DOCUMENT RESUME

ED 038 302 24 SE 008 249

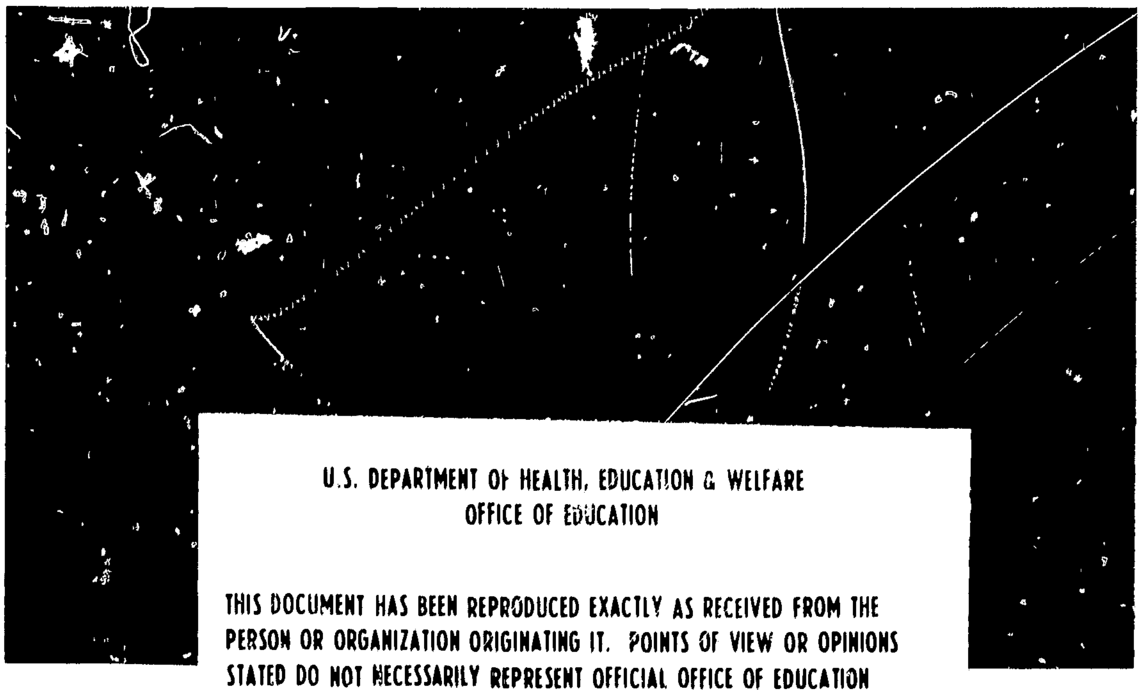
| | |
|---|---|
| AUTHOR | Shepler, Jack L. |
| TITLE | A Study of Parts of the Development of a Unit in Probability and Statistics for the Elementary School, (Part 1). |
| INSTITUTION | Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning. |
| SPONS AGENCY | Office of Education (DHEW), Washington, D.C. Bureau of Research. |
| REPORT NO | TR-105 |
| BUREAU NO | BR-5-0216 |
| PUB DATE | Nov 69 |
| CONTRACT | OEC-5-10-154 |
| NOTE | 146p. |
| EDRS PRICE | EDRS Price MF-$0.75 HC-$7.40 |
| DESCRIPTORS | *Curriculum Development, Doctoral Theses, *Elementary School Mathematics, Grade 6, *Instruction, Mathematics Education, Research, Statistics |

ABSTRACT

Reported is research conducted as a part of the Project on Analysis of Mathematics Instruction. The study had two main purposes: to test the feasibility of teaching topics in probability and statistics to a class of sixth grade students; and to construct a set of instructional materials and procedures in probability and statistics for sixth graders. A unit of instruction was prepared and the order in which behavioral objectives were to be taught was determined from a content outline and a task analysis. The results of the study support the feasibility of teaching most of the topics covered in the unit to average and above average sixth graders. The study also lends support to the use of the systems model employed for developing curriculum materials. (FL)

BR 5-021
PA-24
OE/BR
SE

ED038302

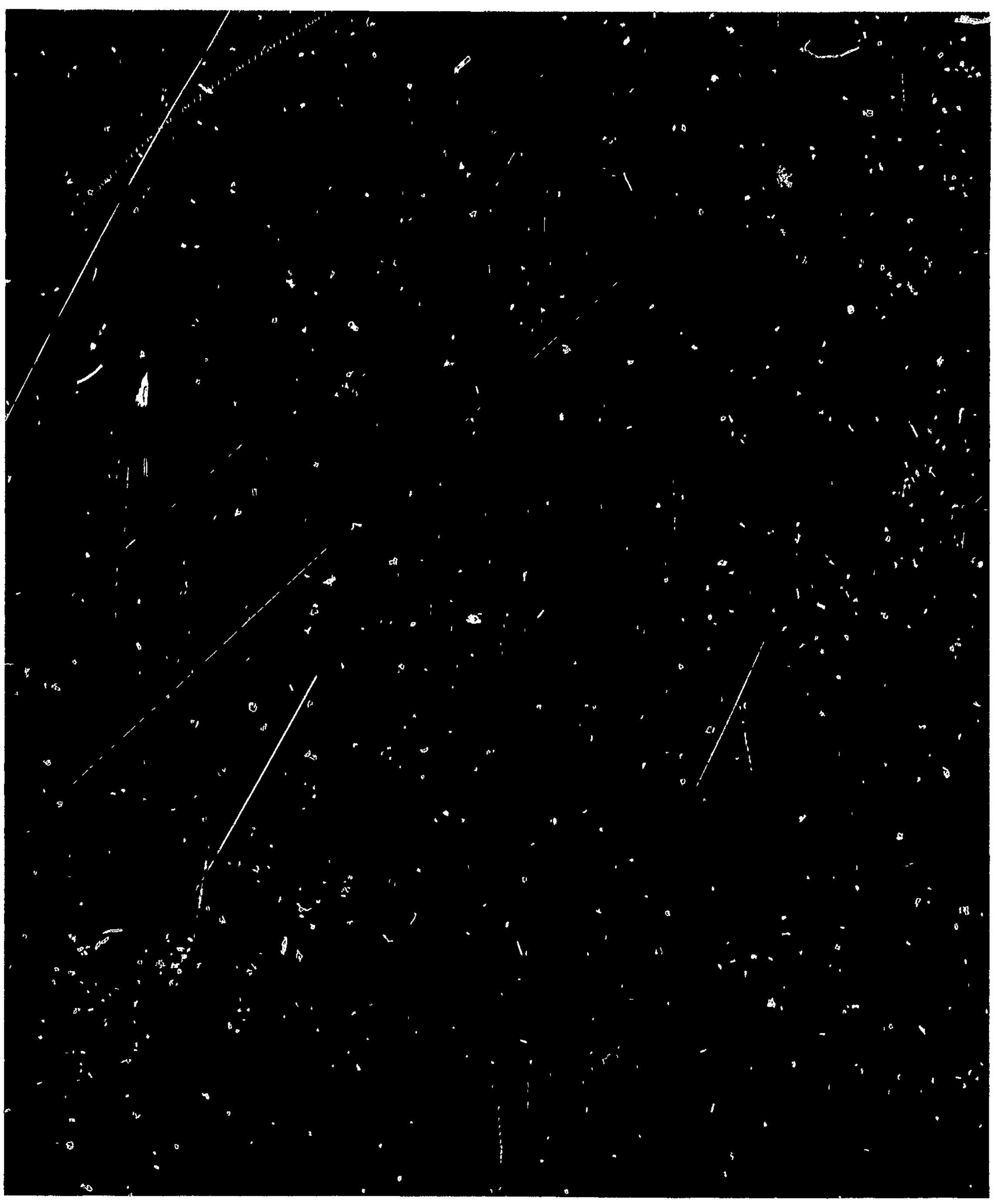




No. 105 (Part I)

# A STUDY OF PARTS OF THE DEVELOPMENT OF A UNIT IN PROBABILITY AND STATISTICS FOR THE ELEMENTARY SCHOOL

Report from the Project on
Analysis of Mathematics Instruction



SE 008 249

U.S. Office of Education Center No. C-03 Contract OE 5-10-154

Technical Report No. 105 (Part I)

# A STUDY OF PARTS OF THE DEVELOPMENT OF A UNIT IN PROBABILITY AND STATISTICS FOR THE ELEMENTARY SCHOOL

Report from the Project on
Analysis of Mathematics Instruction

By Jack L. Shepler

Thomas A. Romberg, Assistant Professor of Curriculum & Instruction
Chairman of the Examing Committee

Thomas A. Romberg, Principal Investigator

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin
November 1969

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in The University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

## NATIONAL EVALUATION COMMITTEE

**Samuel Brownell**
Professor of Urban Education
Graduate School
Yale University

**Launor F. Carter**
Senior Vice President on Technology and Development
System Development Corporation

**Francis S. Chase**
Professor
Department of Education
University of Chicago

**Henry Chauncey**
President
Educational Testing Service

**Martin Deutsch**
Director, Institute for Developmental Studies
New York Medical College

**Jack Edling**
Director, Teaching Research Division
Oregon State System of Higher Education

**Elizabeth Koontz**
President
National Education Association

**Roderick McPhee**
President
Punahou School, Honolulu

**G. Wesley Sowards**
Director, Elementary Education
Florida State University

**Patrick Suppes**
Professor
Department of Mathematics
Stanford University

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

## UNIVERSITY POLICY REVIEW BOARD

**Leonard Berkowitz**
Chairman
Department of Psychology

**Archie A. Buchmiller**
Deputy State Superintendent
Department of Public Instruction

***James W. Cleary**
Vice Chancellor for Academic Affairs

**Leon D. Epstein**
Dean
College of Letters and Science

**John Guy Fowlkes**
Director
Wisconsin Improvement Program

**Robert E. Grinder**
Chairman
Department of Educational Psychology

**H. Clifton Hutchins**
Chairman
Department of Curriculum and Instruction

**Clauston Jenkins**
Assistant Director
Coordinating Committee for Higher Education

**Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

**Donald J. McCarty**
Dean
School of Education

**Ira Sharkansky**
Associate Professor of Political Science

**Henry C. Weinlick**
Executive Secretary
Wisconsin Education Association

**M. Crawford Young**
Associate Dean
The Graduate School

## EXECUTIVE COMMITTEE

**Edgar F. Borgatta**
Brittingham Professor of Sociology

**Max R. Goodson**
Professor of Educational Policy Studies

**Russell J. Hosler**
Professor of Curriculum and Instruction and of Business

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

**Wayne Otto**
Professor of Curriculum and Instruction (Reading)

**Robert G. Petzold**
Associate Dean of the School of Education
Professor of Curriculum and Instruction and of Music

**Richard L. Venezky**
Assistant Professor of English and of Computer Sciences

## FACULTY OF PRINCIPAL INVESTIGATORS

**Ronald R. Allen**
Associate Professor of Speech and of Curriculum and Instruction

**Vernon L. Allen**
Associate Professor of Psychology
(On leave 1968-69)

**Nathan S. Blount**
Associate Professor of English and of Curriculum and Instruction

**Robert C. Calfee**
Associate Professor of Psychology

**Robert E. Davidson**
Assistant Professor of Educational Psychology

**Gary A. Davis**
Associate Professor of Educational Psychology

**M. Vere DeVault**
Professor of Curriculum and Instruction (Mathematics)

**Frank H. Farley**
Assistant Professor of Educational Psychology

**John Guy Fowlkes** (Advisor)
Professor of Educational Administration
Director of the Wisconsin Improvement Program

**Lester S. Golub**
Lecturer in Curriculum and Instruction and in English

**Max R. Goodson**
Professor of Educational Policy Studies

**Warren O. Hagstrom**
Professor of Sociology

**John G. Harvey**
Associate Professor of Mathematics and Curriculum and Instruction

**Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

**Burton W. Kreitlow**
Professor of Educational Policy Studies and of Agricultural and Extension Education

**Richard G. Morrow**
Assistant Professor of Educational Administration

**Wayne Otto**
Professor of Curriculum and Instruction (Reading)

**Milton O. Pella**
Professor of Curriculum and Instruction (Science)

**Thomas A. Romberg**
Assistant Professor of Mathematics and of Curriculum and Instruction

**Richard L. Venezky**
Assistant Professor of English and of Computer Sciences

## MANAGEMENT COUNCIL

***Herbert J. Klausmeier**
Director, R & D Center
Acting Director, Program 1

**Mary R. Quilling**
Director
Technical Section

**Thomas A. Romberg**
Director
Programs 2 and 3

**James E. Walter**
Director
Dissemination Section

**Dan G. Woolpert**
Director
Operations and Business

* COMMITTEE CHAIRMAN

## STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from Phase 2 of the Project on Prototypic Instructional Systems in Elementary Mathmatics in Program 2. General objectives of the Program are to establish rationale and strategy for developing instructional systems, to identify sequences of concepts and cognitive skills, to develop assessment procedures for those concepts and skills, to identify or develop instructional materials associated with the concepts and cognitive skills, and to generate new knowledge about instructional procedures. Contributing to the Program objectives, the Mathematics Project, Phase 1, is developing and testing a televised course in arithmetic for Grades 1-6 which provides not only a complete program of instruction for the pupils but also inservice training for teachers. Phase 2 has a long-term goal of providing an individually guided instructional program in elementary mathematics. Preliminary activities include identifying instructional objectives, student activities, teacher activities materials, and assessment procedures for integration into a total mathematics curriculum. The third phase focuses on the development of a computer system for managing individually guided instruction in mathematics and on a later extension of the system's applicability.

ACKNOWLEDGEMENTS

Many people have contributed to the completion of this study. However, I would like to thank in particular Dr. Thomas A. Romberg for the guidance and knowledge he imparted to me during my three years of graduate study. Appreciation is also extended to the other members of the reading committee, Drs. John G. Harvey and Robert E. Davidson for their help and comments concerning the study.

The author is deeply grateful to the staff of the Research and Development Center for their support in providing a teacher, technical advice, and clerical help. The author is particularly grateful to Mrs. Carolyn Gornowicz who taught the unit.

My deepest appreciation is expressed to my wife, Gloria, for her typing of rough drafts, proofreading, etc; but more importantly for her understanding and encouragement.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

## LIST OF FIGURES

LIST OF TABLES

ABSTRACT

From a content outline and a task analysis the behavioral objectives for a unit of instruction in probability and statistics for sixth-grade students and the order in which objectives would be taught were determined. An instructional analysis of the unit was undertaken to select or develop materials and procedures for teaching the unit.

Data from a pilot study conducted in the fall of 1969 were used to identify a set of nine lessons that could be formatively evaluated to test the feasibility of the instructional analysis. The lessons were used to teach a class of 25 sixth-grade students of average to above average ability. The topics developed through experiments, games and exercises were subjective probability notions, empirical probability, counting techniques, _a priori_ probability including simple and compound events, and comparison of two events using probability.

On the basis of the overall pretest and posttest the instructional treatment was generally successful. The pretest percentage was 37.9% and the posttest percentage was 92.8% with all 72 items successful for 11 of the 14 measured objectives. Instruction was unsuccessful in getting students to specify the estimated probability; number the outcomes of an event; and estimate the probability successful for these three objectives because of a lack of stress and practice. Two learning hierarchies were also tested. One hierarchy was validated and the other was not. The results of the study support the feasibility of teaching most of the included topics in probability and statistics to average and above average sixth-grade students given high quality of teaching. The study lends support to the use of the systems developmental model employed in this study for developing curriculum materials for the schools, especially when used in conjuction with Bloom's "Mastery Learning" techniques.

## Chapter I

## THE PROBLEM: CAN A UNIT IN PROBABILITY BE CONSTRUCTED AND SUCCESSFULLY TAUGHT TO SIXTH GRADE STUDENTS?

His technological society has inundated man with probabilistic and statistical statements, materials and decisions. But has he been prepared for such a fusillade? Practically no provision has been made for giving elementary and secondary school students formal instruction in probability and statistics. No reliable evidence exists to show whether elementary school students can learn probabilistic or statistical concepts. The main purpose of this study is to demonstrate the feasibility of teaching a class of sixth-grade students a unit in probability and statistics.

During the past decade several recommendations for including probability and statistics in the elementary years have been made by various mathematicians and mathematics educators. In general, a comprehensive program for probability and statistics beginning in the elementary school and continuing through the junior high and high school has been recommended. During the elementary years, the concepts of probability and statistics should be approached in a very intuitive fashion with the child as he participates in experiments and games. However, the present elementary textbooks are practically void of probabilistic and statistical concepts. A secondary purpose of this study is to develop a set of instructional materials on probability for the elementary school.

The remainder of Chapter I is devoted to presenting the rationale, the recommended content and approach, the existing conditions, and the related research for teaching probability and statistics in the elementary school. In Chapter II the development of materials is presented. Chapter III contains a description of the design and conduct of the study. Chapter IV reports results of the study, and in Chapter V, the conclusions and recommendations of the study are presented.

## Rationale

The immediate question raised by many people concerning the teaching of probability and statistics in the elementary school is "why begin in the elementary school?" The following are eight reasons for including probability and statistics in the elementary school.

1. Probability is one of the most widely-used branches of mathematics; it should be taught to all students.
2. Many probability and statistical concepts are deep and subtle. To learn and understand them requires many exposures with varying degrees of experience and rigor.
3. Probability and statistics are not falsified by exposition in elementary terms. One can approach probability and statistics in an intuitive fashion using games and experiments and still teach good mathematics.

4. There is a need for children learning mathematics to see and study instances of uncertainty before they become overly enamoured with "getting the exact answer."

5. Probability and statistics properly taught from an experimental, game-oriented framework should be exciting and interesting to children.

   It should provide a novel and interesting context for standard mathematical ideas and thereby deepen the student's understanding of and interest in these ideas (Page, 1959, p. 230).

6. Probability and statistics can provide students many new and interesting situations to practice old skills and concepts. For example, much of the drill on learning concepts and computation skills with rational numbers could be disguised within probability and statistical units.

7. Probability is a good mathematical model of the real world and offers children considerable practice in creating mathematical models which approximate reality.

8. Children should be able to understand and profit from the type of instruction implied in Reason 5.

Of the eight reasons, the first two need to be considered in detail.

Reason 1, the utility rationale actually consists of two parts.

A. Probability and statistics concepts are of great importance to many occupations.

For example, they are used in making decisions in such diverse fields as government, the military, design and quality control of manufactured products, scientific research, agriculture, weather forecasting, education and insurance calculations. The CEEB Commission on Mathematics (1959) recommended the inclusion of a course in probability and statistics in the secondary curriculum. The report stated:

> So great is the current scientific and industrial importance of probability and statistical inference that the Commission does not believe valid objections based on theoretical considerations can be offered to its inclusion in the curriculum. . . dissent can only arise, it feels, on the ground of the difficulty of carrying out the task. (p. 32)

This statement might apply equally well to the elementary curriculum if one could show that children can learn some of the intended concepts.

<u>B. Probability and statistics concepts are important and useful to the ordinary citizen, including the school child.</u>

In everyday life people make decisions based on what is essentially statistical data. However, they have many misconceptions and lack background concerning the concepts of chance and statistical procedures. They have little knowledge of ways statistical data can be manipulated to support various arguments.

Since probability and statistics concepts have wide utility, it seems important that all people receive training in the concepts. Yet in every grade after the elementary school, a certain number of students either drop out of school or stop studying mathematics. Hence the training in probability and statistics must begin early.

Reason 2,

Many probability and statistical concepts are deep and subtle and to learn and understand them requires many exposures with varying degrees of experience and rigor.

Bruner (1960) stated in his discussion of the "spiral curriculum":

> If the understanding of number, measure, and probability is judged crucial in the pursuit of science then instruction in these subjects should begin as intellectually honestly and as early as possible in a manner consistent with the child's forms of thought. Let the topics be developed and redeveloped in later grades (pp. 54, 55).

The Cambridge Conference (1963) seems to agree with Bruner's statement and has recommended that probability be taught in "four doses" through the curriculum, beginning with an intuitive, empirical approach to the concepts (pp. 71, 72). The feeling expressed by Page (1959) is that a student should be familiar with some fundamental, intuitive ideas of probability and be exposed to real and imagined experiments before proceeding to a thorough and rigorous study of the subject.

The other six reasons have similarly been supported by the Cambridge Conference Report (1963), the Cambridge Conference Summer Study Report (1966), School Mathematics Study Group (SMSG) (1966), Page (1959), or Engel (1966). In summary, the eight reasons lend strong support for including probability and statistics in the elementary curriculum. However, these reasons do not convey what the intended concepts should be and what approaches might be used.

## Recommended Content

The study of probability as it is known today was launched in the mid-17th century. However, many of the results in probability and statistics used today were discovered only in this century. It is not surprising that before the Commission of Mathematics Report, probability and statistics were almost nonexistent in contemporary school textbooks. Descriptive statistics was sometimes included in the junior high regarding constructing and reading graphs of different types. Some intermediate and advanced high school algebra texts included a small unit on probability. But significant changes in curricula have occurred since 1959.

In probability, the curriculum proposals of the Cambridge Conference (1963), the Cambridge Conference Summer Study Report (1966) and the proposals of Page (1959) and Engel (1966) were examined in making a list of recommended topics in probability. The more frequently recommended topics were:

1. The use of experiments by children to provide an empirical approach to the subject.
2. Counting paradigms (trees, tables, lattices).
3. Formal counting procedures (sum rule, product rule, permutations, and combinations).
4. Relative frequency of an event to approximate the probability of that event.
5. Computation of probabilities in a finite sample space.

The Cambridge Conference Report and the Cambridge Conference Summer Study Report made recommendations about what topics in statistics should be taught in the elementary school. Included was the idea that students carry out experiments to find answers to problems. Such an experiment would involved deciding what data to collect and how to go about collecting it, organizing and graphing it, and finally interpreting the data to answer the question. Such work naturally leads to the inclusion of measures of central tendency (mean, mode, median) and measures of dispersion (range, absolute average deviation, quartiles) so students can talk meaningfully about the data. The graphs suggested are bar graphs, histograms and line graphs (relative frequency against trials and relative frequency against cumulative frequency of trials). The recommendations viewed the skills of graphing as a necessary pre-requisite before the student could observe the long-term stability of the ratio of successes to number of trials and that probability has to do with long-term ratios, not individual events. So descriptive statistical skills should precede certain probability concepts and should assume an integral part in building intuitive understandings of probability concepts. (At a later time the opposite will be true. When statistics is used for predictive purposes, statistics then demands a background of probability.)

## Recommended Approach

If one accepts the rationale and the recommended content, the next question is: "How should the teaching of the concepts be approached?" The Cambridge Conference Summer Study Report states:

> It is our belief that the study of probability (as well as the early study of other mathematics) ought to be closely associated

> with the real world. This means that the children will perform many experiments, and will attempt to draw mathematical conclusions from those experiments. In the early grades, the mathematics will be of a very informal nature, and the children will be getting a feeling for certain concepts, without necessarily stating them explicitly. At a later time, more explicit, quantitative conclusions will be drawn and analyzed (p. 1).

This approach expects the student to perform experiments and play games, and here the teacher should use an informal, inductive teaching procedure to present the intended concepts. However the recommended content and approach differ radically from the present status of probability and statistics in the schools.

## Existing Materials in Probability and Statistics

Today, a number of high schools offer an elective course in probability and statistics to students who usually have had three years of high school mathematics. In the junior high school, probability and statistics units have been published by SMSG (Introduction to Probability, 1966), Secondary School Curriculum Improvement Study (1967), and some commercial publishers. A more recent development has been the inclusion by a few publishers of short units on probability and descriptive statistics in elementary school textbook series. The books published by Standford University Press, American Association for the Advancement of Science (AAAS), Singer, and Scott Foresman contain such units. These books, along with a sampling of commonly used elementary textbook series published by Science Research Association (SRA); Addison-Wesley; Holt, Rinehart and Winston; and Harcourt, Brace and World, are reviewed for probability and statistics content.

School Mathematics Study Group

The most extensive materials presently available for use in the elementary school are the two enrichment books written by SMSG (1966), one for the primary grades and the other for the intermediate grades titled Probability for Primary Grades and Probability for Intermediate Grades. The materials do not presume any prior knowledge of probability concepts. Each concept is developed through a series of suggested class activities and written exercises. Many games and experiments using spinners, dice, etc. are performed by the child. The grade level of the materials is not specified.

In Probability for Primary Grades the authors state, "experience indicates that Lessons (1-5) can be understood by children in kindergarten and grade one." Lessons (1-5) are concerned with the concepts of certainty and uncertainty, comparing the likelihood of simple events, combining events--this one and that one; combining events--this one or that one, and considering the number of possibilities. Lessons (6-8), they report, seem to be able to be comprehended quite well by second-grade children. These lessons include combinations of 2, 3, or 4 things and ordered arrangements. Lessons on arrangements and repeated trials (i.e., 2, 3 and 4--tuples) and the probability of these events, they feel, should be completed by the end of the third-grade.

