DOCUMENT RESUME

ED 135 807 TM 005 899

AUTHOR Shann, Mary H.
TITLE Measuring Problem Solving Skills and Processes in Elementary School Children.
INSTITUTION Boston Univ., Mass. School of Education.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE Jun 76
NOTE 265p.; Material has been removed from the appendices due to copyright restrictions

EDRS PRICE MF-$0.83 HC-$14.05 Plus Postage.
DESCRIPTORS Cognitive Processes; Elementary Education; *Elementary School Students; Interviews; Item Analysis; Manuals; Measurement Techniques; Models; Observation; *Problem Solving; Scoring; *Test Construction; Test Reliability; *Tests; Test Validity
IDENTIFIERS *PROFILES; *Test of Problem Solving Skills; Unified Science Mathematics for Elementary Schools

ABSTRACT

Grounded in a review of existing theories and research on problem solving, the theoretical base and new instrument development efforts discussed in this publication have been sounded against the needs of an innovative, interdisciplinary curriculum project called Unified Science and Mathematics for Elementary Schools (USMES). It is the work on new instrument development for problem solving which is the focus of this document. The report is addressed to those concerned with the evaluation of USMES but also to a wider audience whose concerns may embrace the evaluation of other curricula for elementary schools, research on child development, or theoretical development of models of problem solving. After establishing the need for new instrument development in problem solving, discussing various views of problem solving, and reviewing existing measures of problem solving, this report details the development of a paper-and-pencil Test of Problem Solving Skills (TOPSS) and the development of PROFILES: an interview/observation technique to assess problem solving processes in children. The chapter on TOPSS includes the search for items, pilot testing the instrument, and technical information on item analysis, reliability and validity. The chapter on PROFILES includes information on its rationale, its development, the importance of observer training and monitoring, procedures for its administration, and the development of a scoring protocol. (RC)

TM

MEASURING PROBLEM SOLVING SKILLS AND PROCESSES IN ELEMENTARY SCHOOL CHILDREN

Mary H. Shann, Ph.D.
Associate Professor

The School of Education
Boston University

June, 1976

PREFACE

Knowledge about problem solving has been an elusive prize of perennial interest to psychologists and educators. Yet without sufficient tools to investigate the complex abilities and processes in efficient problem solving, researchers have been limited in their ability to enhance the state of knowledge about problem solving, and educators have been pressed to implement unproven techniques to develop this highly prized, higher mental process in students of all ages. The USMES Evaluation Project, from which this book evolved, was directed by the writer who, as a research psychologist and an educator, is sympathetic to the difficulties experienced by both groups regarding problem solving. It is hoped that the theories and measurement techniques presented herein can help to advance the field of research on problem solving, with results that can be directed toward the improvement of educational practice.

Grounded in a review of existing theories and research on problem solving, the theoretical base and new instrument development efforts discussed in this book have been sounded against the needs of an innovative, interdisciplinary curriculum project called Unified Science and Mathematics for Elementary Schools (USMES). As such, this book, built on theory tempered in practice, may be useful as a supplement to courses for prospective and inservice educators, expecially in the study of educational psychology, measurement, curriculum development, and curriculum evaluation.

While the text is essentially the work of one author, it could not have been written without the efforts of many more people. First, I would like to thank the National Science Foundation for extending its auspices and funding to the USMES Evaluation Team for the investigation of the USMES curriculum's effects and for the new instrument development in problem solving. It was this work and these responsibilities which captured my interest and heightened my concern for the study of problem solving processes. Dr. Raymond J. Hannapel is the person at the Foundation who sanctioned the new instrument development work by the evaluation team whose initial set of responsibilities included only the evaluation of USMES.

The USMES Central Staff at the Education Development Center in Newton, Massachusetts, developers of USMES, were thoughtful and gracious in their response to requests from the USMES Evaluation Team for clarification of their views on problem solving as practiced in USMES. Their ambitious, dedicated efforts have been directed toward the development of a curriculum which might help educators in turn help children to address the complex problems of an increasingly complex environment. Their efforts created for the USMES Evaluation Team a forum and field sites in which to discuss and test views on problem solving.

The Advisory Board to the USMES Evaluation Project offered thought provoking assessments of the evaluation team's chances for "success" in the enormous work of new instrument development. Whether his preference was the route chosen, each advisor offered genuine support

and encouragement for new instrument development. I would like to extend sincere thanks to Professor Jeremy Kilpatrick of the University of Georgia who encouraged the Team to devise new measuring instruments and to Professor Fletcher Watson of Harvard University and Professor Wayne Welsh of the University of Minnesota who renewed our concerns as to why other competent investigators have not yet been able to measure complex problem solving satisfactorily. The balanced perspectives they offered were most helpful.

Various members of my project staff, "the USMES Evaluation Team," deserve the greatest acknowledgment for their work on new instrument development, without which the first three chapters of this book would not have been written, and without which the last two chapters of this book could not have been written. Mrs. Susan Rogers is credited for her work on the supervision of item selection and item writing for TOPSS, the Test of Problem Solving Skills, presented in Chapter IV and for her creative efforts in the conception and development of PROFILES, the observational/interview technique to study children's problem solving processes, which is discussed in Chapter V. Ms. Linda Hench assisted Mrs. Rogers in both of these efforts, and she pilot tested the PROFILES technique in several classes in the Greater Boston area. Miss Anita Paci, the former Administrative Assistant for the Evaluation Team, deserves thanks for distributing, receiving, and processing the testing materials used in the pilot testing of TOPSS.

The teachers and their students who graciously agreed to participate in the trial of both problem solving instruments should be

acknowledged with sincere gratitude because they offered their time with no incentive other than the distant hope of enabling better research and better education in problem solving. Appreciation is also expressed to Mrs. Dorothy Bowler, Ms. Christine Reali, and Mrs. Jean O'Connor who facilitated our entree into the sample classes for data collection.

Dr. Norma Reali, a great friend and respected colleague, offered tremendous encouragement and humored critique as I wrote the pages of this book. The Associate Director for the USMES Evaluation Project, she served as a consultant upon completion of her tenure in that position. I credit her with the management of data collection and analysis of the problem solving tests, but more importantly for her advice in how I should shape this book into a manageable and yet meaningful piece of writing.

Another colleague, Professor Hilary Bender of Boston University offered the benefit of his "way with words" in the improvement of earlier drafts of the chapters on the new instruments themselves. His able, timely review made my tasks with these chapters easier.

It has been a great pleasure to work with Ms. Beth Ingram whose outstanding speed and skill in typing, proof reading, and assembling this document has made the completion of this book a surprisingly pleasant and relatively painless experience for me. Her very able assistance at this normally demanding, stressful phase of a project has heightened my sense of satisfaction in the completion of the very long, arduous efforts reflected and reported herein.

