the Natural Environment for Use in the Geographic Strand of the Social Studies Curriculum." Unpublished Ed.D. dissertation, Stanford University, 1954.

Ebel, Robert L. "Blind Guessing on Objective Achievement Tests." Journal of Educational Measurement, V (1968), 321-25.

_____. Essentials of Educational Measurement. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972.

Edwards, Allen L. Experimental Design in Psychological Research. 3rd ed. New York: Holt, Rinehart and Winston, Inc., 1968.

Edwards, John H. "How Well Are Intermediate Children Oriented in Space." Journal of Geography, LII (April, 1953), 133-43.

Ellis, A. K. "Fourth and Sixth Grade Students' Knowledge of Topographical Terms." Journal of Geography, LXVIII (September, 1969), 344-46.

Forsyth, Elaine. "Map Reading: Testing Skills." Journal of Geography, XLIII (May, 1944), 168-74.

Fuller, Kenneth A. "Developing Map Reading Skills for Global Emphasis." Journal of Geography, XLII (September, 1943), 216-20.

Gabler, Robert E. "Introduction: History, Scope, and Purpose." A Handbook for Geography Teachers. Edited by Robert E. Gabler. Normal, Ill.: National Council for Geographic Education, 1966.

"Geography--Examinations, Questions, Etc., and Tests and Scales." Education Index. New York: The H. W. Wilson Company, 1932-1974.

Gerberich, J. R. Specimen Objective Test Items: A Guide to Achievement Test Construction. New York: Longmans Green, 1956.

Gould, Peter R. "The New Geography." Harper's Magazine, March, 1969, 91-100.

Green, John A. Teacher-Made Tests. New York: Harper and Row Publishers, 1963.

Gulliksen, Harold. "The Relation of Item Difficulty and Inter-Item Correlation to Test Variance and Reliability." Psychometrika, X (1945), 79-91.

Haley, Francis. "Sociological Resources for the Social Studies." Social Education [special issue], XXXVI, No. 7 (November, 1972), 765-67.

Hanna, Paul R. et al. Geography in the Teaching of Social Studies: Concepts and Skills. Boston: Houghton Mifflin Company, 1966.

Hansen, John H. "The Social Studies Program of a Representative Sample of Wisconsin Junior High Schools and the Preparation of Social Studies Teachers." Unpublished Ph.D. dissertation, University of Wisconsin, 1964.

Helburn, Nicholas. "High School Geography and What Is Being Done About It." Social Education, XXX (1966), 631-32, 645.

Hones, Gerald H., and Ryba, Raymond H. "Why Not a Geography of Education?" Journal of Geography, LXXI (March, 1972), 135-39.

Howe, George F. "A Study of Children's Knowledge of Directions." Journal of Geography, XXX (October, 1931), 298-304.

_____. "A Study of the Ability of Elementary-School Pupils to Read Maps." The Teaching of Geography. Thirty-Second Yearbook of the National Society for the Study of Education. Bloomington, Ill.: Public School Publishing Company, 1933.

Hromyk, William John. "An Evaluation of High School Senior's Knowledge of Anglo-American Geographic Concepts." Unpublished Ph.D. dissertation, University of Oregon, 1972.

James, Preston E. "The Hard Core of Geography." New Viewpoints in Geography, Twenty-ninth Yearbook of the Nation Council For the Social Studies. Edited by Preston E. James. Washington, D.C.: National Council For the Social Studies, 1959.

Kaltsounis, Theodore. "An Analysis of Teaching Strategies in Emerging Geography Curricula." Focus on Geography: Key Concepts and Teaching Strategies. Fortieth Yearbook of the National Council For the Social Studies. Edited by Phillip Bacon. Washington, D.C.: National Council For the Social Studies, 1970.

Kennamer, Lorrin, Jr. "Emerging Social Studies Curricula: Implications for Geography." Focus on Geography: Key Concepts and Teaching Strategies. Fortieth Yearbook of the National Council For the Social Studies. Edited by Phillip Bacon. Washington, D.C.: National Council For the Social Studies, 1970.

Kerlinger, Fred N. Foundations of Behavioral Research. 2nd. ed. New York: Holt, Rinehart and Winston, Inc., 1973.