In Probability for Intermediate Grades the authors again assume that children have had no previous instruction concerning probability. Many of the same primary book concepts are taught in the intermediate book. However the intermediate book includes material on bar graphing to help interpret experiments and experiments which are not game oriented as they are for the primary years. New topics introduced in Probability for Intermediate Grades include the use of a fraction to denote a probability, notation symbolizing the probability of an event, use of "trees" as a listing technique, probability of a complementary event, the evaluation of a combination via "Pascal's Triangle", conditional probability, and the meaning of the "law of averages". Again, no grade recommendations are made, but it is suggested that teachers space materials throughout the year rather than using all of them at once.

There are weaknesses in these materials, particularly in the intermediate book. Since experiments are not game oriented, some children when asked to do many of these activities, probably become bored. Also, no attempt is made to use cumulative frequencies. Hence, the graphical comparison of unpredictability, say of the single toss of a thumbtack, cannot be compared with the results from a large-scale, cooperative experiment by the entire class in perhaps 2000 tosses. Thus the "swamping" effect of very large averages would not be observed, a rather crucial omission since the "law of average" or large numbers is discussed in the book.

Still another fault of the materials is the overstructuring of many of the problems (i.e., providing the table of "tree" to be filled in) without providing problems where the student must find his own structure for the solution of the problem. Nor is an attempt made to link probability and statistics. Other than bar graphing and tabulating results from experiments, no use is made of descriptive statisticts.

These books are concerned with the development of probability and seem to ignore opportunities to develop, in a more systematic way, the child's ability to investigate a problem. Why can't he set up an experiment and decide how he would collect, organize and interpret data to answer the question the problem raises? Many experiments are posed, but the tables and methods of graphing and the questions to be answered are explicitly given. Some structuring of materials is fine, but the student must be given some opportunity to practice these statistical skills in an unstructured situation.

Claims for the feasibility of getting children to learn the concepts need to be investigated at each age level. If the writers had empirically evaluated the materials, their statement that "Experience shows that..." might be quite different. However, the books possess many more strengths than weaknesses and are presently the most extensive and best suited materials for probability in the elementary years. One should note that there are wide discrepancies between the Cambridge Report's recommendations and these materials, although SMSG originally shouldered the task of developing materials based on recommendations in the Cambridge Report.

AAAS

Science--a process approach written and published by AAAS (1965) has included topics in probability and statistics in three of their intermediate, experimental teacher manuals (Part 5B, Part 6B and Part 7A) to aid in the development of the processes of experimenting, measuring and interpreting data. As in this study, the objectives are stated behaviorally.

The objectives for booklet Part 5B are for the child to:

1. identify all possible outcomes of an event in which there are a finite number of outcomes.
2. state which among all possible outcomes are favorable and which are unfavorable outcomes.
3. demonstrate the computation of the probability of an event for which all favorable and possible outcomes can be listed and counted. (p. 195)

Four activities were written so that teachers could accomplish the objectives. The activities begin by teaching students to assign probabilities in one-dimensional sample spaces. The probability of an event such as drawing a red marble from a box is defined as P(drawing a red marble) = $\frac{\text{number of favorable outcomes}}{\text{number of possible outcomes}}$. An activity in analyzing and conducting an experiment in drawing a marble from a box is discussed. This activity is followed by an experiment in a two-dimensional sample space. Both experiments are first analyzed logically, with probabilities being assigned to various events, followed by a group experiment in which data is gathered.

In the area of statistics, the booklet includes materials concerned with computing and using the mean, mode, median, and range to interpret a set of measurements.

In Part 6B the objectives for the child are to:

1. construct a probability chart when the number of possible outcomes involves two or three successive events.
2. state and demonstrate a procedure for determining the probability of an event when the number of possible outcomes and the number of favorable outcomes are known.
3. state and demonstrate a procedure for determining the probability of two events connected by "or".
4. state and demonstrate a procedure for determining the probability of two events connected by "and" using a probability chart. (p. 111)

The emphasis in this unit is on logically analyzing chance events. The unit is introduced by using a top with four possible outcomes and considering the number of outcomes when the top is spun once, twice and three times. A list of the possible outcomes is made and probability problems in one, two and three-dimensions are considered. A die and a pair of dice are used as models for further practice in listing the possible outcomes and in assigning probabilities. No experiments with the models are actually performed.

The objective of the last unit (Part 7A) which includes probability and statistics is:

> A child should be able to describe and demonstrate an experimental procedure to approximate the probability of an event where there is a very large number of possible outcomes or an unknown number of possible outcomes. (p. 29)

For this objective, four experiments are introduced ranging from dropping toothpicks to choosing a marble from a box containing an unknown number of marbles. In the latter case the empirical results are then compared to the theoretical result based on probability assignments.

The major weaknesses of the materials are that they do not go far enough. No attempt is made to graph probability data[1]; little is done to help students understand what probability means; and no written pupil exercises are provided. The books are more interested in developing processes than subject matter competencies in probability and statistics. Except for the objective of Part 7A, all the previously mentioned objectives are included in this study.

---

[1] Bar graphing and line graphing of other types of data are included in the other activities of the series.

Singer

Another source for the elementary school is the Sets and Numbers series authored by Suppes, et al (1966). In this series probability concepts are introduced in the fourth grade. The initial concepts are outcome, event, probability of possible outcomes and probability of an event. The lessons begin talking about the meaning of "probably," a few demonstrations by the teacher, and then proceed immediately with assigning a probability number (1/n) to an outcome of a symmetrical spinner with n possible outcomes that can occur. The lesson materials immediately link one out of n chances with the fraction 1/n. All of this is done in the first lesson. The next few pages introduce, intuitively, the notion of an event. The events considered involve two and three-tuples. Set notation is used to designate an event. The student is told that the union and intersection of two events is also an event. The notion of the intersection of two disjoint sets being the impossible event is introduced. No probability assignments are made to events involving n - tuples.

The fifth-grade book reviews these materials and introduces the probability of events involving two-tuples. Similarly in the sixth-grade book, the preceding materials are reviewed and the probability of the union and intersection of events is introduced. Asymmetrical spinners are considered for the first time.

In all three grades set notation is used extensively. The only models considered for assigning probabilities in the fifth and sixth grades are spinners.

The materials assume that the child can list all possible outcomes and that he can recognize when events are equally likely. There are no exercises included where the child performs an experiment. Is it assumed that the child has or will develop easily the notions of "impossible," "probable," and "certain?" The materials and examples are extremely limited in the fifth and sixth grades since they only consider problems involving spinners. Perhaps this is done because of the number of pages allocated to this material, i.e., grade 4, six pages; grade 5, five pages; grade 6, eight pages, in the final chapters of the respective books.

Regarding statistics, a unit of graphing including the making and reading of bar graphs and line graphs is contained in the fifth-grade book. In the fourth, fifth, and sixth grade, computing the average (i.e., the mean) of a set of numbers is developed and practiced. This concept is not used in connection with graphing. The gathering and analyzing of data to find an answer to a problem does not seem to be a goal of these units.

Scott Foresman

The *Seeing through Arithmetic Series* authored by Hartung, *et al* (1968) includes a chapter on statistics and probability in the sixth-grade book. The unit begins by teaching students to read and plot bar graphs and to compute the mean. The probability materials begin with the concepts of experiment, outcomes, equally likely outcomes and chance. The probability of a favorable outcome is defined as $P = \frac{\text{number of favorable outcomes}}{\text{number of possible outcomes}}$. The probability of a complementary event is approached intuitively. The chapter ends by using the frequency of occurences to estimate probabilities and then in turn using these estimates to predict frequencies for a different number of occurences. The materials are of an informal, problem-oriented nature.

The materials do have some shortcomings. In the student's book the child is never asked to carry out an experiment. All data is furnished. In statistics, the development of the skill of asking a good question and of collecting, organizing and interpreting the data to answer the question is not considered. In probability only one-dimensional sample spaces are discussed. Since no practice in carrying out experiments is required, the concept of the randomness of the actual outcomes is not developed. Yet the materials include problems such as, "in flipping a coin you can predict that 1/2 x 1600 will be heads." The materials give the impression that this is what will happen. The concept of expected value is developed before that of randomness. This difficulty could be avoided if the students carried out experiments, graphed data and compared results. Like Sets and Numbers published by Singer, this commercially prepared unit still leaves the field wide open for someone to develop a well planned, validated program in probability and statistics for the elementary school.

## Other Publishers

The Greater Cleveland Mathematics Program (GCMP) published by SRA (Educational Research Council, 1964) includes no work on probability in their elementary series. In the sixth grade, histograms are introduced along with median, quartiles and range. Similarly the Addison-Wesley series, Elementary School Mathematics, written by Eichoiz, et al. (1963, 1964) does not include any work on probability and very little on statistics other than simple graphs such as the line graph.

The Elementary Mathematics: Pattern and Structure series of Nichols et al (1966), published by Holt, Rinehart and Winston, has no work in probability in the elementary schools. In the sixth grade, an 18-page unit on statistics includes organizing data, line and bar graphs and averages (mean, mode, median).

The Harcourt, Brace and World series, Elementary Mathematics, written by Payne, et al., (1966) includes no work in probability and statistics in its elementary program.

If one compares the rationale, recommended content and approach to the existing materials, one finds them to be inadequate. Thus the investigator decided to develop a new unit using the best ideas contained in the existing materials. However, before considering the development of these materials, the question, "What evidence exists to support the inclusion of the recommended content in the elementary curriculum?" must be considered. The evidence or related research is divided into two parts: (1) status studies and related experiments and (2) curriculum experiments.

## Summary of Status Studies and Related Experiments

There are a number of studies which shed some light on the questions about what children may learn from their environment concerning certain concepts of probability without formal instruction and when this learning takes place. Of the status studies concerning probability and statistical concepts possessed by children, the most diverse and influential have been those conducted by Piaget and Inhelder. Through a series of ingenious tasks given to children, these scientists observed three stages in children's development of the ideas of chance. (They also have approximated age intervals for these stages of development.)

The first stage (approximately up to age seven) is identified as the preoperational stage. In this stage the child lacks the intellectual operations necessary to recognize events which are certain, much less those which are uncertain. As Flavel (1963) in summarizing Piaget states:

> In order to identify a set of phenomena as "chance events" one first has to identify a set of phenomena which are not chance events, a nonchance ground against which chance can emerge as figure (pp. 341-342).

There is a generalized non-differentiation between chance and nonchance, between the possible and the necessary, during this development period. In this stage the child would show little or no consternation over getting twenty heads in twenty flips of a coin. He is unable to grasp the extreme unlikelihood of such an event.

The second stage of development, the concrete operational stage, occurs when the child recognizes events he can know and those he can only guess. He clearly recognizes chance events when he encounters them. However, when faced with predicting an outcome, he will do this in terms of the absolute number of favorable or unfavorable outcomes. He typically cannot list all the possible outcomes for a complex event.

In the third stage, the period of formal operations (ages 11 and up), the child is able to think in terms of combinations, permutations and proportion. Thus he now possesses the intellectual tools to handle many chance events adequately where previously he could only recognize them. In the previous stage, the child is able to quantify simple events but makes his decisions in terms of absolute differences.

The child at the formal operational stage is able to quantify events and make correct decisions in terms of proportions. Some individuals are known to progress incompletely through all three stages and thus only partially acquire the concepts of chance.

As the result of these and subsequent investigations Bruner (1960) quotes Inhelder as saying:

> The teaching of probabilistic reasoning, so very common and important a feature of modern science, is hardly developed in our educational system before college. The omission is probably due to the fact that school syllabi in nearly all countries follow scientific progress with a near-disastrous time lag. But it may be due to the widespread belief that the understanding of random phenomena depends on the learner's grasp of the meaning of the rarity or commonness of events. And admittedly, such ideas are hard to get across to the young. Our research indicates that the understanding of random phenomena requires, rather, the use of certain concrete logical operations well within the grasp of the young child--provided these operations are free of awkward mathematical expression. Principal among these logical operations are disjunction (either A or B is true) and combinations. (p. 45)

Many critiques and interpretations of Piaget and Inhelder's studies have been written. A most pertinent comment in relation to this thesis was made by Sullivan (1967) in regard to Piaget's Stage Theory as an aid to the structuring and sequencing of subject matter in a curriculum.

> The Piagetian contribution to the structure and sequencing of subject matter is more apparent than real. This is clearly not the fault of Piaget, but rather of his educational followers. Uncritical extrapolation of Piaget's observations and his metatheoretical considerations (e.g., logico-mathematical model) is, in the opinion of the present author, harmful to the advancement of educational knowledge. The use of Piaget's stages as indicators of "learning readiness" seems most premature and needs more careful consideration on both the research and theoretical levels. (p. 23)

Some studies have been conducted to test Piaget and Inhelder's conclusions. Of these, only a few will be cited briefly. Yost, et al (1962) report a relationship between mental age and understanding of probability. Davies (1965) confirmed Piaget's notion that the concept of probability is a developmental phenomenon. However, she felt Piaget's interviews depended highly on verbal ability. She concludes that the nonverbal behavior reflecting event probability appears earlier than the verbalization of the concept of probability or of its applications to tasks. Her study also indicated there were no sex differences.

The studies of Davies (1965), Goldberg (1966), and Yost, et al (1962) seem to show that young children, 4-6 years of age, do demonstrate some understanding of probability under appropriate conditions, contrary to Piaget and Inhelder's results.

The three status studies of Doherty (1965), Leffin (1968), and Shepler (1968) indicated that students in grades 5 and 6 had acquired considerable knowledge concerning:

1. sample points of a sample space.
2. probability of a simple event.
3. probability of a compound event.
4. decisions between two boxes to maximize one's chance of winning.

Leffin's study also has several implications for teaching these topics. Among these were:

1. Children must be taught the difference between "odds" and "probability."
2. One cannot be certain that a child who gives a correct response to a question about the probability of a simple event actually recognizes all the elements of the sample space which contains the event.
3. Probability of a simple event which involves combinations was an extremely difficult topic for all grades.
4. Young children may be able to understand No. 3 after training, but this needs further study with regard to quantification of probabilities; the children often base their answers on the number of winners rather than on the probabilities of success.
5. There is need for a further study to show the effects the teaching of these concepts have on the performance of children in applying concepts.

Leake's (1962) study with junior high school students implies similar implications for teaching probability.

Also of interest to this study are Cohen's[2] experiments regarding subjective probability, the use of the concept of probability in risk

---

[2] John Cohen, C. E. M. Hansel, Risk and Gambling (London: Longmans, Green, 1956); John Cohen, "Subjective Probability," Scientific American, 197 (November, 1957), pp. 128-138; J. Cohen, E. J. Dearnaby, C. E. M. Hansel, "Measures of Subjective Probability," British Journal of Psychology, 48 (1957), pp. 271-275; J. Cohen, Chance, Skill and Luck: the Psychology of Guessing and Gambling (Baltimore: Penguin Books, 1960).

and gambling situations and the concept of independence. His works show young children are greatly influenced by subjective preference, ideas of fairness and superstitious behavior in dealing with situations involving probability. Of pertinent interest for this study are his studies on independence of events. He concluded that this concept emerged at 12+ years and continued to be strengthened from that time on.

There are a host of research studies dealing with probability learning experiments. Primarily these studies are concerned with the application of subjective judgments in risk situations or learning experiences which include some aspects of probability as it is applied to a specific task. These studies have little evidence to offer about the status of a concept, the construction of a curriculum in probability and statistics, or the behaviors of students who have taken such a course.

The repeated themes running through research cited here seem to suggest that children are learning a significant amount concerning concepts of probability without formal instruction and that this learning is taking place at an age younger than many educators expected. Many of the studies point to the need for empirical curriculum studies to test the feasibility of teaching probability concepts in the elementary school. Some studies also show that mental age, achievement, and grade level may be helpful in predicting the status of these concepts.

## Summary of Curriculum Experiments

Since this study is a curriculum experiment, the curriculum experiments of Wilkenson and Nelson (1966), Smith (1966), Ojemann, *et al* (1965), Grass (1965), and Girard (1967) need to be considered in detail. The most pertinent curriculum experiment was conducted by Wilkenson and Nelson using 22 sixth-grade students from the Laboratory School State College of Iowa. The experiment was conducted late in the first semester for three weeks, 45 minutes per day, to test the feasibility of students' learning certain concepts of probability and statistics. No testing was done and only subjective information was reported. One author did all of the teaching and the other acted as an observer-recorder. The concepts they wished students to have experience with were randomness, probability, odds, bias, variation, law of large numbers, levels of confidence, sampling, mean, median, mode, distributions, equally and unequally likely events. The skills they wished to develop were recording and interpreting data, making and testing hypotheses, and observing and differentiating between relevant and irrelevant data.

They subjectively judged the experiment to be a qualified success. They believed that the probability experiences were worthwhile, and that it is important to expose students to events which have a degree of uncertainty. Wilkenson and Nelson conjectured that activities in this grade should enable students to compare their individual guesses with what they find taking place as they carry out an experiment and that a series of experiments involving probability concepts would help them assess and interpret uncertainties.

Among the recommendations Wilkenson and Nelson made were the following:

1. Avoid pre-prejudiced situations such as the tossing of coins or the throwing of dice. (They felt that students had formed

a variety of prejudiced opinions about these and that experimental evidence did not overcome this. In the situations where the materials were unfamiliar to the student, they reported few such problems.)

2. Keep vocabulary useful and simple. Give special treatment to words such as always, never, usually, sometimes, etc.
3. Do not overstructure students' experiences.
4. Use experiences meaningful to the student.

While the results of this study are useful, the population of a laboratory school and the nature of the subjective evidence collected make generalizations from this study very tenuous. Recommendation 1, the avoiding of pre-prejudice situations such as coins and dice, would seem to be the opposite of what one would want to do to make the instruction meaningful and useful to the child. What might be a more helpful approach would be to introduce a variety of models of the same probability space (introducing the unfamiliar first) and then help the student to discover that the models are equivalent.

A second study of some importance to this formative evaluation was conducted by Smith (1966). The two-fold purpose of the study was (1) to develop materials in probability and statistics for seventh graders and (2) to evaluate the effectiveness of these materials.

Development consisted of making a teacher's and student's manual covering elementary topics. The objectives for the eleven topics included in the unit were formulated under the divisions of knowledge, understandings and abilities. Smith used an objective test consisting of fifty items in order to obtain a pre- and post-test measure. The eleven topics were:

1. Equally likely events.

2. Events that are not equally likely.

3. Mutually exclusive events.

4. Independent events.

5. Models.

6. Sampling.

7. Pascal's triangle.

8. Continuous and discrete data.

9. Histograms and frequency polygons.

10. Central tendency, including mean, mode and median.

11. Variation, including range and average deviation.

The materials were taught to four classes (97 students) for seventeen days. The control group consisted of two seventh-grade classes (53 students).

One conclusion of the study was that high and middle ability students learned significantly more than those designated as low ability students. However, the low ability group did learn a significant amount. The study indicated all but three of the eleven topics, namely, independent events, sampling and measures of variation (including average deviation), were appropriate for these low ability students. However, these three topics did seem to be appropriate for students of high ability.

Smith states that his study seemed to demonstrate the feasibility of seven-grade students' learning certain topics in probability and statistics. The "median proportion passing" on the pretest was .32, and on the posttest, .62; however, passing is not defined. The posttest mean was .58. The mean on the pretest was not reported. One would have to judge the instructional treatment to have produced poor results in relation to mastery of topics, particularly in comparison to an 80/80 or 90/90 criterion. To both teach the eleven topics and obtain good performance levels would no doubt take more than 17 days of instruction. Smith seems to have shown that a low level of overall mastery was possible using his unit of instruction. Many of his objectives were not stated behaviorally (e.g., a student should know ... or a student should understand...), and no data were gathered to support or reject the attainment of these objectives.

Ojemann, _et al_ (1965) conducted two very similar experimental studies to teach certain probability concepts to third grade and fifth grade students. The major abstractions that were taught were:

> (1) when a random selection is made from a group of equally available alternatives, each alternative has an equal chance of being chosen; (2) when there are more of one kind of item than another and the selection is random, the former has the greater chance of being selected; (3) additional pertinent information usually helps in making a decision; (4) as the number of factors affecting a situation increases, combining their effects tends to become more difficult. (p. 416)

The instructional treatment for both studies consisted of five sessions in which experiments using physical models were conducted by the teacher and the students.

Ojemann, _et al_ concluded that the third grade and fifth grade students had developed considerable ability in relating their predictions to the information available. In relation to students ability to maximize success when prior knowledge of the sample space is or is not available, they concluded that after treatment third grade students tended to maximize success in prior information situations without the use of extraneous rewards. The fifth-grade students tended to maximize success in situations without prior information of the proportions in the sample space and also without the use of extraneous rewards.

The treatment did not include teaching students to assign a probability to an event. The study was more concerned with developing students' subjective notions of probability as influenced by their own observations of patterns of outcomes of various experiments. Thus, actual probabilities were never employed in the decision-making processes investigated.

A curriculum experiment concerned only with statistics was performed by Grass (1965) using a fourth and fifth-grade class. The teaching unit included the concepts of central tendency--mean, mode and median. He reports that he helped the students to devise and evaluate their own statistical study to determine whether boys or girls have greater extrasensory perception. Actual hypotheses were made and tested using mean, mode and median. He reported students' interest and enthusiasm to be very high for the unit. However, Grass offers no objective basis for his conclusions.

Girard (1967) describes a curriculum experiment involving the development of a unit on descriptive statistics and graphing with a class of elementary students. The main purpose of the instruction was to develop children's ability to critically interpret statistics and graphs. Girard reported that students were enthusiastic about the topic and one for which they willingly carried out special projects on their own. Other than the teacher's description of the proceedings, no empirical evidence was gathered and reported.

In retrospect, the amount of empirical evidence available for objectively judging the feasibility of teaching probability and statistics concepts in the elementary school is quite small. However, the status studies and curriculum experiment do suggest that teaching certain of these concepts may be feasible at certain age levels of children in the elementary school.

## Summary of Recommendations, Existing Materials, and Research As They Relate to This Study

Most mathematics educators feel that probability and statistics should be a part of the school program in mathematics; but there is a wide difference in opinion about what should be included, when it should be taught, and what approach should be used. Some argue for the inclusion of probability and statistics as a formal course for the high school and others for an informal course at that level; others have argued for the inclusion of probability and statistics beginning in the junior high school. A more recent trend has been to recommend a comprehensive program of instruction which begins in the elementary years and extends throughout the high school years. Advocates of this approach see the elementary years as a time when concepts of probability and statistics should be approached in a very intuitive fashion in which the child learns through experiments and games.

In considering the recommendations, existing curriculum materials and research, one finds quite a bit of discrepancy between them. With respect to existing curriculum materials, none of those reviewed include all of the suggested topics. In fact, only the SMSG's Probability for Primary Grades and Probability for Intermediate Grades contain material close to that suggested for probability for the elementary school and even these materials are woefully weak in the development of statistical concepts and processes. Thus, a new set of materials had to be developed for this study.

With respect to research, the hope of changing through instruction the age at which children learn the concepts of probability and statistics has some support. Children may learn probability concepts at least a year to three years earlier than the age at which Piaget and others concluded that children form these concepts. Whether earlier ages are realistic will only be known after much experimentation in teaching the concepts to children of different ages, ability and background. However, the status studies did influence certain factors in this study.

In the Statement of the Problem at the beginning of the chapter, the sixth grade was selected as a starting point for teaching probability and statistics in the elementary school. Leake (1962) supports this choice in his summary of the research by commenting that the results of these experiments add evidence to the theory that the age range centered around 12 to 13 years is a formative period for the concepts dealing with probability. Since a high degree of mastery learning was desired for the short unit of instruction developed in this study, the status of the development of probability concepts and prerequisite skills (e.g., ordering of two ratios) played an important role in choosing

Grade 6 (age 12) in which to perform this experiment and in deciding which concepts to include and exclude (e.g., independence is excluded).

The following chapter presents the content selected and the procedures used to develop a unit in probability and statistics for use with sixth grade students.

## Chapter II

## DEVELOPMENT OF THE INSTRUCTIONAL PROGRAM

The purpose of this study was two-fold: (1) to test the feasibility of teaching topics in probability and statistics to a class of sixth-grade students; and (2) to construct a set of instructional materials and procedures in probability and statistics for sixth-grade students. The latter purpose of the study is discussed first.

The author used a working paper by Shepler, Harvey and Romberg (1969), and the developmental model of Romberg and DeVault (1967) to build the unit. Shepler, *et al*, constructed a framework for developing an instructional system in probability and statistics for use in the elementary school. The paper included a content outline, a task analysis of content, and specific grade recommendations for topics used in elementary school. The present study was designed to test the feasibility of parts of the working paper.