Finally, my deepest gratitude goes to Bob Shann who endured with little complaint the problems of measuring problem solving, to Ryan Shann who was cheerful and understanding like his father, and to the Lazzari family, especially "Gram," whose loving care of Ryan enabled me to work productively, confident that Ryan was going to settle the world's problems without me at least as well as he could with me.

June, 1976
Boston University

Mary H. Shann
Associate Professor

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE NEED FOR NEW INSTRUMENT DEVELOPMENT IN PROBLEM SOLVING

The Problem of Definition

The term problem solving has been used variously to describe a method of attack, the synthesis of isolated experiences, the reorganization of cognitive structure, productive thinking or simply a moment of intuition. Some authors, such as Getzels (1964), take the position that all learning is problem solving. Davis (1966) argues that problem solving is any semi-complex task which has not already been identified by another name. Others reserve the term to label a more complex mediating process which individuals must pursue to reach a goal which is not readily or immediately attainable. (Duncker, 1945; Wertheimer, 1959; Duncan, 1959; Travers, 1967; Ausubel and Robinson, 1969; Bourne, Ekstrand, and Dominowski, 1971; Feldhusen, Houtz, and Ringenback, 1972; and Speedie, Treffinger, and Feldhusen, 1973.) Even among the later group, however, what exactly constitutes this process is a matter for considerable debate.

The fact that this construct eludes easy definition or description probably accounts for the enormous, amorphous literature on "problem solving." More than a decade ago, Weir commented:

> "The psychological literature on problem solving and related topics such as concept formation is vast and appears to be growing exponentially. Attempts to adequately review this literature

> have been hampered by the noticeable absence of systematic research programs which attempt to isolate important variables influencing this type of behavior" (Weir, 1964, p. 473).

The Problem of Measurement

Research on complex problem solving has been hampered too by the inadequacy of measuring instruments and observational assessment techniques to study problem solving behaviors. In reviewing the tasks available for use in human problem solving research in the early 1950's, Ray was struck with the observation that many of these tasks have been of the parlor puzzle or game variety. Even more disheartening was his observation that: "Reliability has been mentioned only twice and validity never by the writers of the reports" (Ray, 1955, p. 148). A more recent review of problem solving tests undertaken by Feldhusen, Houtz, Ringenback, and Lash (September, 1971) revealed only little advancement in the state of the art of measuring problem solving since Ray's review was published. A greater assortment of problem solving tests is now available. These tests include measures of components, of processes, and of performance on simulated real-life problem tasks. However, reliabilities and validities remain typically low or unstudied (Speedie, Treffinger, and Feldhusen, 1973, p. 34).

Theory and Practice in Teaching Problem Solving

With serious obstacles to the empirical testing of models of problem solving, educational practice continues to outstrip theory in this area. Psychologists can continue to pursue the "perennial challenge"

offered by the description of problem solving (Merrifield, Guilford, Christensen, and Frick, 1962, p. 1), but educators have had to respond to the more immediate challenge of helping students to deal effectively with the complex problems of an increasingly complex environment.

The interdisciplinary, process curriculum called Unified Science and Mathematics for Elementary Schools (USMES) constitutes one major response to that challenge. Funded by the National Science Foundation and coordinated by a staff at the Education Development Center in Newton, Massachusetts, the USMES project purports to develop the problem solving abilities of elementary school students. The goal of the USMES program is the development of thirty-two interdisciplinary units engaging the students in long-range investigations of real and practical problems taken from their school or community environment. By acting on these problems, called "challenges," the student is supposed to develop his problem solving abilities, and do so in a manner that gives him an experiential understanding (learning-by-doing) of the problem-solving process, as well as the acquisition of its basic skills and concepts.

The Problem of Evaluating a Curriculum for Problem Solving

The National Science Foundation has also sponsored the independent evaluation of the USMES program, and NSF's overriding concern for this curriculum evaluation project has been the investigation of "proof of concept." Yet, the determination of whether or not USMES

is increasing students' abilities in complex problem solving is dependent on the availability of adequate appropriate measures of children's performance on complex problem tasks. Apprised of the inadequacy of existing assessment devices, the Foundation agreed to support the work of the writer and her staff at Boston University not only for the evaluation of USMES but also for the development of new instruments to measure the problem solving abilities and processes engaged in by elementary school students.

A separate report on the evaluation of a variety of student effects of the USMES program during the 1974-75 academic year has been prepared (Shann, et al., December, 1975). An earlier report for the 1973-74 academic year is also available (Shann, August, 1975). It addresses broader issues including patterns of USMES usage, teacher's appraisals of the program's effectiveness, the utility of USMES materials, the problems with USMES teacher training, and indirect effects of USMES implementation. Non-technical summaries of both of these lengthy evaluation documents can be obtained from the Project Director.

It is the work on new instrument development for problem solving which is the focus of the present document. We address this report to those concerned with the evaluation of USMES but also to a wider audience whose concerns may embrace the evaluation of other curricula for elementary schools, research on child development, or theoretical development of models of problem solving.

CHAPTER II

VIEWS ON PROBLEM SOLVING

General Perspectives for USMES

The theoretical base of a test selected or developed for use in the evaluation of USMES should be compatible with the theory of problem solving to which the USMES developers subscribe. However, the evaluators' success in providing a criterion measure which the USMES developers could endorse for the evaluation of the USMES curriculum was frustrated on two fronts. Neither could we locate any reliable test of complex problem solving appropriate for use with children, nor had the developers yet articulated a careful description of their view of the components of problem solving or of the structure of the problem solving process.

If the point of USMES is the development of children's abilities to solve real, complex problems, then, arguably, there must be a generalizable structure of problem solving which children can learn through their involvement in USMES challenges. Furthermore, the structure must be transferable to aid the children's solution of other real world challenges experienced outside of the USMES curriculum, especially outside of the school environment. Evidence of such transfer of the structure of problem solving would constitute the ultimate "proof of concept" for USMES. The challenge to the evaluation team was to identify the critical features of that structure, or, at the

very least, to enumerate the components of real complex problem solving thought to be the essence of USMES.

The USMES developers assisted us in this effort. As we renewed our literature search on problem solving theory and measurement, the developers worked to clarify their thinking on the components of the USMES-styled, real problem solving process and on the interaction among the phases. Our plan was to use the USMES Guide and other printed USMES materials, visits to USMES classes, and most importantly, conferences with the USMES developers as initial check points for determining what notions of problem solving in the literature were consistent with USMES.

The body of literature on human problem solving is enormous and disorganized; its review is a very difficult task. Several comprehensive critiques are helpful for acquainting the reader with the research in this field (Duncker, 1945; Duncan, 1959; Davis, 1966; and Gagne, 1970). Undoubtedly the lack of agreement on definitions and the limited state of the art of measuring problem solving, especially complex problem solving, contribute heavily to impeding the advancement of sustained research efforts in the field.