Kohn, Clyde F. "Foreward." Geographic Approaches to Social Education. Nineteenth Yearbook of the National Council For the Social Studies. Edited by Clyde F. Kohn. Washington, D.C.: National Council For the Social Studies, 1948.

Kurfman, Dana G. "A Curriculum Revision Evaluation Model: Its Tryout and Revision." Dissertation Abstracts International, XXX (1969).

Lindquest, E. F., and Hieronymus, A. N. Iowa Test of Basic Skills: Multi-Level Edition for Grades 3-9 (Form 4). Boston: Houghton Mifflin Company, 1964.

Lord, F. M. "The Relation of the Reliability of Multiple-Choice Tests tothe Distribution of Item Difficulties." Psychometrika, XVII (1952), 181-94.

Magnusson, David. Test Theory. Reading, Mass.: Addison-Wesley Publishing Company, 1967.

Manson, Gary A. "The Effects of Immediate and Posponed Observer Feedback on the Acquisition of Higher Order Questioning Skills by Prospective Teachers." Unpublished Ph.D. dissertation, University of Washington, 1970.

Marshall, Jon Clark, and Hales, Loyde Wesley. Classroom Test Construction. Reading, Mass.: Addison-Wesley Publishing Company, 1971.

Mehrens, William A., and Lehmann, Irvin J. Measurement and Evaluation in Education and Psychology. New York: Holt, Rinehart and Winston, Inc., 1973.

_____, and Lehmann, Irvin J. Standardized Tests in Education. New York: Holt, Rinehart and Winston, Inc., 1969.

Michigan Educational Assessment Program. Local District and School Report: Explanatory Materials. Lansing, Mich.: Michigan Department of Education, 1973.

_____. Local District Results: The Fourth Report of the 1971-72 Michigan Educational Assessment Program. Lansing, Mich.: Michigan Department of Education, 1972.

Michigan Educational Statistics: A Statistical Summary of Michigan Education. Lansing, Mich.: Michigan Department of Education, 1972.

Nunnally, Jum C. Educational Measurement and Evaluation. New York: McGraw Hill Book Company, 1972.

Patton, Donald J. "The Product." From Geographic Discipline to Inquiring Student: Final Report on the High School Geography Project. Edited by Donald J. Patton. Washington, D.C.: Association of American Geographers, 1970.

Peace, Barbara A. "Bibliography of Social Studies Tests." Evaluation in Social Studies. Thirty-Fifth Yearbook of the National Council For the Social Studies. Edited by Harry D. Berg. Washington, D.C.: National Council For the Social Studies, 1965.

Pelow, Randall Arthur. "Predicting Aptitude for Modern Geographic Concepts in Sixth Grade Pupils." Unpublished Ed.D. dissertation, The Pennsylvania State University, 1971.

"People Taking Part in the High School Geography Project." From Geographic Discipline to Inquiring Student: Final Report on the High School Geography Project. Edited by Donald J. Patton. Washington, D.C.: Association of American Geographers, 1970.

Pratt, Robert B. "High School Geography Project: Geography in an Urban Age." Social Education [special issue], XXXVI, No. 7(November, 1972), 750-52.

"Questions and Answers About the National Assessment of Educational Progress." National Assessment of Educational Progress. Washington, D.C.: Government Printing Office, 1970.

Rice, Marion J. "Improving Elementary Geography; Tests in Geography." Elementary School Journal, LXVI (December, 1965), 134-36.

Richardson, Marion W. "The Relation Between the Difficulty and the Differential Validity of a Test." Psychometrika, I (1936), 33-49.

Rushdoony, Haig A. "A Child's Ability to Read Maps: Summary of the Research." Journal of Geography, LXVII (April, 1968), 213-22.

_____. "Achievement in Map Reading: An Experimental Study." Elementary School Journal, LXIV (November, 1963), 70-75.

Sax, Gilbert. Empirical Foundations of Educational Research. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.

Scarfe, Neville V. "Designing the Curriculum to Develop Geographic Concepts." Journal of Geography, LII (March, 1953), 98-108.

_____. "Geography as an Autonomous Discipline in the School Curriculum." Journal of Geography, LXIII (October, 1964), 297-301.

_____. "Our Educational Environment and Geography Teaching." Journal of Geography, LIX (March, 1960), 103-12.

Skretting, J. R., and Sundeen, J. E. "Social Studies Education." Encyclopedia of Educational Research. Edited by Robert L. Ebel. New York: The MacMillan Co., 1969.