From strands of the task analysis, the author decided upon behavioral objectives for the unit of instruction and the order in which objectives would be taught. Using this basis, an instructional analysis of the unit was undertaken next. The purpose of this analysis was to select or develop materials and procedures for teaching the unit of probability to sixth-grade students.

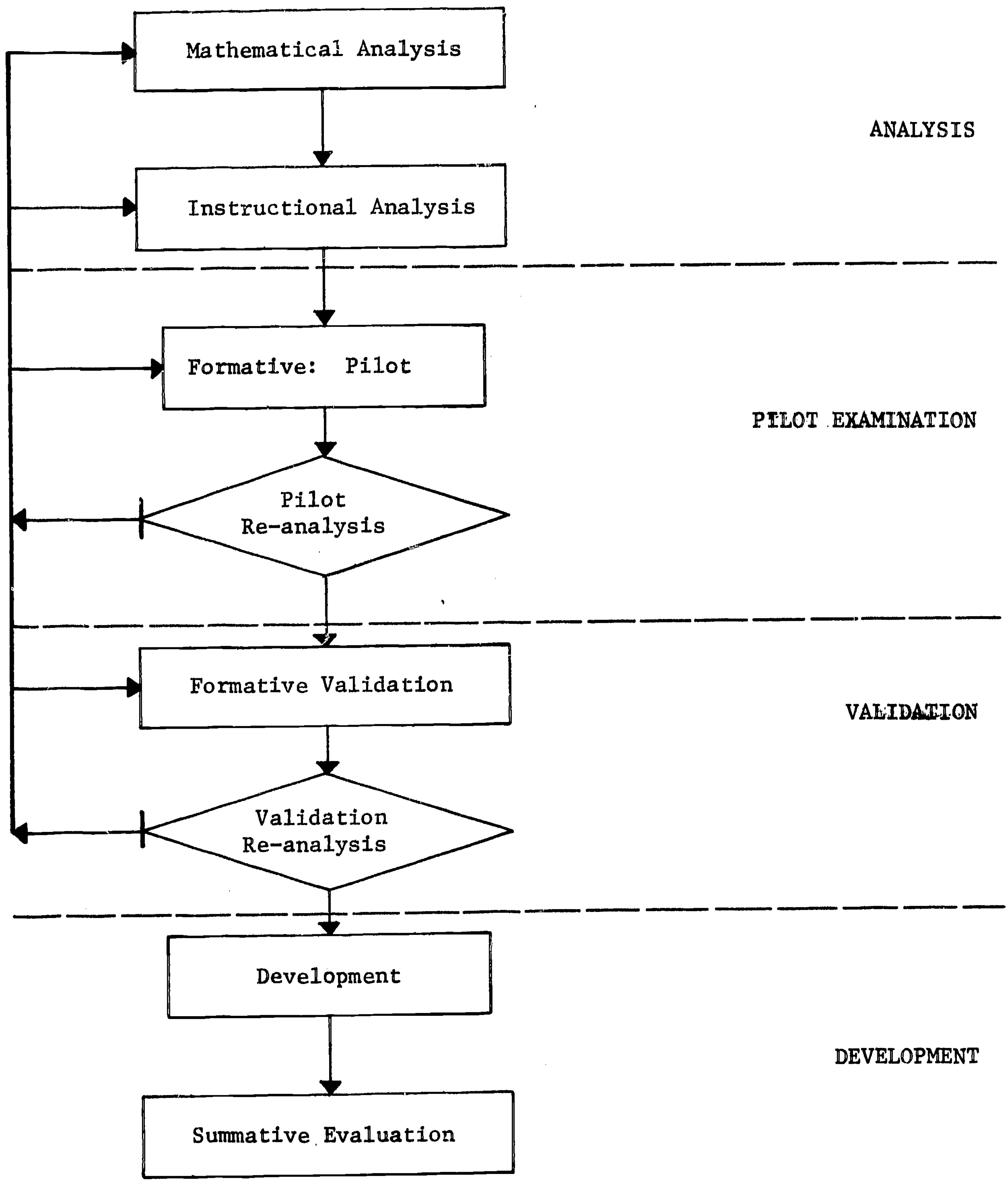
To aid in the developmental processes of task analysis and instructional analysis, a pilot study was conducted in the Fall of 1968. The data from the pilot study was used to identify a set of nine lessons that could be formatively evaluated to test the feasibility of the instructional analysis. The lessons were to be used to teach a class of sixth-grade students of average to above average ability. The goal of instruction was to demonstrate mastery learning of the behavioral objectives. By employing a pretest and posttest designed to measure the behavioral objectives, the feasibility of the unit was to be tested. In addition, certain learning hierarchies within the task analysis strands were examined using Walbesser's paradigm (1968).

## Mathematical Analysis

The first stage of the Romberg and DeVault model (Figure 1) is the analysis stage which involves an analysis of the mathematical content followed by an analysis of how to communicate the content to students successfully. Romberg and DeVault state the mathematical analysis stage includes listing the goals, content and behavioral objectives. The goals of the study have been discussed in Chapter I. The content and behavioral objectives are further subdivided for this study as follows: (1) the content outline, (2) the definitions of important mathematical terms, and (3) the behavioral objectives of the unit. Also considered are the parts of the task analysis for probability and statistics pertinent to this study and a description of two learning hierarchies that are tested.

FIGURE 1

STEPS IN DEVELOPING AN INSTRUCTIONAL SYSTEM



CONTENT OUTLINE

The following content was selected to be investigated in this study (Shepler, et al, 1969):

1. Subjective Probability (distinguishing between certain, uncertain and impossible events).
2. Probability (*a posteriori*).

   P (A) = s/n---where s = number of successes and n = number of trials of an experiment.
3. Probability (*a priori*)--finite sample space for equally likely outcomes: $P(A) = \frac{\text{number of favorable outcomes}}{\text{number of possible outcomes}}$
4. Counting technique--listing possible outcomes via a tree diagram.
5. Probability of an event to include
   a. the certain event
   b. the impossible event
   c. simple event (one outcome)
   d. compound events (two or more outcomes)
      1. problems involving A or B (where A and B are events in a finite sample space S)
      2. problems involving A and B
   e. any of the above in a two or three-dimensional sample space where a tree diagram is employed in counting the number of outcomes in A and the number of possible outcomes in S.

6. Comparison of two or more events
    a. (Most--least) likely event
    b. Equally likely events
7. Performance of an experiment
    a. Collecting data
    b. Interpreting data

It is obvious that one could not develop and formatively test all of the proposed content for grades (1-6) in one study. The investigator was convinced that an intuitive approach to probability as recommended in Chapter I, using basic skills of statistics, was the most appropriate way to begin. Experiments to help develop understanding of probability were considered important. This and the desire for an application of probability led to the inclusion of the topic of choosing between two games. In order to insure that probability would have to be used in making the best decision between games, the author included those which had one and two dimensional sample spaces. The report of the pilot trial presented on p.(49 ) was helpful in making final decisions as to what specifically would be included.

The prerequisite skills identified for learning the previously mentioned content are:

1. identifying the order of two or more ratios (e.g., 1/2 > 3/7).
2. counting.
3. drawing a bar graph.
4. interpreting a bar graph.

Although students might not possess skills 1, 3, and 4, instruction on these skills was planned as need for them arose in the unit.

Basically this study was concerned with certain principles and concepts of probability. The following pages define these terms.

DEFINITIONS

The term "probability" as it is used in this study needs to be clarified. Cohen (1956) quotes L. J. Savage as indicating that "There must be dozens of different interpretations of probability defended by living authorities and some authorities hold that the concept of probability may have different meaningful senses in different contexts (p. 23)." Newman in The World of Mathematics identifies three main interpretations of probability (p.395). For this study there are four definitions that are useful.

1. Classical (a priori) view: If an experiment can result in any one of n different, equally likely outcomes, and if exactly m of these outcomes correspond to event A, then the probability of event A, $P(A) = \frac{m}{n}$ .

2. Generalization of (1): Given an experiment with a set S of finite outcomes. Let 0j be an outcome of the experiment. To each outcome 0j of the experiment one assigns a number P(0j) called the probability of the outcome 0j such that the following two conditons are satisfied:

(1) The probability of each outcome is a non negative number; $P(0j) \geq 0$; $(j = 1, 2, \ldots n)$, n = cardinality of S.

(2) $$\sum_{j=1}^{n} P(0j) = 1$$

3. A posteriori or empirical probability: An experiment is repeated n times, denoting the number of successes as S. The number s/n is an approximation to the probability of success which is defined as the $\lim_{n \to \infty} s/n$ if such a limit exists.

4. Subjective probability: This refers to someone's state of mind, the certainty or uncertainty of his beliefs. It is intuitive and yet there are degrees of uncertainty which are distinguishable by an individual.

Only the first definition is made explicit although all four definitions will be used in this study by the students. Definitions 4 and 1 are the first interpretations of probability encountered in instruction. An attempt is made to link the subjective notions of degrees of uncertainty and *a priori* probability. On the basis of experiments repeated a large number of times and relative frequency plots using line graphs, it is hoped that the relationship between *a posteriori* probability and *a priori* probability would be established. Definition 2 (Generalization of 1) is used in connection with models such as asymetrical spinners (e.g., 3/4 red, 1/4 blue).

None of the following terms are explicitly defined in the lessons. Rather the terms are used in context to build implicit meaning. This is in keeping with the intuitive approach to probability that this study has assumed.

Experiment—any act that can be repeated under given conditions. (Usually the exact result of the act cannot be predicted with certainty and there are only a finite number of outcomes—flipping a coin, tossing a die, etc.)

Outcome—an observable occurrence resulting from carrying out an experiment.

Sample space—the set of all possible outcomes of an experiment.

Event—a set of possible outcomes of an experiment (a subset of a sample space)

Equally likely outcome—when any outcome of an experiment is just as likely to occur as any other outcome of the experiment.

Unequally likely outcome—when one outcome of an experiment is more (less) likely to occur than other outcomes of the experiment.

Certain event—an event which must occur when a given experiment is performed.

Possible event—an event which may occur when a given experiment is performed.

Impossible event—an event which cannot occur when a given experiment is performed.

Estimated probability of an event (A) = s/n where s = the number of successes (i.e., the number of times one gets an outcome of A) and n is the number of times the experiment is performed.

Tree diagram—a structured diagram that may be used to list the possible outcomes of an experiment. (E.g., The tree diagram in Figure 2 for spinning the spinner below twice is

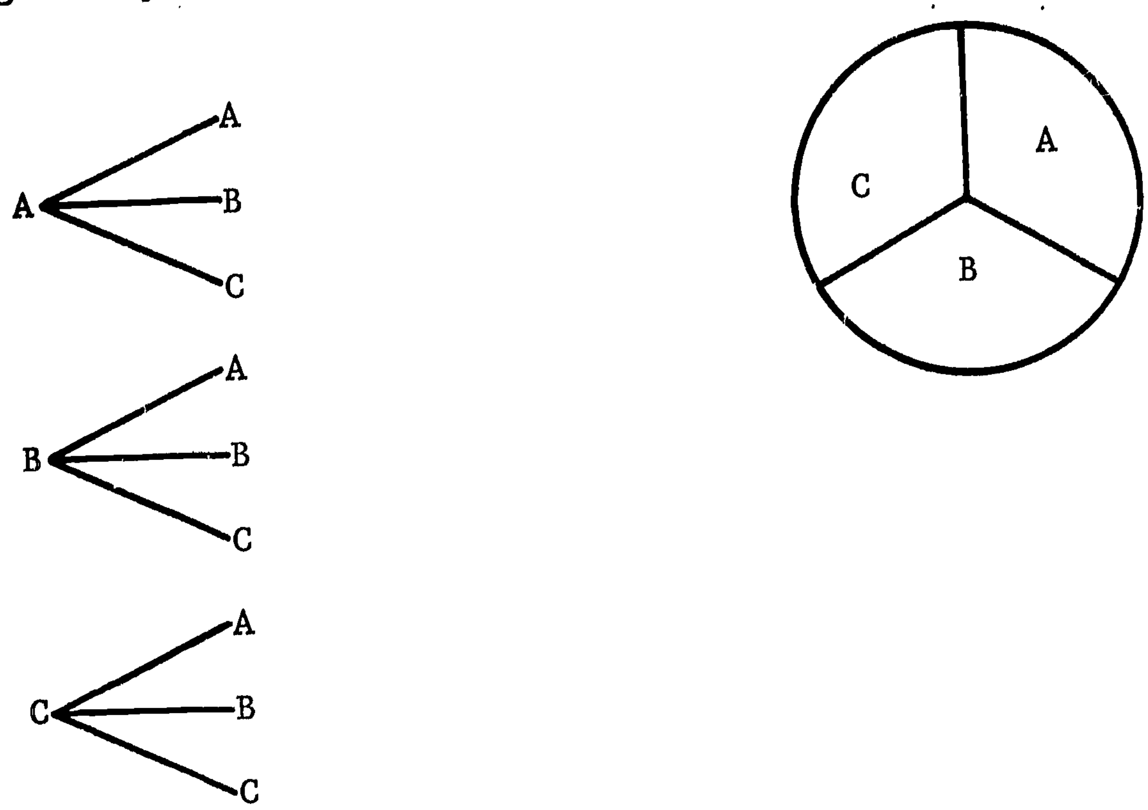



Figure 2. Tree Diagram for Spinner

One-dimensional outcome—an outcome resulting from spinning a spinner once, tossing an object one time, choosing one object from a box, etc.

N-dimensional outcome—an outcome resulting from spinning a spinner n times, tossing an object n times, choosing from a box n times with or without replacement, etc.

BEHAVIORAL OBJECTIVES AND TASK ANALYSIS

Statement of behavioral objectives and the task analysis of these objectives are complementary acts. Behavioral objectives are stated so they describe specific observable behaviors. A key part of behavioral objectives is the use of action words (verbs or phrases) which denote observable activities. A first step in this stage is to construct a task analysis of the identified terminal objectives.

Gagné (1965), who has been very instrumental in the development of task analysis, states:

> The planning that precedes effective design for learning is a matter of specifying with some care what may be called the learning structure of any subject to be acquired. In order to determine what comes before what, the subject must be analyzed in terms of the types of learning involved in it (p. 25).

The principles of task analysis imply the following activities:

(1) An objective to be learned should be analyzed into component objectives which may be learned in different ways and which require different instructional practices.

(2) The successful achievement of the component objectives is required for performance of a terminal objective.

(3) The component objectives have a hierarchical relationship to each other so that successful achievement of one component objective is required for successful achievement of the subsequent component objective.

The task analysis on which this study is based was constructed by examining certain probability and statistics concepts and principles, breaking these concepts and principles into their simpler learning components, then arranging the components into a hierarchy. These components are stated as behavioral objectives. Key steps in writing the behavioral objectives included determining terminal objectives and performing a task analysis of terminal objectives based upon the chosen action words.

## Action Words

The action words were used as operational guides in the construction and evaluation of the instructional objectives as well as in the task analysis. For this study they were:

1. Identifying. Selecting (by pointing to, touching, marking with a pencil, or picking up) the correct object of a class in response to its name or to an implicit instance of its class name. (E.g., identifying a sample space of an experiment for this study would be in terms of identifying all the possible outcomes of the experiment.)

2. Distinguishing. Identifying objects or events which are potentially confusable (e.g., equally likely outcomes, unequally likely outcomes) or for which two contrasting identifications (such as certain and impossible events) are involved.

3. Listing. Writing down explicitly all possible members of a class (set) or of a subset of the class.

4. Constructing. Putting together geometrical objects such as axes, rectangles and/or line segments in a particular, systematic fashion.

5. Counting. Numbering one by one or using a numbering technique to find the total number of members of a class (set).

6. Specifying. Assigning the appropriate rational or natural number to a class in response to either the cardinality of the class, the probability of the class, or a transformation on the class.

## Task Analysis

The chart in Figure 3 shows the task analysis of the probability concepts (Shepler _et al_., 1969) included in this study. The chart reads from the bottom of the page to the top, following the indicated arrows. The task analysis does not specify the instructional analysis of the tasks. It only specifies the hypothesized prerequisites to a terminal behavior. Since this study only involved certain boxes on the chart, boxes and paths actually used are indicated. Using the content outline, the action words, and the task analysis, most of the behavioral objectives of the unit were established.

## Behavioral Objectives

The following are the major behavioral objectives of the unit. The child should be able to

1. Distinguish whether an event is an instance of certainty, uncertainty or impossibility.

2. Count the number of outcomes of an event.

3. Count the number of possible outcomes of a sample space.

4. Specify the probability of a(n)

    a. simple event

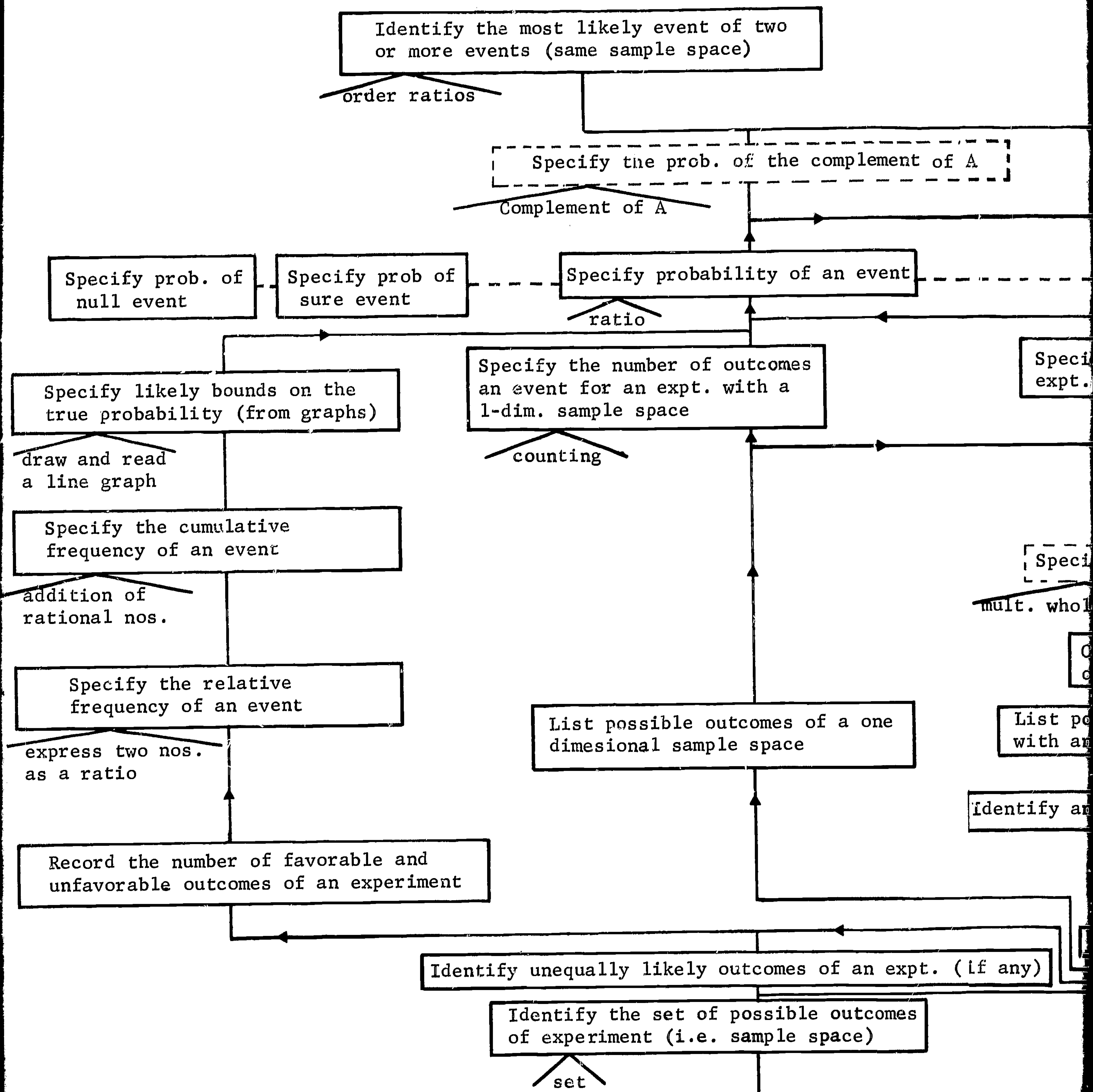
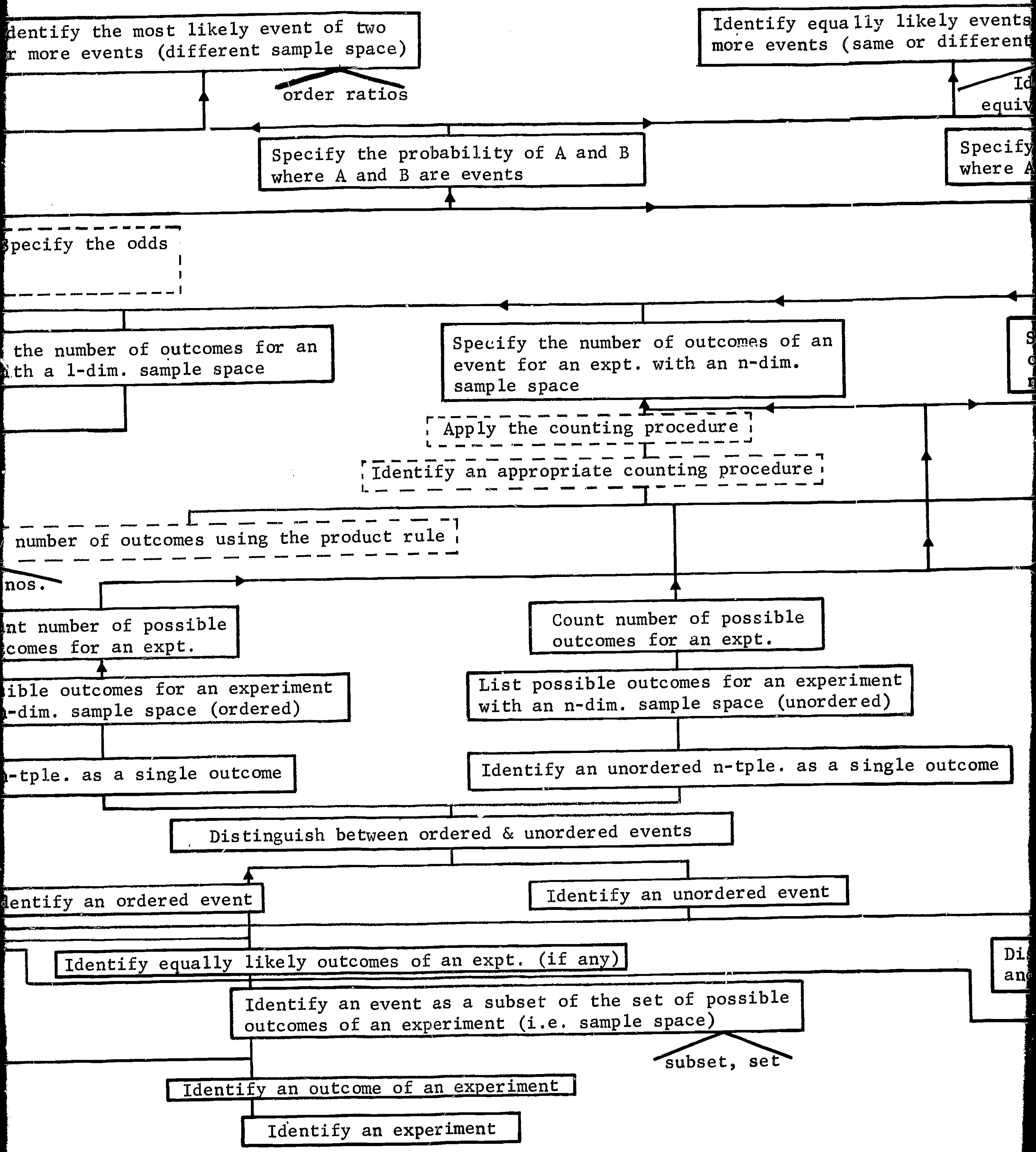
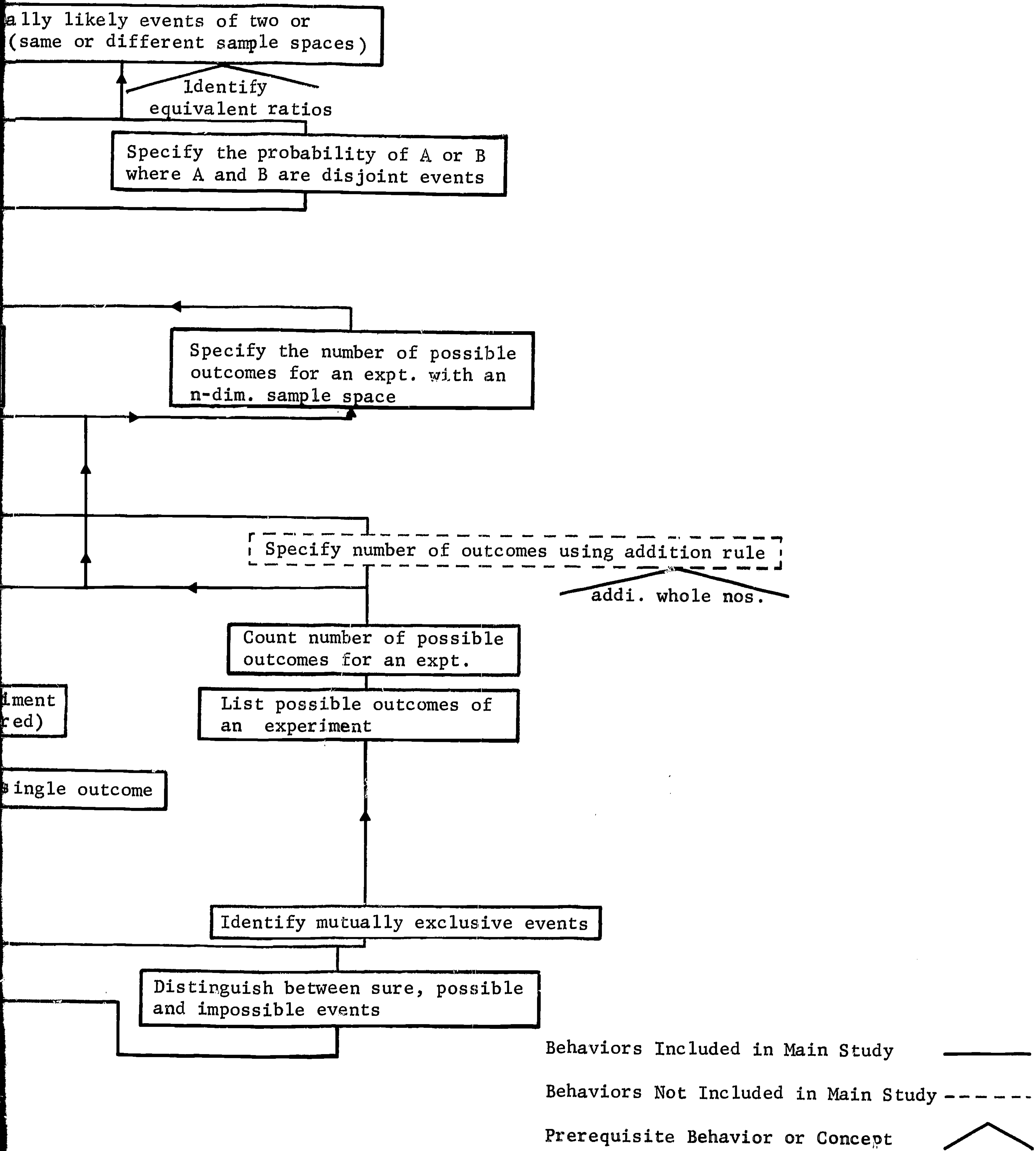