Problem solving may be used for responding to complex situations, doing homework exercises, solving puzzles, creating inventions, or resolving interpersonal conflicts. Problem solving activities may include planning, analyzing, creating, evaluating, persuading or detecting. Like the overworked, often misunderstood concept, "intelligence," problem solving is commonly used in a variety of contexts.

Although problem solving by its semantic nature is itself a process, "good" problem solving ability is generally believed to be an attribute not unlike "high intelligence" in its potential applications and benefits to the individuals who possess it. However, unlike "intelligence," which can be defined in terms of its value for predicting academic success, the concept of problem solving has been derived largely through logical analysis, and the contexts in which problem solving is discussed shed some light on meanings ascribed to the term.

We surveyed theoretical and empirical studies from basic and applied disciplines -- philosophy, psychology, and education, including science education, mathematics education, and interdisciplinary education. Our purpose in researching this vast, disorganized literature was not to summarize a review of the works, many of which had little relevance for USMES. Rather, our purpose was to sift out theories, models, and hopefully measures of problem solving which seemed to be consistent with the USMES notion of problem solving and which might help us clarify the process as it is applied by children in the USMES curriculum.

Problem Solving in Philosophy

The USMES literature itself suggested a review of the term problem solving in the writings on educational philosophy and on philosophy of science. The USMES philosophy is an eclectic one. It encompasses features of the theoretical positions expressed by Dewey, Bruner, Gagne, and others. Most consistently evident in the USMES developers' written

statements about the USMES approach, however, are references which call to mind John Dewey's "five logically distinct steps" of the problem solving process:

1. a felt difficulty
2. its location and definition
3. suggestion of possible solution
4. development by reasoning of the bearings of the suggestion
5. further observation and experiment leading to its acceptance or rejection, that is, the conclusion of belief or disbelief (Dewey, 1910, pp. 72-77).

The USMES developers' parallel to Dewey's conceptualization is illustrated in the following tements from The USMES Guide (EDC, May, 1974):

> "Since its inception in 1970, USMES has been developing and carrying out trial implementations of interdisciplinary units which are based on long-range investigations of real and practical problems taken from the local school/community environment.... In responding to these real problems, called challenges, students are involved in all aspects of problem solving: observation, collection of data, representation and analysis of data, formulation and trial of successive hypotheses, and decision on a final action to be taken" (p. 1).

And from the second chapter of the Guide:

> "The children must develop the various aspects of good problem solving, including definition of the problem, observation, measurement, collection and analysis of data, acquisition of needed skills, development of judgement, formulation and trial of possible solutions, and decision on final action to be taken" (p. 15).

A. Problem Solving as Logical Thinking

Clearly the USMES developers based their descriptions of problem solving on John Dewey's logical analysis of the reflective method of thinking. His theory of problem solving has had an impact upon education that is as influential today as it was in 1910. The resilience of Dewey's theory is attested by Guilford. In his analysis of creative production, Guilford (1965, p. 8) compared several more recent theories and concluded that the most remarkable thing about them was their similarity to those of Dewey.

Dewey's basic interest lay in contrasting the scientific method of thinking to "chance and idle thinking" (Dewey, 1910, p. 3) and other methods of arriving at belief, such as reference to authority or tradition. The method of thinking outlined by Dewey is a set of logical guidelines for arguing that a particular hypothesis should be accepted or rejected. A "true" conclusion is not guaranteed in any absolute sense, but a strong logical case can be made by this method.

Despite this frequent paraphrasing of Dewey's five steps in the descriptions of USMES, one cannot equate the generalizable structure that presumably underlies and unifies USMES challenges with Dewey's method of reflective thinking. At first glance, it appeared that the aim of USMES was to teach the structure of problem solving as logical or reflective thinking. Closer scrutiny dispelled this notion. At times, USMES uses the category "Problem Solving/Logical Reasoning" (USMES Bicycle Transportation, 1974, p. E-11) as a major activity.

However, the essential element of supporting or rejecting hypotheses through the rules of logic is not a major activity in any USMES challenge. In some units, the formulation of hypotheses which are capable of refutation does not occur at all.

B. Problem Solving as the Scientific Method

The descriptions of problem solving in USMES also called to mind the scientific method. It was possible that the USMES developers considered "real problem solving" to be the application of the scientific method to everyday dilemmas. Certainly many USMES activities are similar to the activities pursued by scientists, such as observation, measurement, and data analysis. In fact, the USMES project was formed in response to the recommendations of the 1967 Cambridge Conference on the Correlation of Science and Mathematics in the School, which supported implicitly the teaching of the structure of the scientific method:

> We believe that the primary goal of science education is an understanding of scientific methodology.... Since we hope that most of the effort will be spent on teaching the scientific method, we must expect that children will acquire only a small sample of scientific knowledge (Goals for the Correlation of Elementary Science and Mathematics, 1969, p. 9).

If the scientific method is posited as the unifying structure of problem solving practiced in and transferable from the USMES challenges, then one must consider the ongoing debate regarding the idea that there is one generalizable scientific method. Conant (1961, p. 45) has argued that there are as many scientific methods as

scientific endeavors. However, we took the position that there is a scientific method whose general characteristics have been identified, explained, and formalized successfully by philosophers, mathematicians, and logicians (e.g., Cohen and Nagel, 1934; Nagel, 1961; Hempel, 1966; Foster and Martin, 1966; and Martin, 1972). As Martin argues in his philosophical analysis of contemporary science education, there are:

> "...important structural similarities among the procedures of the anthropologist, astronomer, sociologist, biologist, and so on. All these scientists test their hypotheses by deducing consequences from them, together with auxiliary hypotheses. Moreover, the general criteria of confirmation or refutation of a hypothesis do not differ in the various sciences" (Martin, 1972, p. 43).

It is the exacting criteria of confirmation or refutation of hypotheses which unifies the work of all scientists not their activities -- observation, measurement, recording, data analysis, and so on. In contrast, USMES uses not the rigorous standards for accepting or rejecting hypotheses but the pragmatic criterion, "If it works, it's right." Concluding that "real" problem solving in USMES was not synonomous with either logical thinking or the scientific method, we looked to the research literature of psychologists to study the process of problem solving and to gather ideas for its measurement. By education and experience biased more toward the psychological literature, the writer and project director was receptive to Gagne's advice which he offered in his review of "Problem Solving and Thinking:"

> "There appear to be the beginnings of a solid experimental literature on this very old psychological topic. The serious student who enters

> the maze of intellectual functions, if he is able to become negatively adapted to the philosophical noise which still persists, learns to thread his way through the obfuscatory traps of terminology, and overcomes the temptations of goal-oriented mathematical cul-de-sacs, will obtain at least substantial aperiodic reinforcement from this literature" (1959, p. 147).