Slater, Francis Ann. "An Inquiry Into the Kind of Geography Being Learned in Grades 3-12." Unpublished Ph.D. dissertation, University of Iowa, 1969.

Social Education [special issue]. XXXVI, No. 7(November, 1972), 723-71.

"Social Studies Objectives." National Assessment of Educational Progress. Washington, D.C.: Government Printing Office, 1970.

"Social Studies Performance Objectives: Elementary Level, Working Draft." Lansing, Mich.: Michigan Department of Education, 1973. (Mimeographed.)

Strahler, Arthur. Physical Geography. 3rd. ed. New York: John Wiley and Sons, Inc., 1969.

Thomas, Katheryne Colvin. "The Ability of Children to Interpret Graphs." The Teaching of Geography. Thirty-Second Yearbook of the National Society for the Study of Education. Bloomington, Ill.: Public School Publishing Company, 1933.

Thorndike, Robert L., ed. Educational Measurement. Washington, D.C.: American Council on Education, 1971.

_____, and Hagen, E. P. Measurement and Evaluation in Psychology and Education. 3rd. ed. New York: John Wiley and Sons, Inc., 1969.

Thorp, Mary T. "Studies of the Abilities of Pupils in Grades Four to Eight to Use Geographic Tools." The Teaching of Geography, Thirty-second Yearbook of the National Society for the Study of Education. Bloomington, Ill.: Public School Publishing Company, 1933.

Thralls, Zoa A. The Teaching of Geography. New York: Appleton-Century-Crofts, Inc., 1958.

Tinkelman, Sherman N. "Planning the Objective Test." Educational Measurement. Edited by Robert L. Thorndike. Washington, D.C.: American Council on Education, 1971.

Towler, John O. "The Elementary School Child's Concept of Reference Systems." Journal of Geography, LXIX (February, 1970), 89-93.

_____, and Nelson, L. D. "The Elementary School Child's Concept of Scale." Journal of Geography, LXVII (January, 1968), 24-28.

Tyler, Leona E. The Psychology of Human Differences. 3rd ed. New York: Appleton-Century-Crofts, 1965.

Veltkamp, James J. "An Analysis of the Status of Geography Education in the Intermediate Grades in a Tri-State Regional Area." Dissertation Abstracts, XXVIII (1967).

Wagner, Louise D. "Measuring the Map-Reading Ability of Sixth-Grade Children." Elementary School Journal, LIII (February, 1953), 338-44.

Wesley, Edgar B., and Wronski, Stanley P. Teaching Social Studies in High Schools. Boston: D. C. Heath and Company, 1964.

Wesman, A. G. "Writing the Test Item." Educational Measurement. Edited by Robert L. Thorndike. Washington, D.C.: American Council on Education, 1971.

Whipple, Gertrude. "Geography in the Elementary Social Studies Program: Concepts, Generalizations, and Skills to Be Developed." New Viewpoints in Geography. Twenty-ninth Yearbook of the National Council For the Social Studies. Edited by Preston E. James. Washington, D.C.: National Council For the Social Studies, 1959.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, Inc., 1962.

Wood, D. A. Test Construction: Development and Interpretation of Achievement Tests. Columbus, Ohio: C. E. Merrill, 1960.

Wood, Susan. "An Evaluation of Achievement Tests in Geography." Evaluation in Geographic Education. The 1971 Yearbook of The National Council for Geographic Education. Edited by Dana G. Kurfman. Belmont, Calif.: Fearon Publishers, 1970.

Yeates, Maurice H. "Hinterland Delimitation: A Distance Minimizing Approach." The Professional Geographer, XVI (1963), 7-10.

Zacur, Maurice M. "A Measure and Analysis of Geography Achievement in Twenty Junior-Senior High Schools of Pennsylvania." Unpublished Ed.D. dissertation, The Pennsylvania State University, 1970.

Zimmer, Sister Stephanie. "A Diagnosis of Map Skill Deficiencies in Elementary School Children." Dissertation Abstracts, XXVIII (1967), 2107A.

Zwicker, Alma. "A Test on Italy and Sicily for Upper Intermediate and Grammar Grades." Grade Teacher, LXX (September, 1952), 88-90.