Figure 3. Task Analysis of Probability
And Statistics Concepts
For the Elementary School


dentify the most likely event of two
r more events (different sample space)
order ratios
Identify equally likely events
more events (same or different
Id
equiv
Specify the probability of A and B
where A and B are events
Specify
where A
Specify the odds
the number of outcomes for an
th a 1-dim. sample space
Specify the number of outcomes of an
event for an expt. with an n-dim.
sample space
Apply the counting procedure
Identify an appropriate counting procedure
number of outcomes using the product rule
nos.
nt number of possible
comes for an expt.
Count number of possible
outcomes for an expt.
ible outcomes for an experiment
-dim. sample space (ordered)
List possible outcomes for an experiment
with an n-dim. sample space (unordered)
-tple. as a single outcome
Identify an unordered n-tple. as a single outcome
Distinguish between ordered & unordered events
entify an ordered event
Identify an unordered event
Identify equally likely outcomes of an expt. (if any)
Di
an
Identify an event as a subset of the set of possible
outcomes of an experiment (i.e. sample space)
subset, set
Identify an outcome of an experiment
Identify an experiment


ally likely events of two or (same or different sample spaces)
Identify equivalent ratios
Specify the probability of A or B where A and B are disjoint events
Specify the number of possible outcomes for an expt. with an n-dim. sample space
Specify number of outcomes using addition rule
addi. whole nos.
Count number of possible outcomes for an expt.
iment red)
List possible outcomes of an experiment
single outcome
Identify mutually exclusive events
Distinguish between sure, possible and impossible events
Behaviors Included in Main Study
Behaviors Not Included in Main Study
Prerequisite Behavior or Concept

GPO 814-476

b. compound event

1. p(A or B)

2. p(A and B)

c. certain event

d. impossible event.

5. Distinguish the order of two fractions.

6. Identify the most likely event of two unequally events.

7. Identify two equally likely events as being equally likely.

8. Specify the estimated probability of an

a. unequally likely outcome

b. equally likely outcome.

* 9. Identify the likely bounds on the frequency of an outcome experiment that is performed n times.

* 10. Identify an instance of the law of averages.

* 11. Identify an estimate of the true probability given a set of data from an experiment.

12. Identify experiments which are equivalent (i.e., those with equivalent sample spaces).

13. Construct a bar graph of data from an experiment.

14. List the possible outcomes of an experiment by employing a tree diagram.

* 15. Identify that the probability of an event E, p(E), tells one that for a large number of trials the event will occur in a ratio approximately equal to p(E).

* [The starred objectives were not task analyzed before instruction.]

* 16. Identify that for a large number of trials the estimated probability approaches the a priori probability.

* 17. Identify that for a large number of trials that the cumulative frequency of a more likely event should be larger than that of the less likely event.

18. Specify the likely bounds on the a priori probability of an event from the graph of the data from an experiment.

The mathematical analysis in this section depends heavily on the task analysis. However, the hierarchies of objectives are only hypothesized relationships and must be validated to be anything more than that.

## LEARNING HIERARCHIES

Doing a task analysis means one is building learning hierarchies. The paths in Figure 3 represent a chain of learning hierarchies.

As stated by Walbesser (1968):

> A learning hierarchy consists of a terminal behavior, its identified subordinate behaviors, and the hypothesized dependencies among these behaviors. . . . Behavioral objectives can be ordered as steps in one or more learning hierarchies that represent a progression of intellectual development of the individual. Learning hierarchies defined by a succession of behaviors (described by the objectives) in instructional units constitute a set of hypotheses about intellectual development which may be empirically tested. These are hypothesized learning dependencies. This hypothesis states that if the learner has acquired the more complex behavior there is a high probability that he has already acquired all the behaviors named as subordinate (pp. 198-200).

Gagnè (1967) demonstrates the concept of a learning hierarchy and gives an example of its usage. Walbesser (1968) has produced a paradigm for testing hypothesized hierarchies.

In this study the following principles in the learning hierarchy will be tested. (1) In order to learn $\boxed{M}$ (identify the most likely event of two or more events or to identify equally likely events) the student must learn $\boxed{P}$ (specify the probability of an event) and $\boxed{R}$ (be able to order two ratios.) (See Figure 4.)

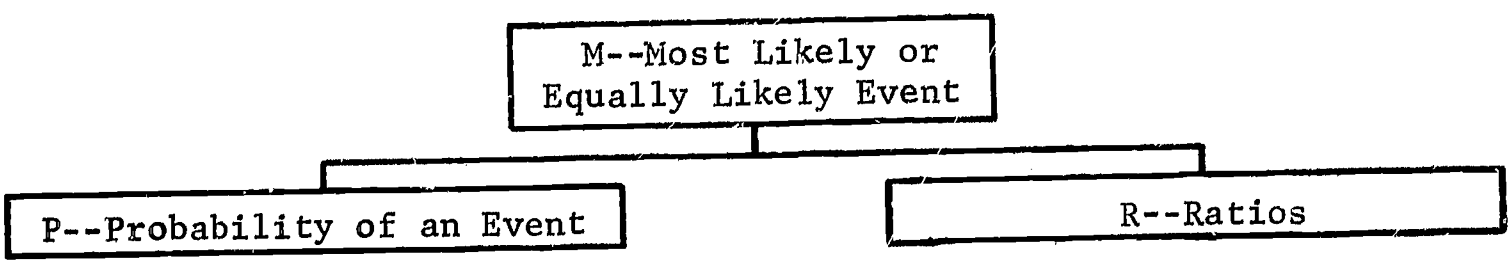



Figure 4. Learning Hierarchy 1

(2) In order to learn $\boxed{P}$ (specify the probability of an event) the student must learn $\boxed{E}$ (specify the number of outcomes of an event) and $\boxed{S}$ (specify the number of outcomes of the finite sample space). (See Figure 5.)

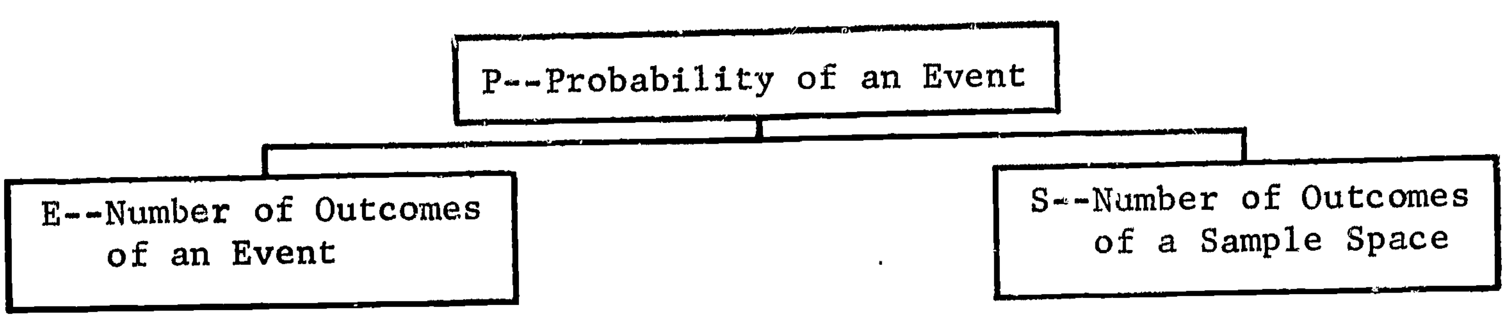



Figure 5. Learning Hierarchy 2

However, there are possible difficulties with the first hierarchy.

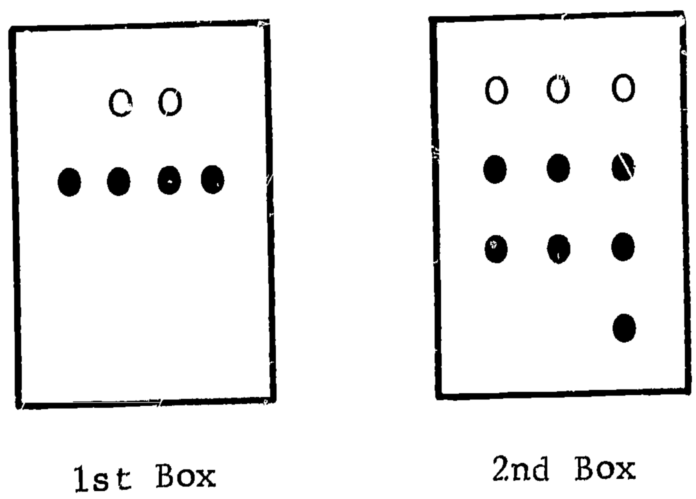



Figure 6. A Sample Problem in Deciding Between Two Boxes

For example, to decide whether picking a white ball is more likely from the first box or second box in Figure 6, some students from Shepler's (1969) status study and the pilot study (see p. 49) have been observed to use the following strategy. When a white marble was in a 1:2 ratio with the black balls, there was one black ball left over in the second box, and they observed that there are more ways to lose in the second box. Hence they chose to pick a white ball from the first box. This strategy involved combining the use of ratios and of absolute differences in the number of losers.

Thus the learning hierarchy for finite sample spaces might be the following:

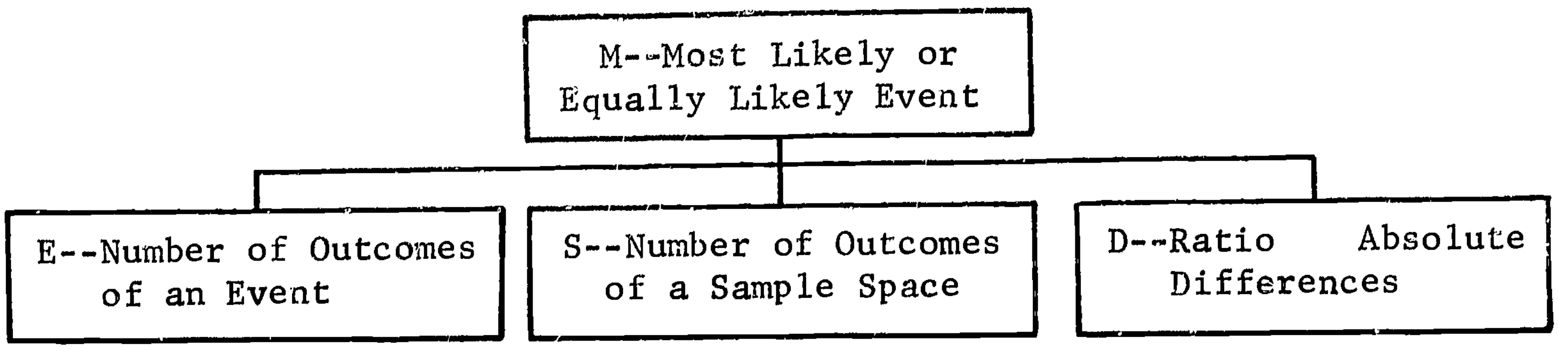



Figure 7. Learning Hierarchy 3

(D denotes a complex process of specifying the absolute difference between the number of favorable outcomes (unfavorable) of two events after applying ratios to matching the favorable and unfavorable outcomes of two events.)

Data was to be gathered for principles $\boxed{M}$ , $\boxed{P}$ , $\boxed{R}$ , $\boxed{E}$ and $\boxed{S}$. Using the procedures developed by Walbesser (1968), this data should give evidence as to whether the first two learning hierarchies are valid.

## INSTRUCTIONAL ANALYSIS

Instructional analysis is the second stage in development. Here attempts are made to conceptualize how best to teach each learning element of the task analysis. Gagnè (1965) states that "a student is ready to learn something new when he has mastered the prerequisites . . . . Planning for learning is a matter of specifying and ordering the necessary prerequisite capabilities within a topic to be learned . . . (p. 25)."

The assessment of the prerequisite behaviors of an individual is important. Individual differences may necessitate different instructional tasks or materials or organizational context of the instruction. From such considerations, including evidence from the pilot study, the instructional analysis and the daily lesson plans were created to form the instructional unit for the main study.

The basic instructional procedure for both studies was to approach probability concepts in an intuitive fashion where the student was to be actively involved using physical models. He was to gather empirical data from experiments and interpret the results. The student was also expected to empirically validate major objectives.

### PILOT STUDY

To aid in the development of materials and in deciding the appropriateness of these materials for sixth grade students, a pilot study was conducted in the fall of 1968. Six students (four girls, two boys) from Huegel Elementary School in Madison were taught probability concepts by the author. The pilot consisted of 12 sessions lasting between 40 and 45 minutes, plus one testing session at the end. At times, concepts were presented in different ways with different models. Students were asked for their subjective evaluation of the procedures. The goal of the pilot was not mastery learning of the concepts presented, but to aid in the selection of materials and procedures for the main study.

The following is a brief description of the lessons employed in the pilot. Many of the lessons were based on ideas or materials from SMSG's Probability for the Intermediate Grade (1966).

Lesson:

1. Discussion of the terminology of probability (e.g., certain, uncertain, more likely, experiment, etc.). A major purpose of the lesson was to identify instances of certainty, uncertainty, or impossibility and to list outcomes of simple experiments.

2. Carrying out of an experiment. The second lesson consisted of students actually doing five simple experiments discussed in lesson one.

3. Graphing data from an experiment. This lesson consisted of organizing and making bar graphs of the results from the experiments in lesson two.

4. Discussion of graphs and the carrying out of new experiments. The students were asked to discuss the results of their graphs. New experiments were carried out as in lesson two.

5. Graphing data from an experiment. This lesson consisted of organizing and graphing (bar graphs) the results from an experiment in lesson four.

6. Specifying the probability of an event. The lesson centered on the assigning of a probability number to a simple event and to an event with more than one element.

7. Choosing whether two one-dimensional events were equally or unequally likely. The lesson applied the idea of specifying the probability in deciding whether two one-dimensional events were equally or unequally likely. The ordering of two rational numbers was discussed.

8. Carrying out an experiment. Experiments using the problems from lesson seven were carried out to see if, over a large number of trials, the event with the larger probability occurred more often.

9. Graphing data from an experiment. An attempt was made to graph the data from lesson eight to show how a cumulative line graph of the relative frequencies would approach the a priori probability, and that there is less variation in the former graph when it is compared to a line graph of the relative frequency of each committee.

10. Specifying the probability of the complementary event. The lesson centered on the assigning of a probability number to the complement of an event.

11. Probability of an event. The lesson centered on the assigning of a probability number to events from a two-dimensional sample space. Also problems calling for assigning probability numbers to

"and" and "or" statements were given to the students for independent study.

12. Counting methods. Counting methods employing trees and tables were introduced to list and count two and three-dimensional outcomes.

13. Product rule. The lesson centered on using the product of the number of elements from two or more sets to find all the possible arrangements of these sets.

14. Review. A review of counting procedures and the assigning of a probability number to an event was emphasized.

15. Test. A modified form of Leake's (1962) test was used to measure, in some sense, the results of the instruction. The concepts measured were the ability of a student to:

(1) list or count the sample points from ordered and unordered events.

(2) specify the probability of a simple event.

(3) specify the probability of the union of mutually exclusive events.

(4) identify if two events are equally likely or if one event is more likely than the other.

Notes were written upon completion of a lesson. The most important observations were the following:

1. Too many imposed experiments in a short period of time bored some students.

2. The initial materials were too easy for these sixth-grade students.

3. The students had no experience in making graphs. However they did have experience in reading graphs. This may be typical of sixth-gradestudents and would need to be ascertained by pretesting.

4. The numerical skills of these students were average to poor. Problems such as changing 41/75 to a decimal or multiplying two digit numbers caused them particular difficulties. The difficulity in division resulted in the elimination of the lesson on cumulative frequency line graphs. These manipulative skills also have to be pretested.

5. The amount of material one can cover effectively in three weeks was less than expected.

6. A change in the initial task analysis was needed. New strands were needed that distinguished one-dimensional sample spaces and events from n-dimensional sample spaces and events, and that distinguished between ordered and unordered events.

7. Students should have had more practice in using newly learned skills in less structured situations.

The results from the pilot showed strengths and weaknesses in instruction and in the test that was used. The results of the pilot testing are contained in Table 1.

The instruction had been moderately successful as measured by this test. However, the test contained many interdependent items and transfer items. It was not a criterion-referenced test. The test also measured only a few of the instructional objectives. No measure of the notions of subjective probability, graphing, and

TABLE 1

POSTTEST RESULTS OF PILOT STUDY

| SUBJECT | IQ LORGE THORNDIKE | CONCEPT 1 TEN ITEMS | CONCEPT 2 TEN ITEMS | CONCEPT 3 TEN ITEMS | CONCEPT 4 THREE ITEMS |
|---|---|---|---|---|---|
| 1 | 142 | 68% | 80% | 60% | 100% |
| 2 | 134 | 90% | 100% | 70% | 100% |
| 3 | 120 | 83% | 70% | 70% | 100% |
| 4 | 119 | 100% | 90% | 60% | 67% |
| 5 | 115 | 76% | 100% | 40% | 100% |
| Mean | 126 | 83.4% | 88% | 60% | 93% |

Overall average--80 percent

Concept 1. List or count the sample points from ordered and unordered event.

Concept 2. Specify the probability of a simple event.

Concept 3. Specify the probability of the union of mutually exclusive events.

Concept 4. Identify if two events are equally likely or if one event is more likely than the other.

analyzing the results of experiments were included in this test. Aside from this, many items were of a transfer nature in that the models were quite different from those used in the study. For example, no blocks other than dice were used in the pilot, while the test used six-sided, ten-sided, and twelve-sided blocks. Some problem involved combinatorial counting procedures which were not explicitly covered in the instruction. The results of these differences were particularly reflected in the results of Concept three, probability of mutually exclusive events. This concept was not explicitly mentioned in the instruction, although students did find the probability of events involving more than one outcome.

The analysis of the errors made on this test had implications for changes in the instruction. With respect to concept (1), sample space, the listing of results of flipping three different coins seemed to be a problem. "Trees" and tables were introduced in the instruction, but no student saw them as appropriate tools to use for this problem. Students need more experience using these tools with diverse problems. Another common mistake, that of adding rather than multiplying to find the number of possible outcomes, necessitates a stronger treatment so that students discriminate when the product rule is appropriate. Students also need more practice using ordered and unordered events.

With respect to the probability of a simple event, some students made the error of saying the probability of getting a "2" when throwing a 12 sided die is 2/12. With respect to Concept 3, the probability of mutually exclusive events, the mistakes were in counting the number of possible outcomes and/or the number of outcomes in the event. The mistakes were a confounding of mistakes in Concept 1, adding instead of multiplying to find the total number of outcomes and not considering all possibilities with respect to outcomes of an event.

With respect to three problems measuring Concept 4, deciding between two boxes, only one person missed a problem--and that was an easily explained mental error. However only one person assigned probabilities to the two events and then ordered the probabilities. All of the others used a combination of ratios and absolute differences. Items calling for the student to choose between two one-dimensional boxes were not of sufficient difficulty to force the need for assigning probability numbers first.

From the results of the pilot study the following modifications were incorporated in the instructional plans for the main study.

1. There would be a need to ascertain whether students knew how to construct and interpret bar graphs. If not, these skills would have to be taught.

2. Because of the student's inability to change numbers such as 41/75 to a decimal, it was decided that some students would be taught to use either a slide rule or a desk calculator as a way for converting fractions to decimals.

3. The amount of material to be covered was cut back. A lesson on the probability of the complementary event and the explicit stating and use of the product rule was dropped.

4. Only six experiments using various models would be used in Lesson 2 (Performing of experiments in 1-Dimension).

5. Exercises which were not highly structured were to be included.

6. In the light of the analysis of the test, more emphasis would be placed on the use of a tree in a counting situation.

7. Distinguishing between one-dimensional and two-dimensional problems would be stressed.

Based on this information, the following instructional plan was designed for the main study. To aid one in viewing the entire plan the following figure presents the instructional plan for teaching the probability unit.

## DESCRIPTION OF INTENDED LESSONS

The instructional plan including the nine lessons to be taught is based on the outline in Figure 8. The purpose of each lesson is briefly summarized. The complete set of lessons are found in Appendix A.

In each of the nine lessons the following are included:

(1) The behavioral objectives of the lesson.

(2) A list of important prerequisite behaviors needed for the lesson.

(3) The materials to be used in the lesson.

(4) New vocabulary introduced in the lesson.

(5) A description of the method of presentation of the lesson.