Problem Solving in Psychology

A. Models of Problem Solving

Psychologists theorizing on problem solving have postulated models of the process which seem to follow either of two approaches: (1) the identification of subgroups of intellectual processes linked in some order, whose linkages may be of the linear or feedback variety; and (2) the consideration of problem solving as the operation of an information processing system.

1. Models Identifying Component Intellectual Processes

Those psychologists and other theorists thinking in the first vein reflect Dewey's influence, since the phases they propose for problem solving are almost invariably the condensation, elaboration, or retitling of Dewey's five logically distinct steps (1910). Wallas (1926), a philosopher whose work is cited in the psychological literature, proposed four phases: preparation, incubation, illumination, and verification. Polya (1948), a mathematician, also listed four steps toward the solution of a problem: understanding the problem; working out connections between the known and unknown, thus deriving a plan of solution; carrying out the plan; and examining the solution. Johnson (1955)

suggested three broad classes of problem solving processes: preparation, production, and judgement. He affirmed these steps as a basis for a systematic presentation of the facts concerning problem solving in a more recent work (Johnson, 1972).

Acknowledging their debt to this kind of theorizing which attempts to isolate logically distinct phases of problem solving, Merrifield, Guilford, Christensen, and Frick (1962) envisaged five phases of the total problem solving process: preparation, analysis, production, verification, and reapplication. They viewed these phases not as clear-cut, successive steps, but as conceptually different operations with an approximation to temporal ordering and with much overlapping of particular events.

Gagne (1959) conceived of five phases of problem solving behavior: reception of the stimulus situation; concept formation or concept invention; determining courses of action; decision making; and verification. But Gagne noted that he could not follow this outline of phases as the framework for his review of problem solving literature "because some phases have not been studied systematically" (1959, p. 148). Later he described problem solving as the complex interaction of a number of subordinate learnings which lead to the learning of a new rule (Gagne, 1964).

A final model of problem solving which reflected the approach of identifying subgroups of intellectual processes was developed by Feldhusen, Houtz, and Ringenbach (1972). Based on their review and

synthesis of the problem solving literature, they postulated a number of specific behaviors which seemed to be involved in all phases of the problem-solving process. Their list included twelve abilities:

(1) Sensing that a problem exists;

(2) Defining the problem;

(3) Clarifying the goal;

(4) Asking questions;

(5) Guessing causes;

(6) Judging if more information is needed;

(7) Noticing relevant details;

(8) Using familiar objects in unfamiliar ways;

(9) Seeing implications;

(10) Solving single-solution problems;

(11) Solving multiple-solution problems; and

(12) Verifying solutions.

2. Models Focusing on the Operation of Problem Solving Processes

The other models of problem solving referenced below might be described as information processing models. Instead of identifying the components or subgroups of intellectual abilities that one employs during the complex process of problem solving, the information processing models focus on the operations which take place during problem solving.

The most elemental of these information processing models was the TOTE feedback system presented by Miller, Galanter, and Pribram (1960).

They postulated that the TOTE was the basic process used to solve all problems confronting human beings. Test a given condition; if the condition is not satisfied, perform some one of a small class of operators and test the condition again; exit when the criterion is met. Thus, the TOTE model does not specify any particular abilities involved in problem solving but only the form the processes take -- test operate, test, exit.

The model proposed by Guilford (1967) is considerably more elaborate. Based on his Structure of Intellect model, the Guilford model of problem solving is an information processing model in which various intellectual processes act upon environmental and somatic input, with continual reference to memory, to generate problem solutions. Filtered by arousal and attention mechanisms, the input for the problem can be ignored, or the problem can be sensed and structured through cognition. The cognition processes may call for new input or yield to evaluation processes. If the problem is sustained after passage through this stage, answers may be generated by convergent or divergent production. Then these answers are evaluated, and either a new cycle of cognition, production, and evaluation is pursued, or the problem is exited. Potential exit points follow each phase, and accession to memory storage underlies all of the processes in Guilford's model.

Newell and Simon (1972) derived an information processing model of human problem solving from their extensive work on computer simulation of human reasoning in three task environments: chess playing,

discovering proofs in logic, and cryptarithmetic. To capture the "bones" of their theory, Newell and Simon offered four propositions:

(1) A few, and only a few, gross characteristics of the human information-processing system are invariant over task and problem solver.

(2) These characteristics are sufficient to determine that a task environment is represented (in the information processing system) as a problem space, and that problem solving takes place in a problem space.

(3) The structure of the task environment determines the possible structures of the problem space.

(4) The structure of the problem space determines the possible programs that can be used for problem solving (1972, Ch. 14).

The Newell-Simon works (Newell, Shaw, and Simon, 1958, 1962; Simon and Newell, 1971; and Newell and Simon, 1972) are couched in difficult language, but a more basic consideration may limit the value of their theory for studying human problem solving, especially in children. The writer is persuaded by Mouly's (1968, p. 382) argument: programmed to locate conflicts in the various projected solutions for data it has received and to isolate solutions that are n o t in conflict with rules of operation, the computer performs essentially the same operations as the human brain, and therefore can provide some understanding of the process of adult human reasoning. But the computer is restricted

to the solution strategies for which it has been programmed, while the human problem solver is not. The human operator is not as thorough or as fast in checking through his solutions, but he has much greater flexibility in shifting from one strategy to another. Thus, even though the computer can simulate human problem solving, one cannot assume that the computer and its human counterpart progress toward solution in the same way. Compounding the uncertainty of the value of Newell, Shaw and Simon's (1962) model for the development of problem solving measures for elementary school children is the question of whether the heuristic they employ to approximate the thought processes of adults also resembles a child's approach to the solution of a complex problem.

The final model of problem solving included in this review of information-processing-type models is one offered by Ausubel and Robinson (1969). They postulated four levels in the problem solving process: problem setting; definition of the problem; gap filling; and verification. Central to this model is the gap filling process. To reduce the gap between the initial proposition(s) -- the given(s) -- and the final proposition -- the goal -- background propositions are manipulated by rules of inference, guided by a strategy.

Both types of models influenced our plans for designing measures of problem solving to use in the evaluation of USMES. The component processes models have logical appeal, and they yield to measurement whose validity might be assessed through content or construct validation

procedures. In fact, we developed a paper-and-pencil test of some of the component abilities involved in the solution of simulated, life-like problem situations. These development efforts are reported in Chapter IV. However, the question of interaction among the phases, or the operation of component abilities, was underscored in the models which viewed problem solving as a system of processing and feedback. This concern for learning about the structure of problem solving was the impetus behind the development of the PROFILES technique for studying how children go about solving complex problems. The development of PROFILES is reported in Chapter V.

B. Empirical Studies of Problem Solving

1. Reasons for the Lack of Progress

The decision to pursue new instrument development did not proceed directly from the review of basic theoretical frameworks offered by psychologists on problem solving. But our search of the empirical studies in the literature on problem solving revealed no attractive alternative. Gagne's comment summarized one kind of failing: "Only infrequently do psychologists include all these phases (he listed five) in their studies of problem solving, and most often only a single phase is studied as a focus of interest" (Gagne, 1959, p. 148). The other serious shortcoming with most of the empirical studies we reviewed for the USMES evaluation was the very limited, highly artifical nature of the tasks used to measure problem solving.

Collectively, the empirical studies of problem solving behavior offered some important information on issues relevant to the investigation of children's problem solving development in USMES, but these studies did not support one particular model of problem solving over others. Sustained research efforts on this "very old psychological topic" are difficult and costly. Most researchers nave had to narrow their empirical investigations to a limited aspect of the problem solving process. These efforts can be productive, but progress is slow.

The tools of the trade constitute a more limiting factor. Existing measures of problem solving are critiqued in greater detail in the next chapter, but a preview can be summarized succinctly -- the tasks are virtually limited to artificial intellectual puzzles, game boards, mazes, mathematical word problems, etc., and/or the reliabilities and validities of the measures are typically low or unstudied.

Elemental tasks themselves are not the limiting factor. With none but the simplest of props, Piaget (1954) applied his highly developed method of systematic observation of children to construct a theory of child development which has had profound influence on psychologists and educators, but his research efforts were disciplined and sustained. Most researchers on problem solving have had to conduct their research unsystematically, and they further limited the scope of their research by choosing measures of problem solving which at best can only tap limited aspects of this complex process.

A limited number of studies comprise the significant literature in the factor analytic studies of problem solving (Merrifield, Guilford, Christensen, and Frick, 1962; Werdelin, 1966; Bunderson, 1967; and Guilford and Hoepfner, 1971). Two studies using regression analyses (Harootunian and Tate, 1960; and Stevenson, Hale, Klein, and Miller, 1968) complement the studies employing a correlational approach to the investigation of human abilities thought to be important to problem solving. While some verbal reasoning and memory abilities appear consistently across these studies, this writer agrees with Speedie and his associates who concluded:

> "There is considerable difficulty in drawing any conclusions from these studies with respect to human abilities important to problem solving. The primary reason for this is the lack of similarity among the problem solving criteria used in the different studies" (Speedie, Treffinger, and Feldhusen, 1973, p. 23).

The later group called for representative measures of human problem solving which are operational definitions of comprehensive models of this complex process, but their well-informed efforts at new instrument development yielded real-life relevant problem tasks with unfortunately low reliabilities (Speedie, et al., 1973).

Empirical studies in the literature on problem solving which did offer important information for the evaluators' work on USMES dealt with developmental stages in problem solving among children and with factors influencing the performance of children on problem solving tasks.

2. Evidence for Developmental Stages in Problem Solving

The research showing the problem solving behavior develops through many different stages in which the child is able to master some skills but not others is persuasive, and it certainly has implications for both the development and the evaluation of USMES, a curriculum designed for elementary school students. The theoretical approach to the growth of logical thinking as described by Inhelder and Piaget (1958) appears to meet some of the rigorous criteria suggested by Newell, Simon, and Shaw (1958, p. 151) and reiterated by Simon and Newell (1971, p. 145) for a theory of human problem solving. In particular, the Piagetian developmental theory seems able to "predict the performance of a problem solver handling specified tasks," and to "explain how specific and general problem solving skills are learned, and what it is that the problem solver 'has' when he has learned them."

Piaget (Flavell, 1963) described various stages in the child's development in which the child's overt behavior shows distinctly different modes of thought from those of adults. While the reasoning of children may be subject to the same conditions and limitations as that of adults (Mouly, 1968, p. 385), there is substantial evidence that Piagetian developmental level significantly predicted problem solving performance. Saarni (1973) found that children classified as formal operational (or transitional) were generally more competent problem solvers than those who were classified as concrete operational.

She noted that

> "if the problem solver is limited to considering the concrete empirical situation at hand, he will be less able to hypothesize solutions which satisfy the constraints of the problem and transcend the empirical given.... On the other hand, the formal operational individual can consider problems involving several variables and their interaction; he can entertain hypotheses and deduce inferences from them and systematically evaluate alternatives. The continual decline in egocentrism that accompanies cognitive development also allows the problem solver to adopt different perspectives on the problem, thus making for still further flexibility and decentering in the strategies employed to solve the problem" (Saarni, 1973, pp. 342-343).

The use of limited, artificial problem tasks -- anagrams -- which may not be appropriate for the evaluation of USMES does not invalidate the evidence offered by Beilin (1967) that problem solving ability increases steadily with age. The work of Stevenson, Hale, Klein and Miller (1968) also supports this observation of developmental differences, and Neimark and Lewis (1968) report evidence compatible with an interpretation of the development of cognitive structures as a progression through a hierarchy of relatively discrete stages.

Some investigators considered the rival hypothesis that developmental differences in problem solving behaviors are observed because the measures of problem solving used are couched in the verbal medium, and that linguistic capabilities can mask the subject's ability to deal with the problem. For example, Weir (1964) in his discussion of age differences in problem solving observed that younger

children do not use the same strategies as older children, but he noted that this may be due to age-related differences in the ability to handle language. Eimas (1969) hailed the recent availability of several techniques that permit the investigation of problem solving behavior in children without undue reliance on interpretation of the child's verbal responses. (Again, the techniques were limited, artificial tasks -- a series of matrices -- which were not appropriate for evaluating problem solving in USMES but which were helpful for studying developmental trends in one aspect of problem solving.) The results of the Eimas study indicated that the efficiency of problem solution was strongly influenced by developmental level. Notably, responses at the first level were characterized primarily by a guessing strategy. The research of Neimark and Lewis (1968) is consistent with this finding.

Can strategies for the solution of complex problems, real or contrived, be taught to children as early as the second or third grade? The research literature suggests not, and our data based on interviews with a nationally drawn, random sample of 120 USMES children support this position. (See Shann, et al., Ch. IV and Ch. VI.) While USMES seemed to be teaching some of the components of problem solving to most of these children, "very often, the challenge was not perceived as a problem by the children, who simply saw what they did as a series of unrelated activities" (Shann, et al., 1975, p. 82). This was especially true of the younger children in our samples.

Clearly this evidence has serious implications for USMES curriculum development, but it is also highly salient to the development of new measures of complex problem solving. The new measures must allow for the study of developmental trends in the problem solving behaviors of elementary school children. The PROFILES technique especially was designed with this consideration in mind.