(6) The student's exercises for that lesson.

| | |
|---|---|
| Day 9 - | Quiz |
| Day 8 - | Lesson 5--Probability of an Event (2-D) |
| Day 8 -<br>Day 7 - | Lesson 4--Probability of an equally likely outcome (1-D) |
| Day 6 -<br>Day 5 - | Lesson 3--Graphing and Interpreting of results from the experiments |
| Day 4 - | Lesson 2--Performing of Experiments in 1-D |
| Day 4 -<br>Day 3 - | Lesson 1--Introduction to Probability<br>Subjective notions and vocabulary<br>Experiments with Spinners and marbles |
| Day 3 -<br>Day 2 -<br>Day 1 - | Pre Instruction - Bar Graphing |
| Day 2 -<br>Day 1 - | Pretest |

Figure 8. Instruction Plan for the Probability Unit--Anticipated Number of Days (20-22), 50 minutes per day

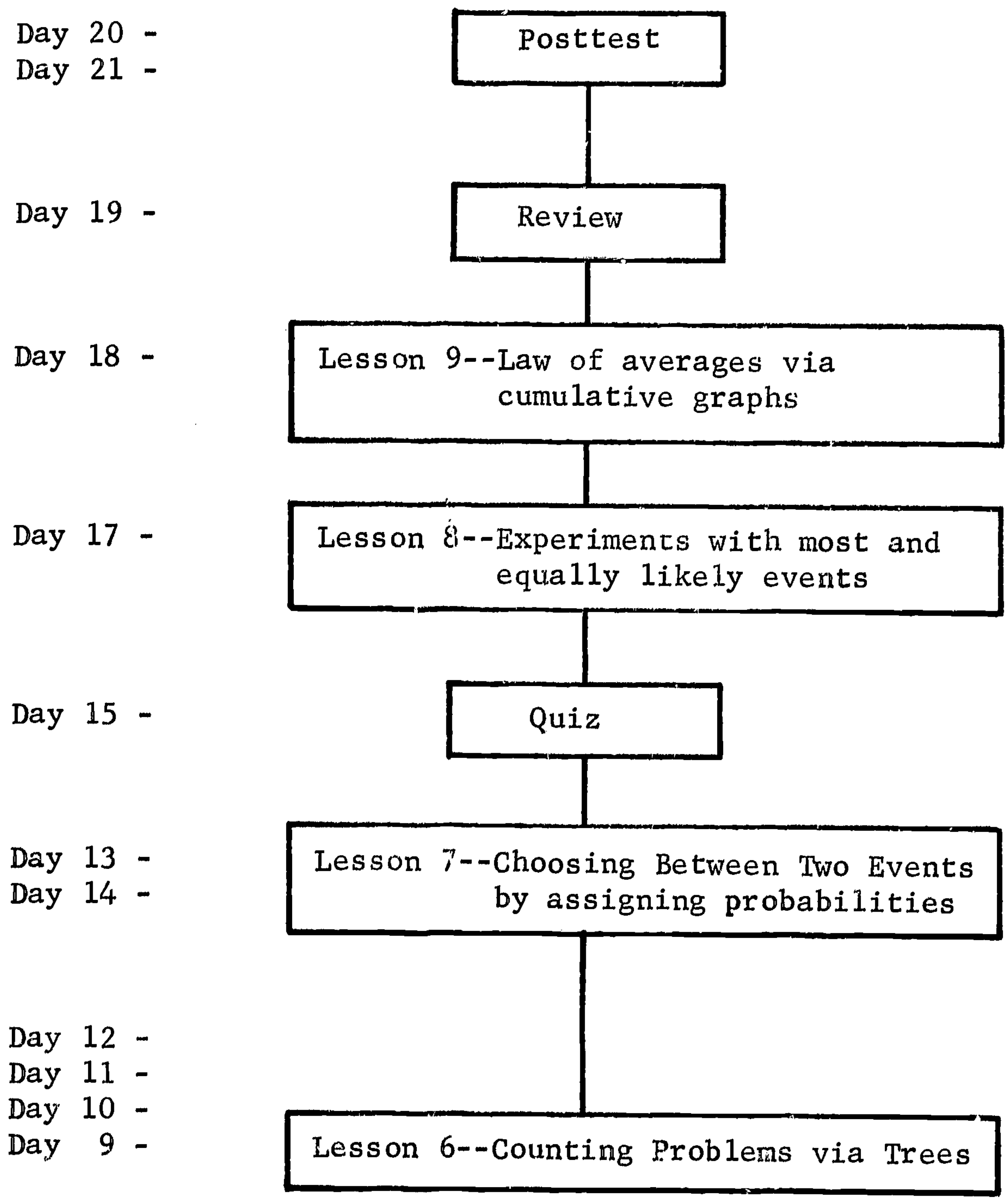



Figure 8. (Cont'd)

Summary of Lessons

Two days were to be set aside for the pretesting. Pre-instruction on bar graphing was planned and would be given after each of the testing sessions during the first two days. The goal of the pre-instruction was to enable students to plot and interpret bar graphs. Instruction on the prerequisite skill of ordering two ratios was also planned. Its use would depend on students' performance on the pretest. If student performance was good (80 percent or better), very little time would be devoted to this prerequisite skill. If needed, it would be done in connection with Lesson 7. Here the student would be able to see why it was important for him to order two ratios in making the best decision between two games.

The first lesson was designed to introduce much of the vocabulary of the unit, using the models that were to be employed in future lessons. Also, the students' concepts of certainty, uncertainty, and impossibility were to be strengthened by class discussion and a written exercise.

The second lesson called for the students to take the models introduced in the first lesson and to carry out six experiments with them. The data from these experiments were to be recorded.

The third lesson was to involve students using the recorded data of Lesson 2 to construct a graph and interpret the results. Both the individual results and the cumulative results were to be graphed for each experiment. From the interpretation of the graphs the students were to observe the following:

(1) The results of an experiment are uncertain.

(2) Two events may or may not be equally likely.

(3) There may be a pattern in chance events over a large number of trials.

The fourth lesson was designed to build on the students' empirical experience with the models by introducing the concept of the probability of an event. The probability of the impossible event, the certain event, and the notation denoting the probability of an event was introduced. (E.g., P (R) = 1/2 could mean that the probability of getting red in one draw from a box is 1/2.) All models and problems in this lesson were to be restricted to models having one-dimensional sample spaces.

Lesson 5 was to introduce probability problems in a two-dimensional sample space. All the problems in this lesson and the exercise accompanying the lesson list the sample points or ask the student to complete the list of sample points.

At the end of Lesson 5, a quiz was to be given testing the more important objectives of lessons 1-5.

Lesson 6 consisted of two parts. By introducing a model which students should not be able to list all the sample points, the need for a better counting procedure was to be created. Lesson 6-Part I then introduced the concept of a tree diagram as a devise useful in counting the outcomes of problems in two and three-dimensional sample spaces. The first exercise for Part I was highly structured with the tree and a table partially drawn. The second exercise for the lesson asked the student to draw his own trees to count the number of possible outcomes. Part II introduced the dice model and a dice game. The exercise accompanying this lesson was designed to give the student practice with probability problems involving dice and a variety of models having two- or more-dimensional sample spaces.

Lesson 7 also consisted of two parts. Part I introduced the idea of using the probability number to make the best decision between two one-dimensional box models containing colored marbles. Part II extended the use of the probability number to making a decision between two games, one of which was a one-dimensional model and the other a two-dimensional model.

Lesson 8 called for students to verify that they had made the best decision when they had employed the strategy of ordering the probability numbers to make their choice between two games. This was to be done by students carrying out experiments and collecting data. Also, using outside activities, the students were to employ certain models (cups, tacks, spinners, etc.) to gather data for analysis in Lesson 9. Relative frequency bar graphs of individual trials and of cumulative frequency of the trials were to be plotted. The concept of the estimated probability (i.e., relative frequency) of an event was to be introduced and used in making the relative frequency bar graphs. The graphing of the estimated probabilities required changing numbers such as 41/75 or 87/150 to their decimal equivalents. To do this, some students were to be taught how to use a slide rule or a desk calculator. A prerequisite skill necessary for the graphing of these values is the ability of the student to order two decimal fractions. The amount of time available would decide how much work would be done with the estimated probability as a decimal and the amount of graphing of these values that could be done.

Lesson 9 asked the students to interpret their bar graphs of the data gathered in Lesson 8. By analyzing the results of the cumulative frequency graphs of an experiment the students were to

(1) Identify that the probability of an event, p(E), tells one that for a large number of trials, the event will occur in a ratio approximately equal to p(E).

(2) Identify that, for a large number of trials, the approximate probability approaches the a priori probability.

(3) Specify the likely bounds on the a priori probability from the graph.

(4) Distinguish instances of the law of averages. By comparing two graphs the students were to observe that for a large number of trials the cumulative frequency of a more likely event should be larger than that of the less likely event. From the exercise, the students were to learn to distinguish instances of the law of averages.

The day before posttesting was to be spent reviewing. The last two days were to be set aside for administering the posttest.

## MASTERY LEARNING

A major goal of the experiment was to demonstrate students in the class could master the objectives of the unit. Bloom (1968) believes most students can master what is to be taught them and that it is the task of instruction to find the means to get this mastery. Bloom cites Carrol's view that "aptitude is the amount of time required by the learner to attain mastery of a learning task. Implicit in this formulation is the assumption that given enough time all students can conceivably attain mastery of a learning task (1968, p. 3)."

For this study, the ultimate level of mastery for a behavioral objective was set at 90/90. This 90/90 criterion means that if 90 percent or more of the group achieve 90 percent or better on a behavioral objective then instruction is judged to be successful for that objective.

No doubt some students will need more time, effort and help than others to achieve this high level of mastery of an objective. However, if mastery is to take place a very important factor for a given learner is his perserverance (the time the learner is willing to spend in the learning task). Bloom states that "frequency of reward and evidence of success in learning can increase the student's perserverance in a learning situation (1968, p. 7)." As a means of decreasing the amount of perserverance necessary for learning a given task, Bloom recommends frequent feedback accompanied by specific help.

In order to achieve mastery learning of the behavioral objectives in this unit the following ideas were incorporated into the instructional procedures:

1. Specification of the objectives of the unit.

a. Goal chart—A large chart of the important objectives of the unit was to be referred to frequently to show how the activities of a given day were aimed at meeting a particular objective of the unit. (See Appendix A p.149)

Bloom feels the ability to understand instruction really is the ability of the learner to understand the nature of the task he is to learn and the procedures he is to follow in the learning of the task. The goal chart was one way the author endeavored to get the students to understand the purpose of a particular lesson.

2. Extrinsic Rewards

a. Grading—All exercises and quizzes were to be graded Masters (M) or Nonmaster (NM). This idea for grading is based on Bloom's feeling that the use of the grades A, B, C, D, and F on progress tests prepares students for accepting less than mastery.

(1) Mastery. "Master" usually meant scoring 90 percent or better.

(2) Nonmastery.

(a) If a student was a nonmaster on an exercise he was to given the opportunity to correct his mistakes and re-submit his paper. If he had corrected his mistakes successfully he then was to be counted as a master for that exercise.

(b) If a student was a nonmaster on a progress quiz he was to be given an opportunity, after further instruction and help, to take a parallel quiz. If he was graded as a master the second or third time he was to be counted as a master for the unit measured by that quiz.

b. Diploma—The students were to be told that if they are a master of 90 percent or more of the graded papers that they would receive a diploma saying that they are a Master of probability concepts in the elementary school. (See Appendix A p.361.)

3. Formative Tests—These were to be given periodically and graded Master or Nonmaster. These tests were to be used to assess individual deficiencies and also the effectiveness of instruction.

4. Practice Exercises— The student's homework was to be frequently graded and marked Master or Nonmaster. If he was a Nonmaster he was to be told that if he corrected his mistakes he would be classified as a Master for that exercise.

5. Prescriptions—Specific prescriptions were to be written on the Nonmaster's exercises and quizzes pinpointing the nature of his mistakes and suggesting what he could do to correct them.

6. Extra Help Sesions—If a sufficient number of people had not been successful after the second iteration of instruction for a set of goals, extra help was to be given after the regular class by the teacher and the author. This was to be a small

group or a tutorial situation. Also at times, the class was to be split into masters and nonmasters with the nonmasters receiving further help during the class period. (The author was to work with one group while the teacher worked with the other.)

7. Master and Nonmaster Teams—For the persistent Nonmaster, a master learner was to be assigned to him to help him outside of class.

8. Prerequisite skills—Diagnostic techniques were employed to try to ascertain whether most students had the preprequisite skills before introducing a task dependent on those skills.

9. Iteration of the objectives—If the instruction resulted in fewer than 80 percent of the students obtaining mastery at the first time of testing, further group instruction was to be given on the objectives measured by that quiz.

By using the previously mentioned procedures, this study intended to demonstrate that most of the group could achieve mastery learning on most of the measured behavioral objectives of the unit.

CRITERIA

There are two criteria to be used in judging whether the instructional treatment has been successful in getting students to demonstrate mastery of a behavioral objective. The first is a 90/90 criterion referred to earlier. This 90/90 criterion was to be the ultimate goal of instruction. However, this arbitrary criterion is a function of the size of the group and of the number of measures of the objectives.

For example, with a group of 30 students, three or less students can score below 90% on a test and instruction be judged successful, while for a group of 25, two or less students can score below 90% on a test. Thus the standard actually becomes a 92/90 criterion for a group of 25.

The criterion also is a function of the number of items used to measure the behavior. For example, for a test of an objective with one to nine items, students must get all items right to score 90% or better on the test. Students who may have mastered the objective could easily miss one item by error and be falsely classified as a nonmaster of the behavioral objective.

For these reasons the author is also adopting a practical criterion based on the sample size and the number of items used to measure an objective. Since there were 25 students used in the main study (See Chapter III, p. 75) and the number of items (n) used to measure an objective varies between one item and eighteen items the following practical criteria was adopted.

For $1 \leq n \leq 4$, 88/100

For $5 \leq n \leq 13$, $88/\frac{n-1}{n}(100)$

For $14 \leq n \leq 20$, $88/\frac{n-2}{n}(100)$

To illustrate the criteria, consider a test of an objective that contains eight items. Then 22/25 students or better would have to score 7/8 or better for instruction to be judged successful for the objective. This criterion is 88/88. The second number in the criterion $\left[\left(\frac{n-1}{n}\right) \times 100 \text{ or } \left(\frac{n-2}{n}\right) \times 100\right]$ will be referred to later as the practical test criterion. A test used to determine whether a student is a master or nonmaster of a behavioral objective according to the set criterion will be referred to as a criterion-referenced test.

## FORMATIVE EVALUATION

Evidence gathered by this study for formatively evaluating the unit was to lead to an ongoing modification of the instructional plan to meet the needs of the learner and to redesign the unit. The evidence gathered was to be used to test the feasibility of teaching topics in probability and statistics to a class of sixth grade students. To gain this information, pretesting and posttesting of the group was planned, using an instrument that was specifically developed for this purpose.

The pretest was to measure what knowledge the children had of 14 objectives before instruction, including whether they possessed the prerequisite skill of ordering two fractions. The posttest results were to help in reanalyzing the unit in order to make suggestions for modifications in the unit and in testing the feasibility of the study.

Two quizzes were also planned; the first quiz was to measure certain objectives of lessons (1-5) and the second quiz was to measure certain objectives of lessons 6 and 7. Also exercises that accompany a lesson were to be used to gain additional information.

Test

The test used in the pretesting and posttesting consists of 72 items; of which 36 measure one-dimensional sample space problems, 19 measure two-dimensional sample space problems, 7 measure one-dimensional and two-dimensional sample problems, and 10 measure the ordering of two fractions. (See Appendix B.) The items on the test are based on one- and two-dimensional finite sample spaces generated by models using coins, dice, spinners, and boxes with objects.

The behavioral objectives included in the test and the number of items employed to measure these are the following:

The child should be able to:

1. distinguish whether an event is an instance of certainty, uncertainty or impossibility (5 items).
2. count the number of outcomes of an event (5 items).
3. count the number of possible outcomes of a sample space (7 items).
4. specify the probability of a (n):
   a. simple event (8 items).
   b. compount event (8 items).
   c. certain event (2 items).
   d. impossible event (2 items).
5. specify the order of two fractions between 0 and 1 (10 items).
6. identify the most likely event of two unequally likely events (10 items).
7. identify two equally likely events as being equally likely (8 items).

8. specify the estimated probability of an event, given the data from an experiment (1 item).
9. identify the likely bounds of the frequency of an outcome of an experiment that has been done n times (1 item).
10. identify an instance of the law of averages (4 items).
11. identify an estimate of the true probability, given a set of data from an experiment (1 item).

The 72 items were divided into 14 parts labeled R1, R2, . . ., R14. The test was designed to be given in two testing sessions. Section A, to be administered first, consists of parts R2, R4, R6, . . ., R14. Section B consists of parts R1, R3, R5, . . ., R13. Each section was assigned one of four random orders from a random number table. The items were then assembled to form section A or section B. This randomizing of the orders of section A and B was done to rule out the effect of taking a test in a specific order and to minimize the opportunities for cheating.

The same set of directions were attached to the front of each section. The directions were to be read to the students after the test was passed out. Models involved in the test were to be displayed briefly at that time. The students were to be given an unlimited amount of time to complete the test.

## SUMMARY OF CHAPTER

Chapter II described the mathematical and instructional analysis employed to develop a unit in probability and statistics for sixth-grade students. A description of the planned formative evaluation was also included. Chapter III describes the design of the study. It also describes how the unit was used to teach a group of sixth grade students and how the formative evaluation procedures were employed.

# CHAPTER III

# THE DESIGN AND CONDUCT OF THE STUDY

The first stated purpose of this study was to test the feasibility of teaching certain topics in probability and statistics to sixth-grade students. To do this feasibility study, one needs an instructional program which includes materials, procedures and tests. Development of these parts was described in Chapter II. Chapter III presents the design of the study by using an instructional systems model. Also included is a brief account of the conducted study. The detailed day by day account of the study is contained in Appendix A.

## Instructional System

This study is viewed as the development and refinement of an instructional system in probability and statistics for the elementary school. The elements in an instructional system used in this study are based on Romberg's (1969) model. (See Figure 9.)

The basic elements in this model of an instructional system are:

1. Input: Students with entering behaviors
2. Resources: Staff, facilities and equipment
3. Mechanism: Instructional program
4. Feedback: Evaluation and decision procedure
5. Output: Students with terminal behavior

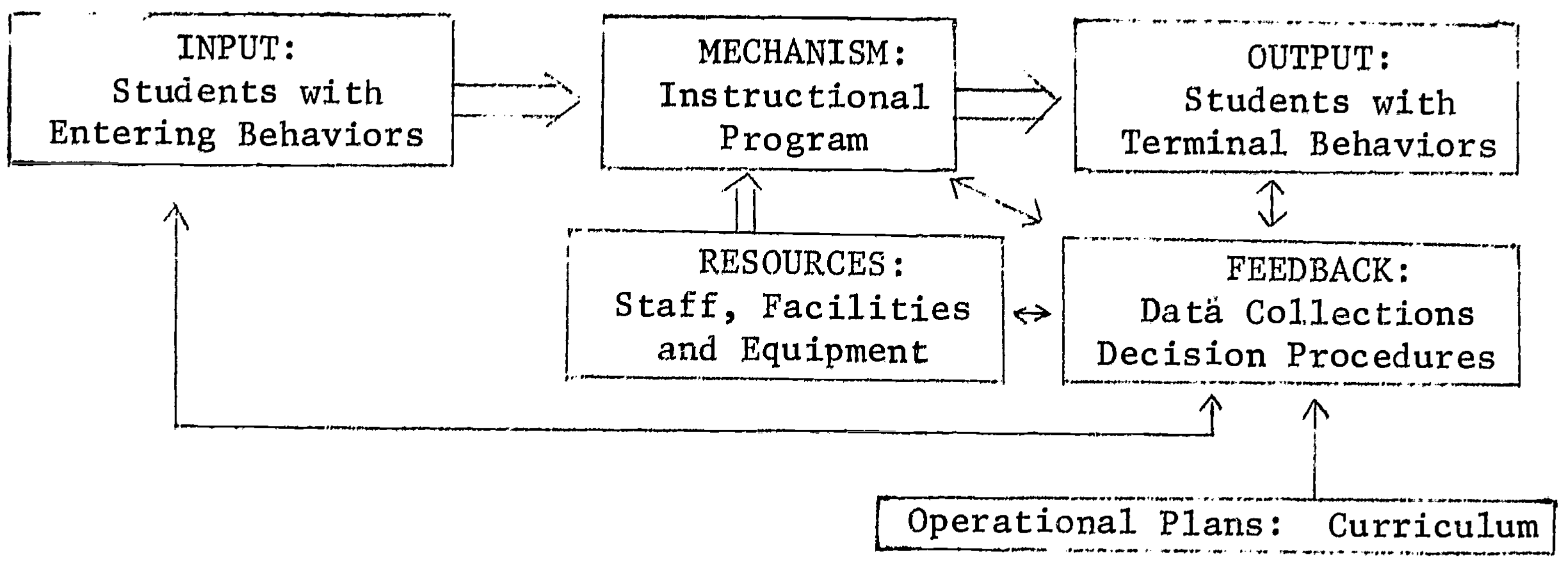



Figure 9. The Basic Elements of an Instructional System

INPUT

To describe student's entering behaviors the following are considered: (a) the population from which the students were chosen, (b) a description of the procedure for choosing the sample from the population, and (c) the student's performance on standardized tests.

Population

The study was conducted in the Waunakee Elementary School, Waunakee, Wisconsin in March and April, 1969. The Waunakee School District comprised an area of approximately 40 square miles and had a school population of approximately 1000 pupils. The Waunakee Elementary School housed 525 students in a modern, well equipped building.

The village of Waunakee is a rural, suburban community 12 miles north of Madison, Wisconsin. The area is socioeconomically middle to lower middle class. Many of the people commute to Madison and are employed by industry or by a government agency. Yet the community's rural characteristics are in strong evidence since 35% of the students in the school district live on farms that surround the community.

The population for this study was chosen from the 67 students contained in two classrooms which comprised the entire 1968-69 sixth grade population of the Waunakee Elementary School. One of the two sixth grade teachers was the arithmetic teacher for both groups.

The Sample

The sample of 25 students was chosen from the population of 67 sixth grade students by the arithmetic teacher, guidance counselor, and elementary principal. The criteria employed by this group were:

1. Scores from The Iowa Tests of Basic Skills (Lindquist and Hieronymus, 1964). These scores included the three separate arithmetic scores and a composite score. If these scores were not above average, the reading score from the Iowa Tests was also considered.
2. Teacher recommendations. The sixth grade teacher chose the top 25 students of both groups based on classroom performance, the child's ability to complete assignments, and the Iowa tests.
3. Study Habits. The criteria of initiative, enthusiasm, and overall performance were considered.
4. A remedial class in arithmetic skills was conducted concurrently with the experimental class by the two sixth grade teachers. The remedial class consisted of all sixth grade students not in the experimental class. If students fell into

the doubtful category on the basis of criteria 1, 2, and/or 3, the question was asked if they would profit more by taking the remedial class. If so, they were chosen for the remedial class.

5. Student Inclination. The children chosen were invited to participate. They were told that they did not have to participate and that there were other children who could take their place if they did not want to take part in the experimental unit.

Table 2 presents the means and variances for the five Iowa Test scores of students selected. The Iowa Tests were administered in the Fall of 1968. The key used by the school in interpreting these scores was that a score from (1 - 35) is weak, (35 - 65) is average and (65 - 99) is strong.

TABLE 2

IOWA TEST SCORES OF THE 25 SUBJECTS IN THE STUDY

| | Arithmetic Concepts | Arithmetic Problems | Total Arithmetic | Reading Comprehension | Composite |
|---|---|---|---|---|---|
| Mean | 69.80 | 73.84 | 74.08 | 73.48 | 76.76 |
| Variance | 269.50 | 454.39 | 290.41 | 340.43 | 250.44 |

The I.Q. test used was the Lorge-Thorndike Intelligence Tests Level 3, Form A (Lorge and Thorndike, 1954) which was administered in April, 1968. (Three students with missing scores were given the test in

March, 1969.) The mean IQ for the group was 117.72, and the variance was 46.79. The average math grade for the first half year of the sixth grade was a B-. Two students selected had grade averages below C, a D-, and a D. (See Appendix D for data on each student concerning the Iowa Tests, IQ, and grades.)

The learners used in this study were average to above average in ability; they had no reading difficulty. These specifications were included in order to test the feasibility of the program with learners who do not have other problems which would affect their learning of probability concepts.

## RESOURCES

The resources for this study are (1) teaching staff, (2) the auxiliary staff and (3) the classroom facilities and equipment.

### Teaching Staff

The teacher used in this study, Mrs. Carolyn Gornowicz, was a Project Specialist on the R & D Center Staff and an experienced teacher. She was assisted by the author. She had a B.S. in elementary education along with two years of mathematics in her undergraduate training including course work in calculus and analytical geometry. The function of the teacher was to implement the planned lessons as precisely as possible. She was to perform the expository and inductive teaching functions as called for and to act as the classroom manager of the learning process. Besides instruction, she was to help in formulating the operational plan for the following day in light of an evaluation of the present day's lesson results, make the assignments, and help in the diagnosing and remediation.

The author acted as an observer, teaching aid (keeping track of the records of each student's performance), and resource person. As a resource person, he began training the teacher in the content of the unit three weeks before the study was to begin. As a teaching aid, he was responsible for diagnosing and planning remediation. Also, the author was to help individual students, to regrade exercises that had been corrected, and to work with small groups when the class was split into various size groups.

### Auxilary Staff

The auxilary staff included two sixth grade teachers in the school and the elementary principal. All were most cooperative and helpful. This cooperation helped create a sound environment for conducting an experimental class. The secretarial and clerical staff of the Research and Development Center helped in typing and duplicating exercises and in recording data so that decisions could be made concerning the next day's lesson.

### Classroom Facilities and Equipment

The classroom used was modern, and fairly well equipped. However, the room did not have blinds. This caused problems in using the overhead projector. Because of too much light, the teacher changed to using posters and the blackboard rather than using the overhead. Later, a more powerful overhead projector with a large screen was rented and, after properly placing the screen, this difficulty was eliminated.