3. Situational Factors Influencing the Performance of Children on Problem Solving Tasks

Other empirical studies which were relevant to the evaluation team's work on the development of new measures of problem solving dealt with factors influencing how children perform on problem tasks. The factor of language has already been addressed above. Two situational factors are discussed in this section: (1) investigating group versus individual efforts at problem solution; and (2) testing with meaningful, relevant tasks rather than with artificial, contrived measures.

USMES purportedly involves group efforts toward the solution of real, complex problems taken from the children's school/community environments. Upon occasion, the children may work individually on subtasks of their USMES units, but they are encouraged to work cooperatively, to interact freely with other children, and to use their peers as resources, not just the teacher.

This picture of USMES suggests that the research on the relative effectiveness of problem solving by groups and by individuals may have

relevance for the design of criterion measures of problem solving to use in the evaluation of USMES or other problem solving curricula. In general, researchers comparing group and individual efforts have found that groups furnish more correct solutions to problems than comparable subjects do working as individuals (Gurnee, 1937; Klugman, 1944; Perlmutter and de Montmollin, 1952; Taylor and Faust, 1952; Gurnee, 1962). The survey of studies contrasting the quality of group performance and individual performance, conducted by Lorge, Fox, Davitz, and Brenner (1958) for the period 1920-1957 supports this generalization, but these researchers offered a noteworthy caution:

> "A common and dangerous practice is to generalize the principles valid for ad hoc groups to traditional groups. The ad hoc group is treated as a microscopic model of the traditioned group. This might be true but has not been experimentally validated. It is equally possible that ad hoc and traditioned groups behave in accordance with their individual principles" (1958, p. 370).

Writing for practicing and prospective teachers, Mouly (1968) argued in a similar vein:

> "A group approach is not superior simply because it involves a group; we need to clarify just what we can expect to be achieved through group methods and what is unlikely. Research, most of it done in a nonschool setting and involving relatively small groups working on an ad hoc basis, suggests that group work is most effective in dealing with complex tasks requiring a background more extensive than any one

> individual is likely to possess. Its effectiveness in problem solving depends on a number of factors including the nature of the problem, the resources, background and involvement of the members, the quality of the leadership, and various other considerations peculiar to the situation" (1968, pp. 392-393).

Studying the effects of group experience on individuals' abilities to solve arithmetic problems, Hudgins (1960) concluded that group members solved significantly more problems than subjects who worked alone, but that group experience did not enhance individual problem solving. Additional significance of Hudgins' study lay in his use of "natural groups." While his groups of fifth grade subjects were ad hoc in the sense that they were organized for purposes of that investigation, the groups operated for three consecutive days and approximated traditioned groups as they exist in the classroom.

The measures of problem solving used in the 1974-75 evaluation of USMES (Shann, et al., December, 1975) were tests of small group efforts toward the solution of real-life-relevant, complex tasks -- the Picnic Problem and the Playground Problem. However, good reliabilities for these tests were difficult to achieve and many problems in the administration and scoring of these complicated tests were not resolved. More is said on the issue of measuring group versus individual performance in problem solving in Chapter III.

The second situational factor influencing the measurement of children's performance on problem solving tasks, about which helpful information could be found in the literature, was the factor of

children's interest in the task. Cronback (1955) urged that problem solving tasks should be meaningful to the child, and Keisler (1969) included task relevancy in his list of criteria for problem solving tests. Research by Simon (1970) on encoding effects on complex problem solving indicated that students had more success in solving problems which were placed in various realistic contexts than they did with similar problems phrased in more abstract terms. The importance of developing problem tasks of interest to children was heightened by the USMES developers' claim that theirs is the only problem solving curriculum in which children work on truly real problems. (Evidence to the contrary was obtained through interviews with 40 USMES teachers, most of whom noted that they introduced their USMES unit challenges to their children in a contrived fashion.)

Such statements about USMES caused us to look briefly at other curricula, particularly science curricula, which number problem solving among their objectives. The interdisciplinary, process, "hands-on," discovery, pupil-centered approach to teaching espoused by USMES suggested a review of some writings from the field of education which these labels called to mind.

Problem Solving in Education

Dogged insistence upon rote memorization of factual information from textbooks and lists characterized the approach to instruction which dominated American schools in the first quarter of this century. The benefits of such disciplined study were claimed to be improved

faculties for clear thinking; the transfer was assumed to be direct and automatic. But as Hudgins (1966) observed, "the curriculum and goals and methods of instruction changed drastically in the aftermath of World War I." He noted two central generalizations which were of consequence:

> "One was a reaction against the long entrenched faculty psychology with its implications of 'automatic transfer' of learning; the other was advancing awareness among educators that our society was in the midst of an age of change. Today's knowledge might be obsolete before it appeared in tomorrow's textbooks" (Hudgins, 1966, pp. iii-iv).

The knowledge explosion affected science education most dramatically. Traditional science courses put forth a "rhetoric of conclusions" (Schwab, 1962, p. 24), but the rapidly increasing accumulation of factual scientific knowledge had brought on an "information crisis" which traditional science courses could not accomodate. Often times void of theoretical bases, the traditional courses were criticized further because they tend to be outdated, organized in a patchwork manner, too massive, and too technical;they did not involve the student in the activities of science (Marshall and Barkman, 1966).

Schwab called for the development of "The Enquiring Curriculum" to replace the traditional approach to teaching science (Schwab, 1962). In the past, he noted, rote learning of facts was thought to be sufficient for the education of the masses, while the mastery of learning was reserved for the elite. The problems facing the world today, he argued, should not be left in the hands of a select minority of educated

leaders because these problems "cannot be solved within the bounds of existing doctrines" (Schwab, 1962, p. 9). Schwab stressed that the general public must "become cognizant of science as a product of fluid enquiry" (Schwab, 1962, p. 3). In order to contribute to the solutions of the problems of an advanced technological society, to make informed political decisions, and to accept the lack of finality in advancements of all branches of science, each man must be taught to be an effective problem solver. In proposing the "Enquiring Classroom," in which the student's task is to analyze and challenge, and the teacher's role is to "teach the student how to learn," Schwab represented the concerns of many in the field of science curriculum at the time (Brandwein, Watson, and Blackwood, 1958; National Society for the Study of Education, 1960; Schwab and Brandwein, 1962; National Science Teachers Association, 1963, 1964).

The post-Sputnik science curriculum revisions initiated in the late 1950's were consistent in their emphasis on the processes of science. The new science curricula usually offered the student some opportunity to experience, first hand, the scientific method in operation. This trend is revealed in Klopfer's summary of basic objectives for science curricula (Klopfer, 1971, p. 567). Among the general objectives, he lists behaviors reminiscent of the components included by several philosophers and psychologists in their models of problem solving. Klopfer's objectives follow:

Processes of Scientific Inquiry I: Observing and Measuring

1. Observation of objects and phenomena
2. Description of observations using appropriate language
3. Measurement of objects and changes
4. Selection of appropriate measuring instruments
5. Estimation of measurements and recognition of limits in accuracy

Processes of Scientific Inquiry II: Seeing a Problem and Seeking Ways to Solve It

1. Recognition of a problem
2. Formulation of a working hypothesis
3. Selection of suitable tests of a hypothesis
4. Design of appropriate procedures for performing experiments

Processes of Scientific Inquiry III: Interpreting Data and Formulating Generalizations

1. Processing of experimental data
2. Presentation of data in the form of functional relationships
3. Interpretation of experimental data and observations
4. Extrapolation and interpretation
5. Evaluation of a hypothesis under test in the light of data obtained
6. Formulation of generalizations warranted by relationships found (Klopfer, 1971, pp. 562-563).

Noting this emphasis on the nature and structure of science and on the processes of scientific inquiry, Goodlad observed:

> "...striking similarity in the aims and objectives of nearly all [new curriculum] projects. Objectives, as they are defined in various descriptive documents, stress the importance of understanding the structure of the discipline, the purposes and methods of the field, and the part that creative men and women played in developing the field. One of the major aims is that the students get to explore, invent, discover, as well as sense some of the feelings and satisfactions of research scholars, and develop some of the tools of inquiry appropriate to the field"(Goodlad, 1964, p. 54).

In this regard, the USMES curriculum was no exception.

The goal of "scientific literacy" for an "informed citizenry" was consistent with the view that the schools of a democratic society should place maximum emphasis on reasoning. Proponents of this view advocated the "problem-solving," "problem-centered," or "reflective" approach to teaching (Hullfish and Smith, 1961; Bruner, 1961; Bayles, 1960). Variously named, the approach refers to a variety of classroom procedures which center on problem situations. Mouly characterized it this way:

> "Problem-centered teaching -- or more correctly, learning -- is a group activity; a problem of interest to the group, collectively and individually, is selected and clarified through discussion. Inherent in the method is the emphasis on discovery: rather than being told the solution, the students discover it. The teacher's task is to act as general consultant and coordinator, keeping the group on the track and on the move" (Mouly, 1968, p. 391).

Mouly's description of the problem solving approach and the effects of its application is quoted at greater length below because Mouly's words bear arresting resemblance both to the USMES developer's description of USMES (The USMES Guide, May, 1974), and to the evaluation results for USMES (Shann, August, 1975; and Shann, et al., December, 1975), even though Mouly's text predates these other documents.

> "This approach is also known as the pupil-centered or even the progressive method.... Where appropriate, it tends to result in increased insight into individual problems as well as increased ability to engage in problem-solving behavior. Ideally, it results in greater interest and motivation; in more penetrating, although perhaps less

> extensive, education; in more meaningful learning; and in greater understanding. Furthermore, it is relatively effective in changing attitudes and behavior [Lewin, 1958]. It also promotes a feeling of group belonging and provides training in democratic resolution of problem situations. On the other hand, it is a rather slow-moving procedure which is sometimes difficult to adapt to a systematic coverage of the curriculum, and it may have to be supplemented by other more systematic approaches if gaps are to be avoided [Hermanovicz, 1961]. It is also difficult to handle well...." (Mouly, 1968, p. 391).

Among the proponents of the problem solving approach, Bruner has been the most visible, and his views on discovery learning are particularly evident in USMES. In their excellent synthesis and critique of research on teaching in the sciences, Shulman and Tamir (1973) cite the Woods Hole conference of experts on the teaching of science, chaired by Jerome Bruner and convened under the auspices of the National Academy of Sciences, as an important milestone in the revolutionary changes and developments undergone by the field of science education in the 1960's. Out of this conference emerged a book, The Process of Education (Bruner, 1960, 90 pp), which Silberman (1970) hailed as the most influential piece of writing to come forth from the curriculum reform movement. Concurring in that judgment, Shulman and Tamir (1973, p. 1098) said the book "provided an unmistakable sign to the rest of the educational community that radical changes in the teaching of science were imminent."

Out of concern for "a well educated citizenry" with the potential for resolving the crises of our times (Bruner, 1960, p. 1), Bruner

advanced recommendations which lend additional support to the notion that it is both appropriate and expedient to teach the process of problem solving in elementary school. Explicating the long-range benefits of "giving students an understanding of the fundamental structure" underlying all academic subject matter (1960, p. 11), Bruner emphasized the necessity for curricula based on "discovery" experiences which lead to the grasp "of regularities of previously unrecognized relations and similarities between ideas" (1960, p. 20). The discovery of structure, according to this approach, is both cognitive and affective -- a notion reiterated by USMES. As Bruner puts it:

> "Mastery of the fundamental ideas of a field involves not only the grasping of general principals, but also the development of an attitude toward learning and inquiry, toward guessing and hunches, toward the possibility of solving problems on one's own"(Bruner, p. 20).

The second major conception introduced by Bruner in The Process of Education was the provocative proposition that "the foundations of any subject can be taught effectively in some intellectually honest form to any child at any stage of development" (1960, p. 33). The challenge thus posed to educators is to develop "intellectually honest forms" of presentation. Bruner refers to his own research and to that of Piaget and Inhelder as foundations upon which meaningful curricula may be built.

For Bruner, the value of the early introduction of discovery experiences to acquire structure lies in the enduring, long-range

benefits to the child. Given the structure of significant subjects, the child can more readily incorporate details as they are presented to him and can transfer new learnings to other situations. Further, according to Bruner, the emphasis on the acquisition of the structure of significant disciplines provides educators with a criterion for streamlining the curriculum. They must ask, says Bruner, of "any subject taught in primary school, whether, when fully developed, is it worth an adult's knowing?" (Bruner, 1960, p. 52).

The USMES response to Bruner's exhortations is characterized in the following quotes from the project's newsletter:

> "USMES is a program which involves elementary school students in investigation and action on real problems from their school and community environment....
>
> USMES is a philosophy which holds that children can themselves design and carry out the investigations and activities needed to solve a problem....
>
> The levels at which the children approach the problems, the investigations that they carry out, and the solutions that they devise vary according to the age and ability of the children. However, real problem solving involves them, at some level, in all aspects of the problem solving process....
>
> The ultimate goal of USMES -- the project, the program, and the philosophy -- is to prepare young people to deal successfully with the challenges of life, to be intelligent voters and consumers, to care about the world they live in, to believe they can make a difference....
> (Education Development Center, USMES News, October, 1975, p. 2)

In characterizing the trends evident among the wave of new science curricula developed during the sixties, Klopfer (1971, p. 565) noted that instead of having material arranged according to the subject areas of a discipline, the new courses were organized around processes of scientific inquiry, particularly at the lower grade levels. Also, the roles of mathematics are stressed in the new science curricula. Both of these generalizations offered by Klopfer are true of the USMES curriculum.