A set of lessons covering the unit was given to the teacher. Various colored marbles, dice, numbered plastic chips, spinners of various types, small boxes to hold the models, thumbtacks, and paper cups were assembled and used in the lessons. These objects were used as demonstration models or as models for experiments that were conducted by the students. Large poster paper was used for graphing data and for displaying the goals of the unit. Rulers, magic markers, and graph paper were also used in connection with graphing. Typed copies of exercises, quizzes and tests were made and administered to the students.

## INSTRUCTIONAL PROGRAM AND FEEDBACK

The instructional program consisted of the classroom organization; the intended lessons, exercises, and quizzes; plus procedures for feedback to modify the lessons and instructional procedures so that mastery learning could be achieved.

### Classroom Organization

The lesson plans called for use of large group instruction (the whole class), and small group activities (two to four children) to carry out experiments. The class would be divided into two or three groups based on achievement. The divisions would segregate the group into masters and nonmasters of a set of behavioral objectives, while the three group divisions would parcel the group into masters, almost masters, and definite nonmasters (those needing a great deal of remedial training).

## Lessons

The intended lessons for these experiments are the set of nine lessons described earlier. However, as Romberg states, "The role of the operational plan is to relate intended and actual learnings (1969, p. 19)." Thus the operational plans were tentatively stated in the lessons, but they could not be the actual plans that would be followed. Because of the formative evaluation procedures, if the intended learning differed quite a bit from the actual learning in a negative way, the operational plan would call for the planning and implementation of a remedial treatment. Also, the nine lessons are not organized for daily use. The operational plan actually implemented had to take into account how one teaches these lessons on the basis of a 50 minute session, day to day basis. Some lessons were actually split up over two to four sessions. The operational plan had to coordinate this fragmentation of lessons to meet the intended learnings that had been planned.

## Evaluation

The evaluation of daily lessons, exercises, and quizzes provided information for decisions concerning individuals, the group, and planned lesson for the next day. Observation of the class in various-sized groups helped in providing information for making these decisions. The mastery learning procedures mentioned in Chapter II were also helpful in the decision making process.

Mastery Learning

All the suggestions for mastery learning mentioned in Chapter II were utilized to aid the instructional process and to help in the remediation needed for the individual learner. For example, after two weeks of instruction, three half-hour extra help sessions were conducted after class to aid individuals who were having a problem in becoming masters of a given set of objectives. When two nonmasters were consistently unresponsive to instruction and the extra help sessions, they were assigned to two master learners to help them whenever possible.

All exercises and quizzes that were collected were graded M (Master) or NM (Nonmaster). If a student was a nonmaster, a prescription was given that tried to point out his mistakes and what he needed to do to become a master. A nonmaster of an exercise simply corrected it and handed it to the experimenter for regrading. On a quiz a nonmaster was told he would have another opportunity to become a master of a set of objectives when the next quiz on those objectives was administered. Also, more group or individual treatment was given before the quiz was administered again.

The teacher frequently alluded to the goal chart to help students see what the major goals of instruction were and what goals a particular lesson was concerned with. The first day after the pretesting, the students were informed of the diploma to be awarded at the end of the unit to masters of the goals. They were reminded of the diploma three days before the posttesting.

Identified prerequisite skills were kept in mind before introducing a particular lesson. How prerequisite skills were used in making decisions, particularly in lessons 4 - 7, and the iteration procedures employed are discussed in the Journal.

OUTPUT

The last component of the instructional system, the output on the terminal behaviors is considered in Chapters 4 and 5.

## Conduct of Study

DESCRIPTION OF THE JOURNAL

The following information is given in the Journal in Appendix A to present the details of the instructional program that were used in the study.

1) The intended lesson.
2) A short description of the actual lesson and its evaluation.
3) An evaluation of the exercise included with the lesson, if any.
4) An evaluation of any quiz given.

If one reads the Journal carefully one can note from the intended lessons and the description that follows, the discrepencies between the intended instructional plan and the actual instructional plan. Altogether nine lessons and thirteen exercises were used. Five quizzes and the pretest and posttest were used to aid in the formative evaluation of the unit. Information from the quizzes and the exercises was used to modify the following day's lesson.

The author attempts to show how the information gathered was used to formatively plan the actual instruction that was implemented.

Figure 10 has been included to aid the reader in viewing the overall sequence of lessons, exercises and quizzes used in the study. All of this information is contained in the Journal (Appendix A).

## DESCRIPTION OF STUDY

The following is a brief account of the conducted study. It may give the reader the flavor of information gained from the pretest, exercises, quizzes and observations and how it was used to modify the instructional plan.

### Pretest

The pretest (72 items) was given in two sessions, the first lasting 35 minutes and the second lasting 25 minutes. The average pretest score in percent was 37.9%. Ten of the items on the pretest measured the prerequisite behavior of ordering two fractions. The average percentage on these 10 items was 78.8%. (This information was used in Lesson 7 - Part I.) Considering the pretest with these 10 items omitted, the 62 items remaining measured 13 of the major objectives of the study. The average percentage on the 62 items was 31.1%.

From an analysis of the results one can conclude that students had some prior knowledge of subjective notions of probability, probability problems in one-dimension (particularly in the specifying the number of outcomes), how to choose between two boxes, and the law of averages. They knew practically nothing about probability problems

| | Date | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 3/3 | 3/4 | 3/5 | 3/6 | 3/7 | 3/10 | 3/11 | 3/12 |
| <u>Pre Test</u> | | | | | | | | |
| Form A | X | | | | | | | |
| Form B | | X | | | | | | |
| <u>Pre-Instruction</u> | | | | | | | | |
| Reading a bar graph | X | X | X | | | | | |
| Constructing a bar graph | | | X | | | | | |
| Exercise (constructing a bar graph of children in a family) | | | X | | X | | | |
| <u>Lesson 1</u> | | | | | | | | |
| Subjective notions of Probability | | | X | | | | | |
| Vocabulary of probability | | | X | | | | | |
| Exercise (homework) 1.1 | | | X | | | X | | |
| <u>Lesson 2</u> | | | | | | | | |
| Review Lesson 1 | | | | X | | | | |
| Performing experiments in one-dimension | | | | X | | | | |
| <u>Lesson 3</u> | | | | | | | | |
| Graphing data from experiments | | | | | X | X | | |
| Interesting results | | | | | | | X | X |
| Exercise 3.1 | | | | | | | X | |
| Exercise (homework) | | | | | | | X | |
| <u>Lesson 4</u> | | | | | | | | |
| Probability of an event (one-dimensional) | | | | | | | | X |

Figure 10. Sequence of Lessons, Exercises and Quizzes

| | Date | | | | | | |
|---|---|---|---|---|---|---|---|
| | 3/13 | 3/14 | 3/17 | 3/18 | 3/19 | 3/20 | 3/21 |
| <u>Lesson 4</u> | | | | | | | |
| Dice game | X | | | | | | |
| Notation (e.g. P(R) = 1/2) | X | | | | | | |
| Exercise (homework) 4.1 | X | X | | | | | |
| <u>Lesson 5</u> | | | | | | | |
| Probability of an event (two-dimensional) | | X | | | | | |
| Exercise 5.1 | | X | | | | | |
| <u>Lesson 6 (Part I)</u> | | | | | | | |
| Introduction to use of trees in counting | | | X | | | | |
| Quiz I (Lessons 1-5) | | | X | | | | |
| Exercise 1 (6.1) | | | X | X | | | |
| Supplementary problems added because of analysis of Quiz I results | | | | X | | | |
| Exercise 2 (6.2) | | | | X | X | | |
| <u>Lesson 6 (Part II)</u> | | | | | | | |
| Dice | | | | | | | |
| Exercise 3 (6.3) | | | | | X | X | |
| <u>Lesson 7 (Parts I)</u> | | | | | | | |
| Quiz II A (Lessons 1-5) | | | | | | X | |
| Quiz II B (Lesson 6) | | | | | | X | |
| Review distinction between one-dimensional and two-dimensional probability problems | | | | | | X | |
| Choosing between two boxes (one-dimensional) | | | | | | X | |
| Exercise 1 (7.1) | | | | | | X | X |
| Supplement A to Lesson 7 (from analysis of Quiz II B results) | | | | | | | X |

Figure 10. (Cont'd)

| | Date | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 3/24 | 3/25 | 3/26 | 3/27 | 3/28 | 3/31 | 4/1 | 4/2 |
| <u>Lesson 7 (Part I)</u> | | | | | | | | |
| Supplement B to Lesson 7 (Review of concepts of Lesson 6) | X | | | | | | | |
| Quiz III (Lesson 6) | X | | | | | | | |
| <u>Lesson 7 (Part II)</u> | | | | | | | | |
| Choosing between one-dimensional and two-dimensional games | | X | | | | | | |
| Exercise 2 (7.2) | X | X | X | | | | | |
| Class broken into Masters and Nonmasters of Lesson 6 | | X | | | | | | |
| Collecting of data (homework) | | X | X | | | | | |
| <u>Lesson 8</u> | | | | | | | | |
| Performing experiments (verifying decisions made in terms of probability) | | | X | X | | | | |
| <u>Lesson 9</u> | | | | | | | | |
| Analyzing data | | | | X | | | | |
| Quiz IV (Lesson 7) | | | | | X | | | |
| Estimated probability | | | | X | X | X | | |
| Convergence of estimated probability as number of trials increase | | | | | X | X | | |
| Exercise (Law of Averages) 9.1 | | | | | X | X | | |
| <u>Review</u> | | | | | | X | | |
| <u>Post Test</u> | | | | | | | | |
| Form A | | | | | | | X | |
| Form B | | | | | | | | X |

Figure 10. (Cont'd)

in two-dimensions. With regard to probability items on the pretest, five students (subjects 10, 11, 13, 16, and 21) consistently used a fraction a/b ($o \leq \frac{a}{b} \leq 1$) to answer questions asking the student to specify the probability. One (subject 25) used "odds" statement consistently. One (subject 8) vacillated between using odds and using fractions. Altogether 7/25 students consistently endeavored to assign a number to a question concerned with questions such as "What is the probability of _____?" The rest of the students either left such questions blank, were inconsistent in the form of their responses, or used terms such as "pretty good," "good," "uncertain," "bad," "awful," "impossible." (See Appendix C, Table 56 for individual results.)

## Actual Instruction and Evaluation

The actual instructional program departed from the instructional plan beginning with the first day of pretesting (March 3). Initial instruction on bar graphing was planned for the remaining part of the period, after administration of the pretest. However, one boy took 10 minutes more than anyone else, and the period was unexpectedly cut short because of a school activity. This meant that teaching students to construct and read a bar graph carried into the second and third day. Because of this change, the amount of time available for the first lesson was cut in half. The dice game which had been planned for either Lesson 1 or Lesson 4 was moved back to Lesson 4. Other ommissions are noted in the Journal.

Lesson 2, on performing experiments, went as planned. However, Lesson 3 on graphing and interpreting the data gathered from Lesson 2 took longer than anticipated (3/7, 3/10, 3/11, 3/12). After Lesson 3 the teacher began Lesson 4 on assigning a probability to a one-dimensional event. The lesson was conducted as planned.

Lesson 5 on assigning a probability to a two-dimensional event was not very successful. (This statement is based on observations and the analysis of the exercise that accompanied the lesson.) The students had a difficult time accepting a two-dimensional outcome as one outcome. Due to a lack of time, Quiz 1, which had been planned for the end of the period, was postponed. This quiz was to have measured some of the more important objectives of Lessons (1 - 5). However, students were obviously confused concerning two-dimensional outcomes. Thus the experimenter decided to postpone the quiz until after the introduction to Lesson 6 - Part I concerned with using a tree to count outcomes of a two-dimensional sample space. More practice conerning two-dimensional outcomes was thus given. However, the results of Quiz I were not up to criteria. (See Table 3 for summary of all quiz results.) A mean of 16.72 (20 items) was less than expected. Also only 60% of the students scored 90% or better and only 80% of the students scored 80% or better.

From an analysis of Quiz I, the experimenter knew that the objective of specifying the number of possible outcomes in a one-dimensional sample space was causing students difficulties. More treatment concerning one-dimensional problems was given the following day before doing Exercise II of Lesson 6. An extra help session was conducted after class Wednesday (3/19) for five students.

Lesson 6 - Part II was done as planned.

Quiz IIA (14 items), a parallel quiz to Quiz I, and Quiz IIB (10 items) measuring the objectives of Lesson 6 were given together on Thursday (3/25). The mean on Quiz IIA was 12.54 (90%) with 18s/24s scoring 12/14 or better. The 18 masters of Quiz IIA and the 90% mean, plus three masters of Quiz I who were not masters of Quiz II convinced the author that no more class treatment on the objectives of Lessons (1 - 5) would be necessary.

The mean for Quiz IIB, 6.96 or 69.6% was poor. Only 6s/24s scored 9/10 or better and only 11s/24s scored 8/10 or better. The principle problem was in specifying the number of possible outcomes in a two-dimensional sample space. Many students were not employing trees to arrive at their answer.

Lesson 7 - Part I on making a decision between two urns was done after the quiz on Thursday (3/20). The exercise for Lesson 7 was started in class Thursday and finished during the first part of Friday's lesson. The experimenter felt that he could procede with Lesson 7 - Part I since two-dimensional problems were not a prerequisite behavior. Since the exercise was done independently in class, the results were analyzed as a quiz. The mean was 14.8 (17 items) or 87.1% with 21s/25s scoring 15/17 or better and 23s/25s scoring 13/17 or better. (See Table 3.) Only subjects 15 (score-5/17) and 20 (score-3/17) had difficulty with the exercise. From these results and the pretest results on ordering two fractions, the investigator decidea that the practice sheet on ordering two fractions would not be used as a class exercise. Instead specific individuals who had problems with ordering fractions would be asked to do the exercise sheet.

From the analysis of Quiz IIB measuring the objectives of Lesson 6, more group practice on the objectives was planned for Friday (3/21) and Monday (3/24). Following the short review on Monday a second quiz on Lesson 6 (Quiz III (5 items) was administered. The mean was 4.0 or 80% with 16s/23s scoring 4/5 or better. Since these results did not meet the criterion and since specifying the probability of an event in two-dimensions was considered a necessary behavior to Lesson 7 - Part II, the experimenter decided to split the class into masters and nonmasters of Lesson 6. On Tuesday (3/25), the masters continued Lesson 7 - Part II (deciding between two games) while the experimenter worked with eight nonmasters on the objectives of Lesson 6. On the basis of student's performance on a set of problems, 3s/8s were classified as masters of the objectives of Lesson 6. Only two subjects ("15" and "20"), were still confused.

The second extra help session concentrating on the objectives of Lessons 6 and 7 - Part II was held Wednesday (3/26) for six children who had been classified as nonmasters of Quiz IIA and/or Quiz IIB.

Lesson 8 (3/26) concerned with experimentally verifying that decisions made on the basis of probability are the best <u>a priori</u> decisions was conducted as planned.

To observe the convergence of the estimated probability, experiments such as tossing a tack were suggested as outside activities. The data from these experiments and from Lesson 8 were pooled to plot

TABLE 3

SUMMARY OF QUIZ RESULTS MEASURING ANY OF THE MAJOR BEHAVIORAL OBJECTIVES

| Quiz | Date Given | Total Number of Items (Includes all Objectives Measured) | Major Behavioral Objectives Included on Quiz | Mean | Variance | Criterion (Ratio Form) |
|---|---|---|---|---|---|---|
| Quiz I (Lesson 1-5) | 3/17 | 20 | 1, 3, 4a-d, 9, 10 (1-D only) | 16.72 | 8.12 | 15/25 scored 18/20<br>20/25 scored 16/20<br>23/25 scored 14/20 |
| Quiz IIA (Lesson 1-5) | 3/20 | 14 | 1, 3, 4a-4d, 9, 10 (1-D only) | 12.54 | 1.83 | 15/24 scored 13/14<br>21/24 scored 11/14<br>24/24 scored 10/14 |
| Quiz IIB (Lesson 6) | 3/20 | 10 | 3, 4a-4d (2-D only) | 6.96 | 4.00 | 6/24 scored 9/10<br>11/24 scored 8/10<br>14/24 scored 7/10 |
| Quiz III (Lesson 6) | 3/24 | 5 | 4a, 4b (1 item 1-D, 4 item 2-D) | 4.00 | 1.83 | 13/23 scored 5/5<br>16/23 scored 4/5<br>19/23 scored 3/5 |
| Exercise (7I) (Lesson 7I) | 3/21 | 17 | 6, 7 | 14.80 | 11.44 | 21/25 scored 15/17<br>23/25 scored 13/17 |
| Quiz IV (Lesson 7I and II) | 3/28 | 10 | 6, 7 | 8.88 | 2.35 | 19/25 scored 9/10<br>21/25 scored 8/10<br>22/25 scored 7/10 |

bar graphs of the cumulative frequency against estimated probability. Due to a lack of time, the activity was teacher centered rather than student centered. This was the opposite of the intent of the lesson.

On Friday (3/28) a quiz on Lesson 7 was given. The results were encouraging in that 19s/25s scored 9/10 or better and 21s/25s scored 8/10 or better. The average percentage was 88.8%.

An exercise on the Law of Averages was assigned as homework.

The last instructional period on Monday (3/31) completed Lesson 9 and reviewed the overall unit very quickly. A third extra help session was held for subjects 15 and 20.

The original instructional plan had called for four days for pre and post-testing and 16 to 17 days of instruction. Lesson 3, graphing data, took longer than anticipated. Assignging probability to an event in two-dimensions was more difficult than anticipated. So instruction went on for 19 days. Also mainly due to the lack of success in attaining the latter factor, five quizzes rather than the planned two were used.

Tuesday (4/1) and Wednesday (4/2) were spent posttesting. It did not take as long as expected, and could have been done in a day.

This chapter presented the study's design as part of an instructional system and gave a short account of the actual procedures used. Chapter IV presents in detail the results of the pretest and posttest.

## Chapter IV

## RESULTS

The purpose of this chapter is to describe the statistical procedures used in this study and to present the data from the pretest and posttest. A brief discussion of the reliability and validity of the test is included. The results for the 14 measured objectives are presented. Also, a short description of how the learning hierarchies discussed in Chapter II are to be tested and the data from those tests are included.

The statistical procedures employed in this formative study are descriptive in nature. To measure the effect of the instructional treatment, a 72 item pre and post-test was administered and the item percentage computed. (The test is in Appendix B.) The pre and post-percentages measuring 14 specific behavioral objectives were also computed by averaging the percentages of the items measuring these objectives on the pre and post-test.

The criterion test of 72 items was employed (1) to measure the effectiveness of the instructional treatment using a (non-random) sampling of behavioral objectives of the treatment and (2) to test the validity of learning hierarchies referred to in Chapter II. The same test was used for the pre and post-testing.

Validity

The test has content validity since the items are criterion items based on the instructional analysis and materials. They were written to test specific behavioral objectives of the instructional treatment.

Reliability

Since the test is a criterion-referenced test, the concept of reliability as measured by norm-referenced statistics is not appropriate. Popham and Husek (1969) establish that one can have a good criterion-referenced test with a zero internal consistency estimate (perfect scores) or even negative internal consistency estimates. Popham and Husek state, "Thus, the typical indices of internal consistency are not appropriate for criterion-referenced tests. It is not clear what should replace them (p. 5)."

For this study, using Hoyt's measure of internal consistency (Baker, 1965), the reliability estimates for the pretest was .86 and for the posttest was .65. The higher internal consistency on the pretest than on the posttest was caused by the much larger variance for the pretest than the posttest. (See Table 4.) This data tends to support Popham and Husek. At present there is no acceptable procedure for computing the reliability of a criterion-referenced test.

Overall Results

Concerning the overall results, the 25 students had a pretest mean of 27.28 and a posttest mean of 66.80 out of a possible raw score of 72. In percentages, the average pretest score was 37.9 percent and the average posttest score was 92.8 percent. There was a dramatic change in the variances of these two tests from 74.13 on the pretest to 11.17 on the posttest. This was caused by the large differences between individual performance on the pretest (low score 7, high score 44) and relatively small differences between individual performances on the posttest (low score 58, high score 71 with 22s/25s scoring between 64 and 71). These results are summarized in Table 4.

TABLE 4

OVERALL RESULTS OF THE PRETEST AND POSTTEST

| | Pretest | Posttest |
|---|---|---|
| Mean (out of 72 items) | 27.28 | 66.80 |
| Mean in terms of percentage | 37.9% | 92.8% |
| Variance | 74.13 | 11.17 |

The change in item percentages for the 72 items for pretest to posttest was encouraging. All items had a positive amount of change. Six items had a 100 point change;[3] 10 items had an 80 to 99 point change; 19 items had a 60 to 79 point change; 12 items had a 40 to 59 point change; 17 items had a 20 to 39 point change; and 8 items had a 0 to 19 point change. On the posttest, 29 items had 100% correct responses, 29 items had 92%

[3]A "100 point change" means the item difficulty has changed from 0% on the pretest to 100% on the posttest.

to 99% correct responses, 9 items had 80% to 91% correct responses, 2 items had 68% to 79% correct responses, and 3 items had percentage of correct responses below 68%.

The item percentages on the pretest and posttest are presented in Table 5. The expected score by chance alone is given for the multiple choice items.

## Behavioral Objectives

As was mentioned in Chapter II, the ultimate goal of instruction was (90/90), that is that 90 percent of the students should score 90 percent or better on each of the measured objectives. However, these percentages are functions of the number of students and the number of items used to measure a behavioral objective. To judge the success of instruction, the author is using both the arbitrary 90/90 criterion and the practical criterion referred to in Chapter II (p. 67).

On the basis of the posttest scores on the 14 measured behavioral objectives (See Chapter II, p. 70) and the criteria, one can conclude that instruction was successful concerning 10 of the objectives, very close to being successful on another and unsuccessful on three of the objectives. (No behavioral objective was achieved for either criteria on the pretest.)

The results for objectives 4a (probability of a simple event), 4d (probability of the impossible event), 5 (ordering two fractions), 6 (most likely event) and 9 (likely bounds) satisfied the 90/90 criterion. Objectives 1 (subjective notions), 3 (number of possible

TABLE 5

ITEM PERCENTAGES OF THE PRETEST AND POSTTEST

Total Number of Students = 25

| Part | Item No. | Expected Score By Chance (%) | PRETEST % of Correct Responses | PRETEST No. of Correct Responses | POSTTEST No. of Correct Responses | POSTTEST % of Correct Responses |
|---|---|---|---|---|---|---|
| R1 | 1 | 33 | 72 | 18 | 25 | 100 |
| | 2 | 33 | 92 | 23 | 25 | 100 |
| | 3 | 33 | 72 | 18 | 23 | 92 |
| | 4 | 33 | 60 | 15 | 24 | 96 |
| | 5 | 50 | 56 | 14 | 22 | 88 |
| | 6 | 50 | 40 | 10 | 25 | 100 |
| | 7 | 50 | 64 | 16 | 22 | 88 |
| | 8 | 50 | 60 | 15 | 22 | 88 |
| R2 | 1 | 25 | 40 | 10 | 23 | 92 |
| | 2 | 25 | 32 | 8 | 25 | 100 |
| R3 | 1 | 33 | 96 | 24 | 25 | 100 |
| | 2 | 33 | 88 | 22 | 25 | 100 |
| | 3 | 33 | 64 | 16 | 24 | 96 |
| | 4 | 33 | 96 | 24 | 25 | 100 |
| | 5 | 33 | 96 | 24 | 25 | 100 |
| | 6 | 33 | 72 | 18 | 23 | 92 |
| | 7 | 33 | 68 | 17 | 24 | 96 |
| | 8 | 33 | 88 | 22 | 25 | 100 |
| | 9 | 33 | 68 | 17 | 23 | 92 |
| | 10 | 33 | 52 | 13 | 21 | 84 |
| R4 | 1 | 33 | 36 | 9 | 23 | 92 |
| | 2 | 33 | 28 | 7 | 22 | 88 |
| | 3 | 33 | 40 | 10 | 24 | 96 |
| | 4 | 33 | 16 | 4 | 23 | 92 |
| | 5 | 33 | 16 | 4 | 23 | 92 |
| | 6 | 33 | 8 | 2 | 23 | 92 |
| | 7 | 33 | 52 | 13 | 20 | 80 |

TABLE 5 (con't.)

Total Number of Students = 25

| Part | Item No. | Expected Score By Chance (%) | PRETEST % of Correct Responses | PRETEST No. of Correct Responses | POSTTEST No. of Correct Responses | POSTTEST % of Correct Responses |
|---|---|---|---|---|---|---|
| R5 | 1 | 33 | 32 | 8 | 23 | 92 |
| | 2 | 33 | 60 | 15 | 24 | 96 |
| | 3 | 33 | 72 | 18 | 25 | 100 |
| | 4 | 33 | 32 | 8 | 25 | 100 |
| | 5 | 33 | 68 | 17 | 25 | 100 |
| | 6 | 33 | 60 | 15 | 24 | 96 |
| | 7 | 33 | 52 | 13 | 24 | 96 |
| | 8 | 33 | 52 | 13 | 25 | 100 |
| | 9 | 33 | 52 | 13 | 25 | 100 |
| | 10 | 33 | 28 | 7 | 23 | 92 |
| | 11 | 33 | 32 | 8 | 23 | 92 |
| R6 | 1 | | 44 | 11 | 24 | 96 |
| | 2 | | 32 | 8 | 25 | 100 |
| | 3 | | 24 | 6 | 25 | 100 |
| | 4 | | 12 | 3 | 24 | 96 |
| | 5 | | 48 | 12 | 25 | 100 |
| R7 | 1 | | 48 | 12 | 25 | 100 |
| | 2 | | 84 | 21 | 25 | 100 |
| | 3 | | 32 | 8 | 25 | 100 |
| | 4 | | 32 | 8 | 25 | 100 |
| R8 | 1 | | 4 | 1 | 25 | 100 |
| | 2 | | 16 | 4 | 16 | 64 |
| | 3 | | 0 | 0 | 25 | 100 |
| | 4 | | 8 | 2 | 23 | 92 |
| R9 | 1 | | 0 | 0 | 23 | 92 |
| | 2 | | 24 | 6 | 22 | 88 |
| | 3 | | 0 | 0 | 25 | 100 |
| | 4 | | 0 | 0 | 21 | 84 |
| | 5 | | 20 | 5 | 24 | 96 |

TABLE 5 (con't.)

Total Number of Students = 25

| Part | Item No. | Expected Score By Chance (%) | PRETEST % of Correct Responses | PRETEST No. of Correct Responses | POSTTEST No. of Correct Responses | POSTTEST % of Correct Responses |
|---|---|---|---|---|---|---|
| R10 | 1 | | 0 | 0 | 23 | 92 |
| | 2 | | 0 | 0 | 25 | 100 |
| | 3 | | 0 | 0 | 25 | 100 |
| | 4 | | 4 | 1 | 25 | 100 |
| | 5 | | 20 | 5 | 23 | 92 |
| R11 | 1 | | 84 | 21 | 23 | 92 |
| | 2 | | 28 | 7 | 17 | 72 |
| | 3 | | 24 | 6 | 24 | 96 |
| | 4 | | 16 | 4 | 23 | 92 |
| R12 | 1 | | 0 | 0 | 22 | 88 |
| | 2 | | 4 | 1 | 19 | 76 |
| | 3 | | 0 | 0 | 25 | 100 |
| | 4 | | 0 | 0 | 24 | 96 |
| | 5 | | 0 | 0 | 25 | 100 |
| R13 | 1 | | 8 | 2 | 13 | 52 |
| R14 | 1 | 20 | 4 | 1 | 7 | 28 |

outcomes of a sample space), 4b (probability of a compound event), 4c (probability of a certain event), and 7 (equally likely event) satisfied the practical criteria.

The results for objective 10 (law of averages) almost satisfied the practical criterion. However the results for objectives 2 (number of outcomes of an event), 8 (estimated probability), and 11 (estimate of the probability) are not close to meeting either criteria. These results are summarized in Table 6.

The average percentage for a behavioral objective on the pretest indicated that no behavioral objective had an average percentage above 80%; two objectives had average percentages between 50% and 80% and twelve objectives had average percentages below 50%. On the posttest, six objectives had averages above 95%, with 11 of the 14 objectives having averages above 90%. Only one objective (11) had an average percentage below 50%. Figure 11 depicts the relationship between the average percentage on the pretest and on the posttest for a behavioral objective. The graph is based on data from Table 7.

To show the particular items measuring a behavioral objective and the number of correct responses on the pretest and posttest for the objective, Tables 16 to 29 are included in Appendix C. To show the percentage of students achieving various levels of performance for each of the fourteen objectives, Tables 30 to 43 are also included in Appendix C.

TABLE 6

LEVEL OF ACHIEVEMENT OF THE FOURTEEN BEHAVIORAL OBJECTIVES

| Behavioral Objective | Criterion Percentage | Criterion Ratio Form |
|---|---|---|
| *1. subjective notions | 96/80 | 24/25 scored 4/5 or 5/5 |
| 2. number of outcomes of an event | 72/80 | 18/25 scored 4/5 or 5/5 |
| *3. number of possible outcomes of a sample space | 88/86 | 22/25 scored 6/7 or 7/7 |
| **4a. probability of a simple event | 96/100 | 24/25 scored 8/8 |
| *4b. probability of a compound event | 92/88 | 23/25 scored 7/8 or 8/8 |
| *4c. probability of a certain event | 88/100 | 22/25 scored 2/2 |
| **4d. probability of the impossible event | 96/100 | 24/25 scored 2/2 |
| **5. order of two fractions | 92/90 | 23/25 scored 9/10 or 10/10 |
| **6. most likely events | 92/90 | 23/25 scored 9/10 or 10/10 |
| *7. equally likely events | 92/88 | 23/25 scored 7/8 or 8/8 |
| 8. estimated probability | 52/100 | 13/25 scored 1/1 |
| *9. likely bounds | 100/100 | 25/25 scored 1/1 |
| 10. law of averages | 84/100 | 21/25 scored 4/4 |
| 11. estimate of the probability | 28/100 | 7/25 scored 1/1 |

* Objective achieved at the practical criterion level.
** Objective achieved at the 90/90 criterion level.

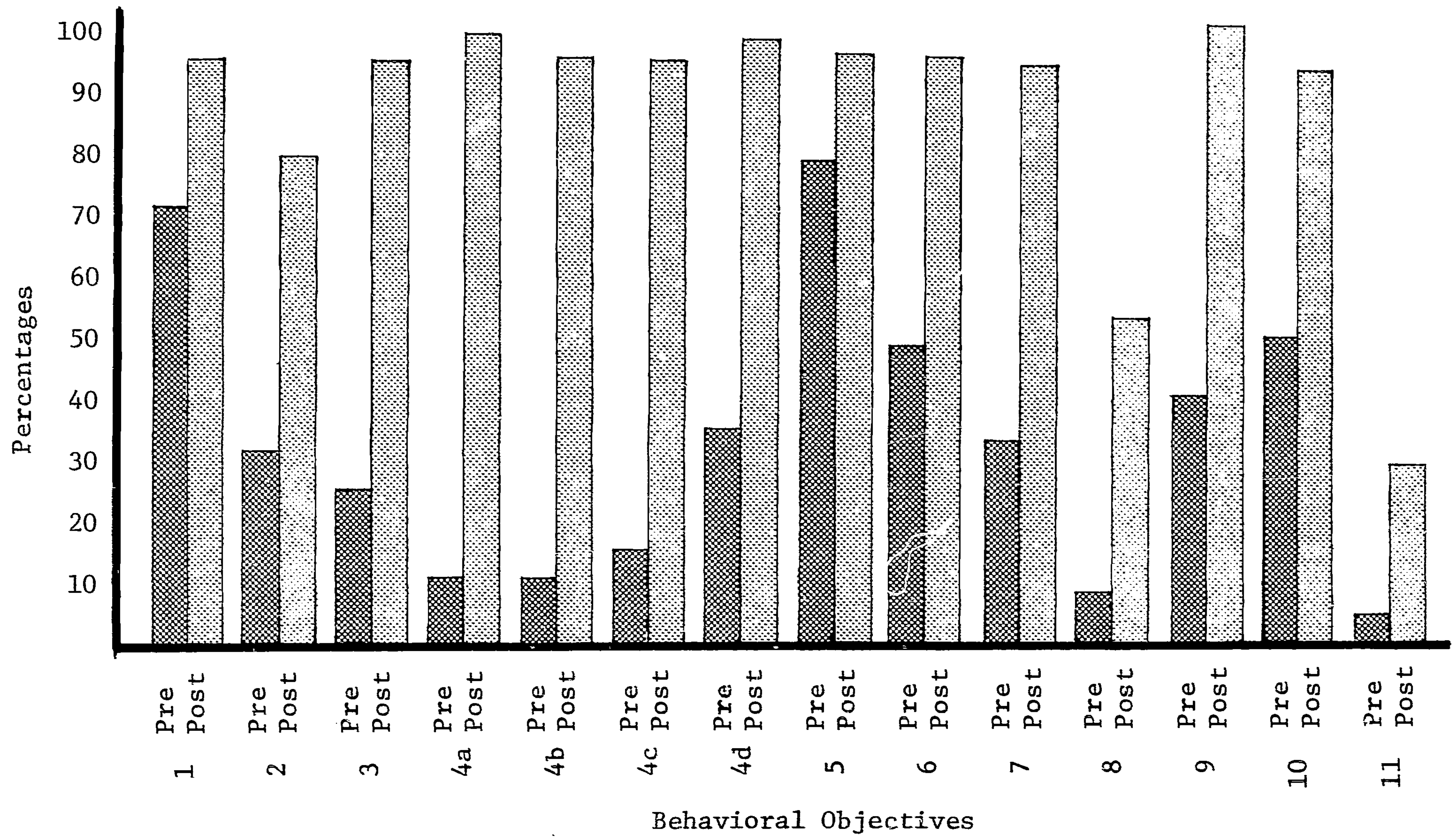



Figure 11. Percentage Graph of Behavioral Objectives
As Measured by the Pretest and Posttest

TABLE 7

SUMMARY TABLE OF BEHAVIORAL OBJECTIVES (PERCENTAGE OF CORRECT RESPONSES)

| Behavioral Objectives | Number of items | | Pretest | | | Posttest | | |
|---|---|---|---|---|---|---|---|---|
| | 1D | 2D | 1D | 2D | 1D and 2D | 1D | 2D | 1D and 2D |
| 1. subjective notions | 5 | | 70.4 | - | - | 95.2 | - | - |
| 2. number of outcomes of an event | 2 | 3 | 56 | 14.7 | 31.2 | 84 | 76 | 79.2 |
| 3. number of possible outcomes of a sample space | 3 | 4 | 58.7 | 1.0 | 25.7 | 96 | 93 | 94.3 |
| 4a. probability of a simple event | 3 | 5 | 28.0 | 0 | 10.5 | 98.7 | 100 | 99.5 |
| 4b. probability of a compound event | 3 | 5 | 24 | 2.4 | 10.5 | 97.3 | 94.4 | 95.5 |
| 4c. probability of a certain event | 1 | 1 | 12 | 20 | 16 | 96 | 92 | 94 |
| 4d. probability of the impossible event | 1 | 1 | 48 | 20 | 34 | 100 | 96 | 98 |

Table 7 (Con't.)

| Behavioral Objectives | Number of items 1D | Number of items 2D | Pretest Ratio | Pretest 1D | Pretest 2D | Pretest 1D and 2D | Posttest Ratio | Posttest 1D | Posttest 2D | Posttest 1D and 2D |
|---|---|---|---|---|---|---|---|---|---|---|
| 5. order of two fractions | (10) | | 78.8 | | | | 96 | | | |
| 6. most likely events | 6 | 4 | | 60.7 | 29 | 48 | | 98.7 | 90 | 95.2 |
| 7. equally likely events | 5 | 3 | | 35.2 | 26.7 | 32 | | 94.4 | 90.7 | 93 |
| 8. estimated probability | | | | 8 | - | - | | 52 | - | - |
| 9. likely bounds | 1 | - | | 40 | - | - | | 100 | - | - |
| 10. law of averages | 4 | | | 49 | - | - | | 92 | - | - |
| 11. estimate of the probability | 1 | | | 4 | - | - | | 28 | - | - |

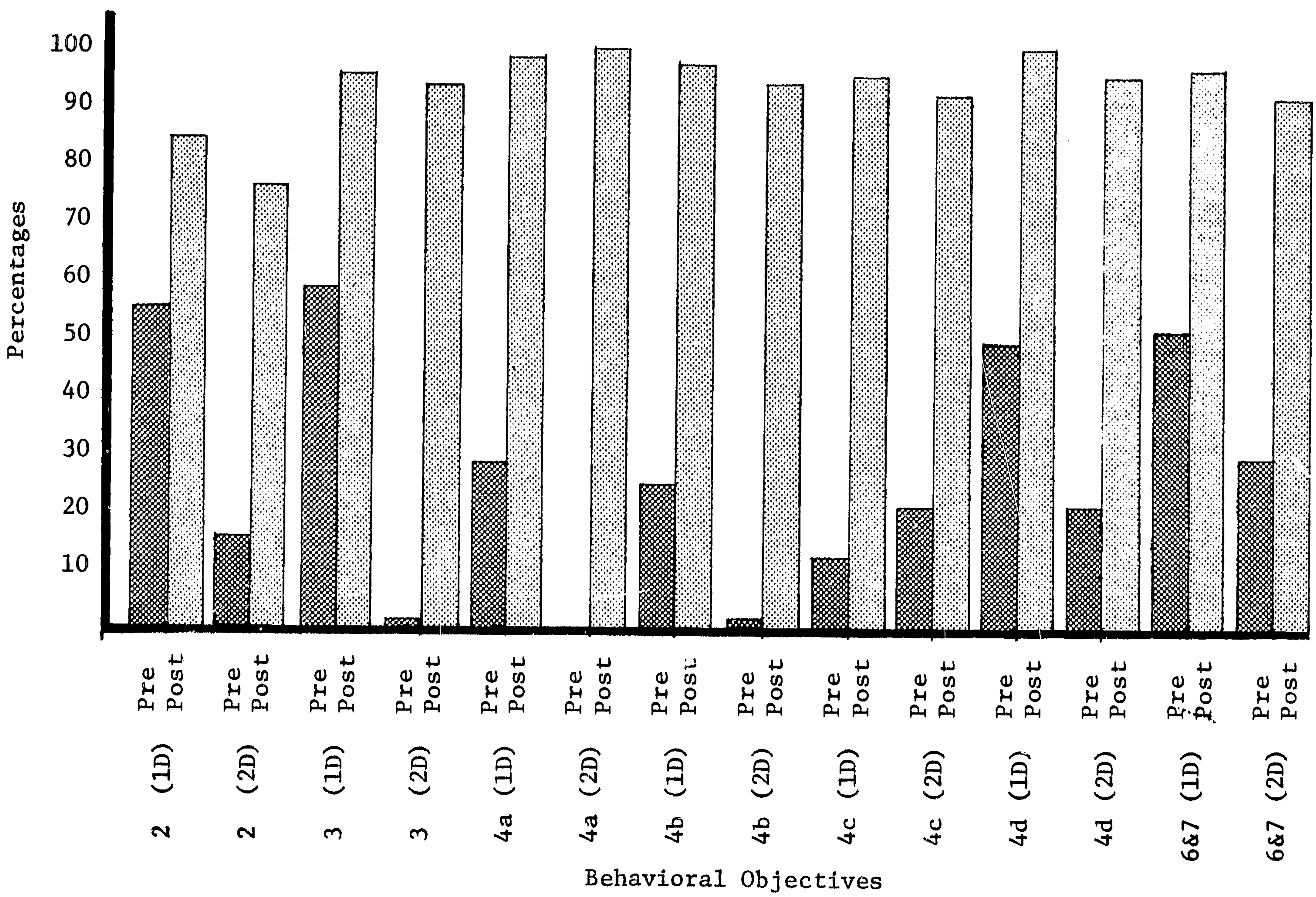



Figure 12. Percentage Graph of Behavioral Objectives Concerned with Both 1-D and 2-D Problems As Measured by the Pretest and Posttest

## One-Dimensional and Two-Dimensional Results

The test was also designed to look at certain objectives as they were measured by one-dimensional (1-D) and two-dimensional (2-D) problems. These results show that on the pretest two-dimensional objectives were much more difficult than one dimensional objectives. However, on the posttest there was little difference. Table 7 summarizes the results for each behavioral objective in percentage form and gives the breakdown for objectives measured by one-dimensional problems and two-dimensional problems. To show the results pictorially, Figure 12 presents a bar graph of the objectives measured in both dimensions. To show when students were achieveing an objective in one or two-dimesnions, Tables 44 to 55 in Appendix C have included information from the quizzes and the pretest and posttest.

## Learning Hierarchy

The following pages present the rationale and the results of testing the two learning hierarchies referred to in Chapter II. The paradigm for testing the feasibility of learning hierarchies has been constructed by Walbesser (1968). Given the following relationships in Figure 13, one assigns a _0_ if the individual does not exhibit the

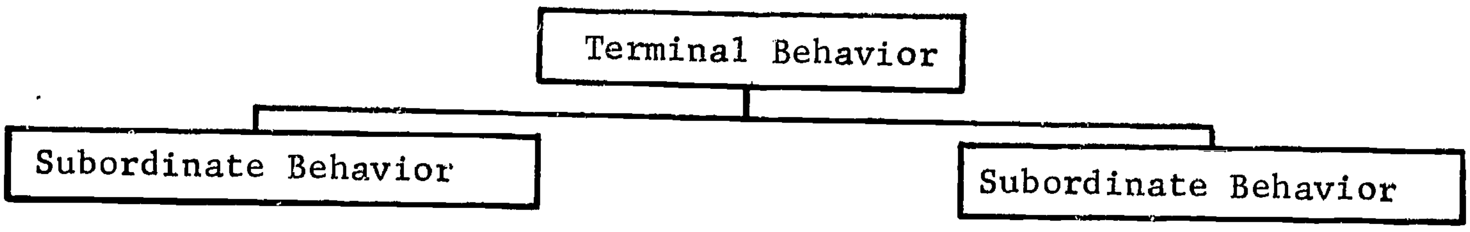



Figure 13. Picture of a Typical Learning Hierarchy

behavior (terminal or subordinate) and a 1 if he does. One can use the following paradigm to describe the hierarchy in Figure 13. Either the persons "possess all behaviors on both levels (+, +); possess the terminal behavior, but not all of the subordinate behaviors (+, -); not possess the terminal behavior but possess the subordinate behaviors (-, +); not possess the terminal behavior, and not all of the subordinate behaviors (-, -) (Walbesser, 1968, p. 204)." For the hierarchy in Figure 13 the components of each column would be as follows:

| (+, +) | (+, -) | (-, +) | (-, -) |
|---|---|---|---|
| 1<br>1 1 | 1 1 1<br>0 1 1 0 0 0 | 0<br>1 1 | 0 0 0<br>0 0 0 1 1 0 |

Figure 14. Learning Hierarchy Components

The (+, +) observations support the hypothesis. The (-, +) observations suggest that there may be something inadequate in the instructional material related to acquiring the terminal behavior and these observations should not be included in a measure of support for the hypothesis. The (+, -) observations clearly refute the hypothesis. This formulation is adapted from the work of Gagné and others on the analysis of learning hierarchies.

In order to obtain a magnitude of the degree of support for the claim of a valid learning hierarchy, the following ratios and the level of acceptable magnitude have been constructed. The three ratios are:

(1) The consistency ratio =

$$\frac{\text{number of members in the } (+, +) \text{ column}}{\text{total number of subjects who acquired the terminal behavior (members of the } (+, +) \text{ and } (+, -) \text{ columns)}} = \frac{\text{number of } (+, +)}{\text{number of } (+, +) \text{ and } (+, -)}$$

"The value of this ratio is a measure of how consistent the data are with the hypothesized dependency (Walbesser, 1968, p. 205)."

(2) The adequacy ratio =

$$\frac{\text{number of } (+, +)}{\text{number of } (+, +) \text{ and } (-, +)}$$

measures the adequacy of the identified subordinate behaviors and the instruction ;if the instruction has been adequate then the expectation is that all subjects who have acquired the subordinate will also have acquired the terminal behavior.

(3) The completeness ratio =

$$\frac{\text{number of } (+, +)}{\text{number of } (+, +) \text{ and } (-, -)}$$

This ratio is a reflection of imcomplete instruction in that if a large number of cases are in the (-, -) category the ratio will be small in magnitude.

For Science - process approach (Walbesser, 1968) ratios of the magnitude of .90 or .95 for all three ratios have been considered necessary to validate a hypothesized hierarchy such as in Figure 13. To measure the validity of the two learning hierarchies referred to in Chapter III, the following ratios have computed.

It has been hypothesized that in order to learn to identify when two events are equally likely or which of two unequally likely events is more likely, $\boxed{M}$, that the child must learn to specify the probability of an event, $\boxed{P}$, and be able to order two ratios, $\boxed{R}$. In order to verify this learning hierarchy (See Figure 4 , p. 46 for a picture of the hierarchy)the ratios considered in Chapter II are computed for the 90% Test Criterion and the Practical Test Criterion.

There are 18 items to measure $\boxed{M}$ (See Appendix C, Tables 24 and 25), 16 items to measure $\boxed{P}$ (See Appendix C, Tables 19 and 20) and 10 items to measure $\boxed{R}$. (See Appendix C, Table 23.)

If a 90% criterion is placed as necessary to being considered a master of any of the three objectives, a student can miss at most one item and still be classified as a master of that objective. Thus for objective $\boxed{M}$, a student scoring 16/18 or 89% is classified as a nonmaster. Similarily for objective $\boxed{P}$, a student scoring 14/16 or 88% is classified as a nonmaster.

With respect to objective $\boxed{R}$, 23/25 students scored 9/10 or better, (one scored 7/10 and one scored 8/10). With respect to objective $\boxed{P}$, 23/25 students scored 15/16 or better (two scored 14/16). With respect to $\boxed{M}$, 18/25 students scored 17/18 or better (five scored 16/18, one scored 13/18 and one scored 12/18); thus 23/25 students scored 16/18 or better.

Using the criterion of 90%, the table used for computing the ratios used to test the hierarchy is:

TABLE 8

PARTITIONING OF STUDENTS IN LEARNING HIERARCHY 1
BY OBJECTIVES: TEST CRITERION--90%

| | (+ , +) | (+, -) | | | (- +) | (-, -) | | |
|---|---|---|---|---|---|---|---|---|
| | 1<br>1 1 | 1<br>0 1 | 1<br>1 0 | 1<br>0 0 | 0<br>1 1 | 0<br>0 0 | 0<br>0 1 | 0<br>1 0 |
| Numbers of pupils | 15 | 2 | 1 | 0 | 6 | 0 | 0 | 1 |
| | | 3 | | | | 1 | | |

On the basis of the formulas on the previous two pages, the following ratios have been computed.

$$\text{Consistency Ratio} = \frac{15}{18} = .833$$

$$\text{Adequacy Ratio} = \frac{15}{21} = .714$$

$$\text{Completeness Ratio} = \frac{15}{16} = .938$$

If the practical test criterion (See Chapter II, p. 69) for classifying subjects as masters is used, scores of 16/18 or better for $\boxed{M}$ and 14/16 or better for $\boxed{P}$ are acceptable. The table and ratios change to the following:

Table 9

PARTITIONING OF STUDENTS BY OBJECTIVES IN LEARNING HIERARCHY 1--PRACTICAL TEST CRITERION

| | (+,+) | (+, -) | | | (-, +) | (-,-) | | |
|---|---|---|---|---|---|---|---|---|
| | (+, +) | | | | | | | |
| | 1<br>1 1 | 1<br>0 1 | 1<br>1 0 | 1<br>0 0 | 0<br>1 1 | 0<br>0 0 | 0<br>0 1 | 0<br>1 0 |
| Number of Pupils | | 0 | 2 | 0 | | 0 | 0 | 0 |
| | 21 | 2 | | | 2 | 0 | | |

$$\text{Consistency Ratio} = \frac{21}{23} = .913$$

$$\text{Adequacy Ratio} = \frac{21}{23} = .913$$

$$\text{Completeness Ratio} = \frac{21}{21} = 1.000$$

For the second hierarchy that is to be tested, it has been hypothsized that in order to learn to specify the probability of an event, $\boxed{P}$, one must specify the number of outcomes of an event, $\boxed{E}$, and specify the number of possible outcomes of the sample space, $\boxed{S}$, (See Figure 5, p.46.)

There are 5 items that measure $\boxed{E}$, (See Appendix C, Table 17) 7 items that measure $\boxed{S}$ (See Appendix C, Table 18) and 16 items that measure $\boxed{P}$. First a table was formed using the 90% criterion (See Table 10). Thus to be classified as a master of $\boxed{E}$, one must score 100% (5/5) on these items. The same is true for $\boxed{S}$ (7/7). With regard to objective $\boxed{P}$, one must score 15/16 or better.

TABLE 10

PARTITIONING OF STUDENTS BY OBJECTIVES IN LEARNING HIERARCHY 2--TEST CRITERION 90%

| (+, +) | (+, -) | | | (-,+) | (-, -) | | |
|---|---|---|---|---|---|---|---|
| 1<br>1 1 | 1<br>0 1 | 1<br>1 0 | 1<br>0 0 | 0<br>1 1 | 0<br>0 0 | 0<br>0 1 | 0<br>1 0 |
| | 10 | 2 | 4 | | | | 1 |
| 7 | 16 | | | 1 | 1 | | |

$$\text{Consistency Ratio} = \frac{7}{23} = .328$$

$$\text{Adequacy Ratio} = \frac{7}{8} = .875$$

$$\text{Completness Ratio} = \frac{7}{8} = .875$$

Using the practical test criterion for objectives $\boxed{E}$, $\boxed{S}$ and $\boxed{P}$ one is classified as a master if he gets $\frac{4}{5}$ or better on objective $\boxed{E}$, $\frac{6}{7}$ or better on objective $\boxed{S}$ and $\frac{14}{16}$ or better on objective $\boxed{P}$. The table and ratios change markedly (See Table 11).

TABLE 11

PARTITIONING OF STUDENTS BY OBJECTIVES IN LEARNING HIERARCHY 2--PRACTICAL TEST CRITERION

| (+, +) | (+, -) | | | (-, +) | (-, -) | | |
|---|---|---|---|---|---|---|---|
| 1<br>1 1 | 1<br>0 1 | 1<br>1 0 | 1<br>0 0 | 0<br>1 1 | 0<br>0 0 | 0<br>0 1 | 0<br>1 0 |
| | 6 | 2 | 1 | | 0 | 0 | 0 |
| 16 | 9 | | | 0 | 0 | | |

$$\text{Consistency Ratio} = \frac{16}{25} = .64$$

$$\text{Adequacy Ratio} = \frac{16}{16} = 1.00$$

$$\text{Completeness Ratio} = \frac{16}{16} = 1.00$$

## Chapter V

## SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

### Summary

Probability and statistics are important mathematical tools used by man in technological society. For numerous reasons recommendations have been made for a comprehensive program in probability and statistics which begins in the elementary school. Research seems to indicate certain topics in probability and statistics may be suitable for elementary students to learn. However, the schools are not teaching these concepts, partly because of a lack of adequate materials.

#### PURPOSE

The purpose of this study was two fold: (1) to test the feasibility of teaching topics in probability and statistics to a class of sixth-grade students; and (2) to construct a set of instructional materials and procedures in probability and statistics for sixth-grade students.

#### METHOD

The study used the working paper of Shepler *et al*, (1969) and the developmental model of Romberg and DeVault (1967) to build the unit. Shepler *et al* constructed a framework for the development of an instructional system in probability and statistics for use in the elementary school. That paper included a content outline, a task analysis of content, and specific grade recommendations for topics used in elementary school. The present study was designed to test the feasibility of parts of the working paper.

From strands of the task analysis, the author decided upon behavioral objectives for the unit of instruction and the order in which objectives would be taught. Using this basis, an instructional analysis of the unit was undertaken next. The purpose of this analysis was to select or develop materials and procedures for teaching the unit of probability to sixth-grade students.

To aid in the developmental processes of task analysis and instructional analysis, a pilot study was conducted in the fall of 1968. The data from the pilot study was used to identify a set of nine lessons that could be formatively evaluated to test the feasibility of the instructional analysis. The lessons were used to teach a class of sixth-grade students of average to above average ability. The goal of instruction was to demonstrate mastery learning of the behavioral objectives. By employing a pretest and posttest designed to measure the behavioral objectives, the feasibility of the unit was tested. In addition, certain learning hierarchies within the task analysis strands were examined using Walbesser's paradigm (1968).

## Conclusions

The conclusions of the study are considered in two parts:

1. conclusions related to the formative study with sixth-grade students.

2. conclusions related to the materials and procedures.

## CONCLUSIONS: FORMATIVE STUDY

On the basis of the criterion pretest and posttest results, one can conclude that the instructional treatment was highly successful regarding achievement on 10 of the 14 measured objectives (See Table 6, p.101). Instruction was moderately successful for Objective 10 (Law of Average-21/25 students scored 100% on Objective 10). In percentages, the average pretest score was 37.9 percent and the average posttest score was 92.8 percent. There was a marked change in student behaviors for all the measured objectives. On this basis the results of the study support the feasibility of teaching most of the included topics in probability and statistics to the group of students used in this study. However, three objectives were not close to meeting the stated criteria.

### Objectives Not Achieved

The following is a table presenting the breakdown for the three objectives that were not achieved. An analysis of each of the objectives follows the table.

TABLE 12

RESULTS IN CRITERIA FORM FOR
OBJECTIVES NOT ACHIEVED

| Behavioral Objectives | Criterion Percentage | Criterion Ratio Form |
|---|---|---|
| 2 | 72/80 | 18/25 scored 4/5 or 5/5 |
| | 88/60 | 22/25 scored 3/5, 4/5 or 5/5 |
| 8 | 52/100 | 13/25 scored 1/1 |
| 11 | 28/100 | 7/25 scored 1/1 |

While the gain from the pretest to the posttest score for behavioral Objective 2 is large, it is below the desired criterion. (See Table 13.) Only items (7, 2) (i.e. Part 7, item 2) and (9, 2) had good item percentages (100 percent and 88 percent respectively). The five items used to measure Objective 2 are in Figure 15.

TABLE 13

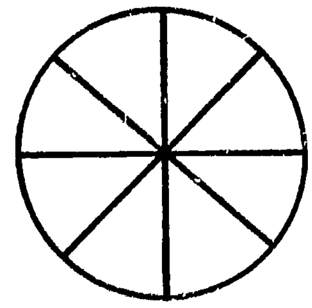
BEHAVIORAL OBJECTIVE 2
(Count the number of outcomes of an event)

| Number of Items | Pretest Percentage | Posttest Percentage |
|---|---|---|
| 5 | 31.2 | 79.2 |

Part of the difficulty in achieving Objective 2 can be explained by the lack of verbal communication of the items. During treatment, the few times students were asked to count the number of outcomes of an event, the question was phrased differently. For example, Item (12, 2) would have been worded "How many outcomes give you a sum of 2 or [illegible]" rather than "How many ways can one get a sum of 2 or 3?" All items on the test were phrased "How many ways . . . ."

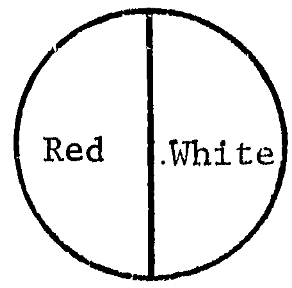
Part 7 For the following problems you spin the spinner below one time.

_____ 2. How many ways can you get a "1"?

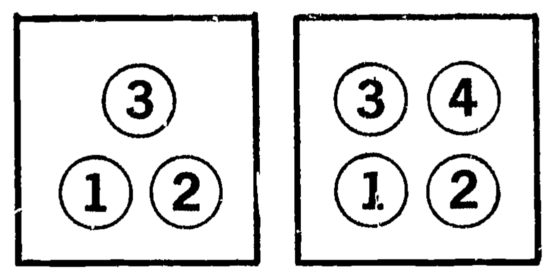


Part 8 For the following problem spin the spinner at the right two times.

_____ 2. How many ways can one get white on the first spin and white on the second spin?

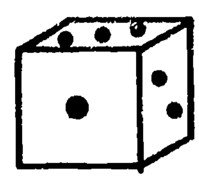



Part 9 For the following problems you pick a numbered chip without looking from Box A and one from Box B.




_____ 2. How many ways can one get a sum of 4 when the two numbers on the chips are added together?

Part 11 For the following problem a dice is thrown and the face which turns up is recorded.



_____ 2. How many ways can one get a "3" or a "5"?

Part 12 For the following problem, two dice are thrown, one red and one white. The sum of the faces turning up is recorded.

_____ 2. How many ways can one get the sum 2 or 3?

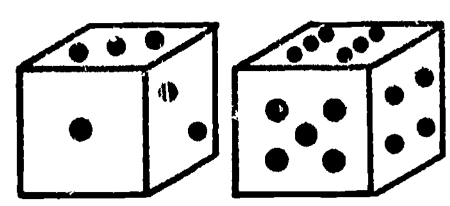


Figure 15. Items Designed to Measure Behavioral Objective 2 (Count the number of outcomes of an event.)

This seems to have led some students to mistranslate the meaning of these items. This mistake was particularly obvious in statements involving "and" and "or" (items (8, 2) (11, 2) and (12, 2)). For item (8, 2) the nine incorrect responses were all the same--1/4, the probability of the event rather than the number of ways of getting the event. Except for item (11, 2), all the incorrect responses for this behavioral objective listed the correct probability of the event rather than the number of ways of getting the event. With regard to item (11, 2) and its eight incorrect responses, two listed the probability and six responded "1."

In summary, possible explanations for these incorrect responses on items designed to measure Behavioral Objective 2 are:

1. The use of "How many outcomes . . . ?" rather than "How many ways . . . ?" during the treatment. During the test, even some of the most capable students asked about the meaning of this question. Some students were unsure of how to translate the statement.

2. The lack of emphasis on this objective. While the students had quite a bit of practice in counting the total number of possible outcomes, they had almost no practice in responding to items which asked them to find the number of outcomes in an event.

With regard to Behavioral Objective 8 (Specifying the estimated probability) the pretest item percentage was 8 percent and the post-test, 52 percent. Only one item was designed to test for this objective. (See Figure 16.)

_____ 1. In 6000 spins of the spinner at the right Bob gets 2653 reds. What is the estimated probability of getting a red on the next spin?

Figure 16. Item (13, 1) Used to Measure Objective 8

This should have been an easy item. Again part of the difficulty in achieving the objective was caused by the students never having had to respond to a written question concerning the objective on an exercise or quiz. Of the 12 incorrect responses, eight seemed to have tried to give an estimate of the probability (e.g., 1/2, 2/5, or 1/3), and two left the item blank.

The analysis of the only item measuring Behavioral Objective 11, (estimating the probability given a set of data) is of interest. (See Figure 17.)

_____ 1. Bob tossed a thumbtack 9,000 times with the following results:

6003--point up
2997--point down

Which of the following statements could Bob make:

a. The chances of the thumbtack pointing down is about 1/2.
b. The chances of the thumbtack pointing down is about 1/3.
c. The chances of the thumbtack pointing down is exactly 2997/6003.
d. The chances of the thumbtack pointing down is exactly 2997/9000.
e. Bob can make no statement at all about the chances of the thumbtack pointing down.

Figure 17. Item (14, 2) Used to Measure Objective 11

The pretest item percentage was 4 percent and the posttest 28 percent. The breakdown of the percentages on the alternates for the item on the posttest is:

a. 0%

b. 28%

c. 4%

d. 24%

e. 16%

(two answers--28%)

Due to lack of instructional time and students' ineptness with decimals, the teacher taught this objective using expository teaching procedures. Again the student was never asked on an exercise or quiz to respond to such a question.

It is interesting to note alternatives a and c are essentially eliminated by the students as being possibilities. All seven of the two letter answers include "b" as a possibility with 4/7 of the two letter answers responding (b, d). The students may be confused by the wording. Perhaps if the alternatives had been worded, "The probability of the . . ." instead, "The chances of . . .," more students would have been able to respond correctly.

In conclusion, the common thread which seems to run through analysis of why the three objectives were not achieved is the lack of practice given to the objectives in a written situation. The problem of verbal communication could, no doubt, be overcome if the students had been given practice and had their responses evaluated and corrected.

One-Dimensional and Two-Dimensional Sample Spaces

Conclusions concerning problems from one-dimensional and two-dimensional sample spaces are based on Table 7 (p.103) and observations of the lessons. The pretest results indicated that, initially, problems in two-dimensions were much more difficult than problems in one-dimension.

The data from the posttest in Table 7 indicated that the treatment was successful in teaching students to solve problems in both one-dimension and two-dimensions. However, there was a slight trend in favor of the one-dimensional percentages being slightly higher than the two-dimensional percentages.

From the Journal one can see that probability problems in two-dimensions were much more difficult to teach. More time, exercises and extra help were needed. The data from the quizzes and tests lend support to two-dimensional problems being more difficult for children to learn than one-dimensional problems.

Alternative Hypotheses

The possible reasons for the large gain between the pretest, posttest could be because of:

(1) the students cheating

(2) testing effect

(3) outside help

(4) treatment effect

(5) pretest-treatment interaction effect

The plausibility of cheating causing this large gain is not acceptable because the students had no opportunity to get a copy of the test beforehand, and the chances of poor students copying from good ones and allowing everyone to get a good score is near zero. Copying was almost impossible since the test was monitored and each of the two parts of the test had four random orderings of the parts which in turn were assigned to every fourth student. This meant that students were working on different items at the same time.

In regard to (2), the testing effect, the author is willing to grant a small part of the gain could be attributed to the test-retest effect, particularly since the same test was used for the pretest and the posttest. However, since the vocabulary of the test was foreign to the student, there probably was little recall of the 72 items over a four week span. It does not seem plausible that a gain in raw score from 27.28 to 66.80 was caused by the testing effect alone.

As far as (3), outside help, many of the assignments were completed in class with no opportunity for outside help. Besides, very few adults are familiar with the concepts of probability to the extent that they could easily aid the students.

Thus the large gain must be attributed to either (4), treatment effect, or (5), pretest-treatment interaction effect. The author has no way of distinguishing between these two possibilities. However, he feels that most of the large gain can be attributed to the treatment.

Testing of Learning Hierarchies

It was hypothesized that in order to identify when two events are equally likely or which of two unequally likely events is more likely, $\boxed{M}$, that the child must learn to specify the probability of an event, $\boxed{P}$, and be able to order two ratios, $\boxed{R}$. (See Figure 4, p.46).

In order to verify the learning hierarchy, three ratios were computed at two different levels of mastery. (See Table 14.)

TABLE 14

RATIOS FOR TESTING HIERARCHY 1
USING TWO CRITERIA

| | 90% Test Criterion | | Practical Test Criterion* | |
|---|---|---|---|---|
| | Ratio | Ratio Decimal Found | Ratio | Ratio Decimal Found |
| Consistency Ratio | $\frac{15}{18}$ | .83 | $\frac{21}{23}$ | .91 |
| Adequacy Ratio | $\frac{15}{21}$ | .71 | $\frac{21}{23}$ | .91 |
| Completeness Ratio | $\frac{15}{16}$ | .94 | $\frac{21}{21}$ | 1.00 |

* ($\frac{14}{16}$ for $\boxed{P}$, $\frac{9}{10}$ for $\boxed{R}$ and $\frac{16}{18}$ for $\boxed{M}$)

Using the arbitrary test criterion of 90%, only the completeness ratio (.938) is above the standard of .90. Using the practical test criterion all three ratios are above .90, and the learning hierarchy is validated under the paradigm of Walbesser.

The validation of the first hierarchy supports the successfulness of the study. Specifying the probability of an event, then using this skill in making a decision between two games were major objectives of the study. The high achievement in ordering two ratios has major implications. No formal instruction was given on the objective. Yet the pretest average score was 78.8% while the posttest average score was 96.0%. One would speculate that other arithmetic skills such as adding two fractions could be improved by using the skills in probability problems. This result also lends support to Rationale 6 in Chapter I (p. 3 ) for including probability and statistics in the elementary school.

The second hierarchy that was tested hypothesized that to learn to specify the probability of an event, $\boxed{P}$, one must specify the number of outcomes of an event, $\boxed{E}$, and specify the number of possible outcomes of the sample space, $\boxed{S}$. (See Figure 5, p. 46.)

The following two sets of ratios were computed as before.

TABLE 15

RATIOS FOR TESTING HIERARCHY 2
USING TWO CRITERIA

| | 90% Test Criterion | | Practical Test Criterion* | |
|---|---|---|---|---|
| | Ratio | Ratio Decimal Found | Ratio | Ratio Decimal Found |
| Consistency Ratio | $\frac{7}{23}$ | .33 | $\frac{16}{25}$ | .64 |
| Adequacy Ratio | $\frac{7}{8}$ | .88 | $\frac{16}{16}$ | 1.00 |
| Completeness Ratio | $\frac{7}{8}$ | .88 | $\frac{16}{16}$ | 1.00 |

*($\frac{4}{5}$ for $\boxed{E}$, $\frac{6}{7}$ for $\boxed{S}$, $\frac{14}{16}$ for $\boxed{P}$)

All three ratios are below .90 under the 90% test criterion while only the consistency ratio is low under the practical test criterion because of relatively poor performance of students on Behavioral Objective 2 (i.e., $\boxed{E}$--Specify the number of outcomes of an event). The mean percentage for the five items measuring $\boxed{E}$ was 79.2%. Because of reasons cited earlier of poor wording of the items measuring $\boxed{E}$ and lack of training given for the behavioral objective, one cannot conclude one way or another whether the learning hierarchy is valid. However, the author feels that a treatment emphasizing $\boxed{E}$ and the use of better test items to measure $\boxed{E}$ would result in the validation of the hierarchy.

Major Objectives Not Measured by Test or Quizzes

The major objectives which were not measured by the quizzes or the pre and post-test are the following:

8b. Specify the estimated probability of an equally likely outcome.

12. Identify experiments which are equivalent (i.e., those with the same sample space).

13. Construct a bar graph of data from an experiment.

14. Interpret the bar graph of data from an experiment.

15. List the possible outcomes of an experiment by employing a tree.

16. Identify that the probability of an event E, P(E), tells one that for a large number of trials the event will occur in a ratio approximately equal to P(E).

17. Identify that for a large number of trials the estimated probability approaches the *a priori* probability.

18. Identify that for a large number of trials that the cumulative frequency of a more likely event should be larger than that of the less likely event.

19. Specify the likely bounds on the _a priori_ probability of an event from the graph of the data from an experiment.

By observation and an analysis of exercises the author has evidence that students could make a list of the possible outcomes of the experiment after drawing a tree diagram (Objective 15) and given minimum help that students could construct a bar graph (Objective 13).

Objectives 8b, 12, 14 and 18 were covered by class discussion. Evidence based on classroom observation shows that the students answering the question met these objectives. How many other students in the class could have met the objectives? The investigator has no evidence to answer this question.

Objective 16, 17 and 19 were covered by expository methods with the teacher emphasizing each objective. The three objectives were covered in this fashion because of time limitations. (The objectives are contained in Lesson 9, the last lesson.) Thus the investigator has no evidence whether students obtained these three objectives.

## CONCLUSIONS: DEVELOPMENTAL PROCEDURES AND MATERIALS

In the opinion of the author, major reasons for the large gain in raw score can be attributed to the developmental analysis used and the mastery learning techniques that were employed. In the author's opinion, developing a curriculum through the following sequence is an excellent way of building research based curriculum materials. Start with a content outline and establish behavioral objectives. Task analyze these objectives and write an instructional treatment to meet them. Proceed to the important step of actually trying these materials with children, while recognizing the possibility of iteration through preceding steps.

The developmental model encourages modifications of materials and procedures based on empirical evidence. The following indicates modifications needed in the unit in light of observations and test analyses.

## Major Suggested Changes

On the basis of the formative evaluation, the investigator recommends the following changes be made in the materials.

1. For Lessons 1-9, identifying and specifying the expected value should have been a behavioral objective.

2. For Lesson 2, more children in a committee would cut down on the number of models needed. This would make the lesson more manageable for a single teacher.

3. For Lesson 3, the use of transparencies or graph paper rather than large poster paper would cut down on the amount of time to make a graph. Also with experiments with two possible outcomes, a bar need only be drawn for one of the possible outcomes.

4. For Lesson 4, distinction between the attributes of manyness and other characteristics of a model needs to be made. More emphasis should be placed on the number of possible outcomes of an experiment and of an event.

5. For Lesson 5, if students know how to graph coordinates, a geometrical picture of a sample space consisting of points might be a better approach than the one used by the investigator. Also the exercise accompanying Lesson 5 should have the first page replaced by an easier problem. More problems like problems on pages two and three of the exercise need to be included for more practice. Greater emphasis on the distinction between one and two-dimensional problems may need to be made in Lesson 5. If so, exercises mixing one and two-dimensional problems will need to be added in the exercise for Lesson 5. (In the study, the mixing of one-dimensional and two-dimensional problems was done in Lesson 6.) If these ideas are incorporated, Lesson 6 might not be so difficult. For Lesson 6 Part II, more problems on three and four-dimensional models using coins, spinners, and other simple models need to be included. Problems concerned with replacement and without replacement of objects drawn from a box should also be included in the exercises. In lesson 6, Exercise 3, more work on verbal problems is needed.

7. For Lesson 8, more time should be spent in doing some experiments as many as 2000 to 3000 times.

8. For Lesson 9, children should have the prerequisite skill of ordering and graphing two decimals expressed in hundreths or thousandths. The children should also be able to express a fraction as a decimal correct to the nearest hundreth or thousandth. This skill could be achieved by teaching students to use a desk calculator.

Lesson 9 should be student centered rather than teacher centered as in this study. The students should graph and interpret their results with the teacher asking questions at appropriate places. Also, another exercise for Lesson 9 needs to be constructed which gives students practice in answering questions concerned with all the objectives of Lesson 9.

9. Better ways of testing many of the major objectives of the unit need to be worked out.

Mastery Learning

The powerful technique of employing Bloom's suggestions for achieving mastery learning has, in the opinion of the author, exciting possibilities for raising the achievement of children's performance. The techniques that seemed to be the most valuable were:

(1) the mastery, non-mastery grading.

(2) the opportunity of the child to achieve mastery of a set of objectives even if he did not succeed the first or second time.

(3) the emphasis placed on the child achieving the prerequisite knowledge before proceeding to a task that requires mastery of that knowledge.

(4) the goal chart that helped the student understand the purpose of instruction.

(5) the three extra-help sessions conducted for the nonmasters.

Criteria

The results of the study indicate that stating an arbitrary 90/90 criterion as the only judge of successful instruction should effect the number of subjects one uses in a study and the number of items one uses to measure an objective. For example, in this study there were two students in the group of 25 who had achieved D and D- averages in arithmetic. For most quizzes and exercises, neither would initially be classified as a master of an objective. This implied that if exactly one more person was classified as a nonmaster then only 22/25 students (88%) achieved mastery. Hence, using the 90/90 criterion, instruction would be judged unsuccessful. A practical criterion based on sample size, the number of items needed to get a stable measure of an objective and the importance of the objective are, in the opinion of the author, the factors that should be used to set a behavioral objective criterion.

LIMITATIONS OF THE STUDY

This study is clearly not generalizable. The presence of two people in a classroom to aid instruction given to a selected group of 25 students in a small suburban-rural school is not typical. The study does not answer the question as to whether an existing elementary teacher could take the materials and achieve the same results with a similar selected group. The question as to whether an existing elementary teacher could teach the materials to a group of children in a deprived area or in middle class suburbia is even more obstruse. The study has only indicated that given ideal conditions for the quality of teaching and a somewhat select group of students that the teaching of these materials could be feasible. The study has demonstrated that inspite of some developmental weaknesses in the materials that the

25 students at Waunakee were able to achieve a high level of performance concerning most of the objectives measured by the posttest.

## RECOMMENDATIONS AS TO THE USE OF THE MATERIALS

The typical elementary teacher should be able to teach the lessons on *a priori* probability in one-dimension with little problem. However in order for her to teach probability problems in two-dimensions successfully, the lessons will need further developmental research to determine a more feasible way of presenting problems in a two-dimensional sample space.

In light of the author's experience in training the elementary teacher used in the study, he would recommend this: put reservations on use of graphing situations calling for subtle interpretations. Lesson 9 was to merge the estimated probability of an event to what the probability of an event means in the real world. The lesson was probably too subtle and difficult for a teacher who has not had substantial background in that area.

## RECOMMENDATIONS FOR FURTHER STUDY

There are many studies that could be done in light of the results of this one. There is a need for developing better lessons for presenting probability problems in n-dimensional sample space. The materials in this study could be tried by other experimenters with different types of children to test the feasibility of using the materials with children of different ages and background.

This study explored only a small part of the proposed content in probability and statistics for the elementary schools. Other formative studies using different concepts and different strands of the proposed task analysis (Shepler, et al, 1969) should be done to test their feasibility. This latter idea also suggests that the whole realm of testing learning hierarchies and better instruments for measuring the various types of behavioral objectives in probability and statistics need to be developed.

What is the attitude of children toward properly taught probability and statistics concepts? What change in attitude might take place in light of such teaching? These could be fruitful areas for further research.

The feasibility of the developmental research model of Romberg and DeVault in creating materials in different areas should be explored.

Related to the development and testing of the feasibility of a new set of curriculum materials is the question of efficiency-whether what has been learned is transferable to new problem situations outside and within mathematics. For example, would it be easier and quicker to teach the group of 25 students in this study a more formal unit in probability in the junior high? Can the use of rational number concepts in doing probability problems increase one's proficiency in using the rational number concepts?

The powerful ideas embodied in Bloom's suggestions for achieving mastery learning should be widely explored using different areas of mathematics and children of various backgrounds and abilities.

There is need for further research regarding effect of behavioral objectives on mathematics teaching, and also concerning achievement and the affective domain of the learner. These suggestions by no means represent an exhaustive list of possible studies. As is usually true with education research, the present study has raised more questions than it answered.