Klopfer (1971, p. 565) also noted that new science curricula at the high school level tended to be organized around unifying conceptual ideas. Interestingly, this distinction between the organizational frameworks for secondary versus elementary science curricula coincides fairly well with Schwab's distinction between two closely related and interdependent aspects of the structure of a discipline -- the substantive and the syntatic. The substantive structure of a discipline consists of "a body of concepts -- commitments about the nature of a subject matter functioning as a guide to inquiry" (1962, p. 203), while the syntactic structure of a discipline involves "the pattern of its procedures, its method, how it goes about using its conceptions to attain its goals" (1962, p. 203). Shulman and Tamir noted that it was structure in the substantive sense which is most similar to what Bruner had in mind, though Bruner did not explicate the concept as clearly as Schwab. However, it is structure in the syntatic sense which is emphasized by USMES.

Problem Solving in USMES

Relaxing the rigorous criteria applied by philosophers who studied the work of adult scientists using the scientific method, this writer concluded that the structure of problem solving implicit in USMES is the syntatic concept of structure -- not the conceptual schemes of science, but its inquiry processes. If this view of the underlying structure of USMES were acceptable to its developers, then the determination of "proof of concept" for USMES might have been a simplified matter. Tests were available which purportedly measured students' ability to apply the scientific method, among them the AAAS (1966) tests listed in Appendix D for the SAPA curriculum, Science -- A Process Approach.

Confronted with this view, the USMES developers claimed that their curriculum could achieve that and much more. USMES distinguished itself from other problem solving, science-process-oriented curricula by its emphasis on real complex problems of the students own choosing -- problems which are truly meaningful to the students. Pressed to enumerate the criteria of "realness" which characterize problems appropriately labeled "USMES-type" problems, the developers offered the following criteria which subsequently appeared in the USMES News:

> "The problems are 'real' in that they
> (1) have immediate, practical effects on students' lives;
> (2) can lead to some improvement by students;
> (3) have neither known 'right' solutions nor clear boundaries;

(4) require students to use their own ideas for solving the problems, and
(5) are 'big' enough to require many phases of class activity for any effective solution (EDC, October, 1975, p. 2).

This perspective for USMES, its developers argued, rendered any contrived problem inappropriate as a measure of proof of concept for USMES. On these grounds, they faulted not only the puzzle-insight tests, mathematical word problems, and multiple-choice tests which portray a limited view of problem solving, but also the Picnic and Playground Problems -- real-life-relevant tests of problem solving processes specially designed for the evaluation of USMES (Shann, et al., December, 1975, Ch. VI). (These tests will be reviewed briefly in the next chapter.)

Since the evaluators attached great importance to matching the focus of our measurement tools with the goals expressed by the USMES developers for their curriculum, we urged the developers to clarify in terms of observable student behaviors, their objectives for USMES. We suggested that a "components approach" to this task might produce the most fruitful set of guidelines for our work on new instrument development in problem solving. The USMES senior staff members responded with the list of components of problem solving/decision making shown in Appendix A. They also offered the set of affective goals shown in Appendix B to complement the cognitive goals. (This list of affective goals was used as a screening device for the selection of items used in the pre-post survey of attitudes of USMES

and control students during the 1974-75 USMES evaluation program. Results of that attitude assessment are given by Shann et al., December, 1975, Chapter VII.)

With great respect for the ability, sincerity, and commitment their efforts represent, we urged the USMES developers to rethink their list of cognitive objectives for USMES with the view toward compiling a simplified, more limited set of goals. The evaluation team members concurred that the objectives listed in Appendix A were much too ambitious, indeed unrealistic, to pose as objectives achievable by the majority of elementary school children, especially those at the primary and intermediate grade levels. While they might be achievable by groups of older children working together on USMES-type problems, these skills were unrealistic to expect of individual students, especially the younger children, even after two or three years of intensive USMES experience. We based this judgement on intuition and on interviews with USMES children, but also on the collective experiences and educational backgrounds of the team members, especially those of the Project Director and her associate. With undergraduate background in science and in mathematics, we had taught in secondary school and in elementary school respectively. Subsequently, we earned doctorates in educational psychology and research. We simply could not believe that USMES, or any other curriculum, could develop the skills at the level listed in Appendix A in individual elementary school students. Yet, the sponsors of USMES sought evidence of individuals' success in mastering the abilities to solve

real, complex problems; that is, NSF wanted "proof of concept," for USMES as the curriculum affects individual's performance in problem solving, not just group performance.

Other factors prompted the evaluators to seek a more limited set of goals for problem solving from the USMES developers. Some of the components listed in Appendix A could be measured quite readily, even by paper-and-pencil tests, for example:

> Distinguishing facts from opinions;
>
> Identifying unsupported assumptions or generalizations.

However, others of the components, strictly speaking, were not stated as observable behaviors, for example:

> Considering the practicality of suggested solutions;
>
> Considering that a problem may have different solutions depending on the values applied;
>
> Deciding on generalizations that might hold true under similar conditions.

Still other components, though stated as observable behaviors, would be very difficult, if not impossible, to measure in the contexts of real problems, for example:

> Trying out various suggestions and evaluating the results;
>
> Applying the process learned to other real problems;
>
> Making suitable simple mathematical models of real situations and refining them.

The USMES developers responded to our critique with diplomacy and dispatch. They offered the revised list of goals for individual children in problem solving shown in Appendix C. This list served as

the basis for our efforts toward the development of new measures of problem solving appropriate for use with elementary school children.

Following a critique of existing measures of problem solving and a discussion of the qualities desirable in tests of problem solving in Chapter III, two new measures of problem solving are advanced. One is a paper-and-pencil test of selected components of complex problem solving. The other is an observation/interview technique called PROFILES designed to identify the processes individual children engage in as they attempt to solve real, complex problems in group situations. These new measures are discussed in Chapter IV and Chapter V, respectively.

CHAPTER III

A REVIEW OF EXISTING MEASURES OF PROBLEM SOLVING

The purpose of this chapter is threefold: (1) to acknowledge criteria for ideal problem solving tests; (2) to critique existing tasks, pointing out their serious limitations for the measurement of USMES-styled, real, complex problem solving; and (3) to identify the challenge for new instrument development thus posed to the USMES evaluators.

Our review of available tests was aided greatly by three important works surveying the tasks for measurement of human problem solving: Ray (1955); Feldhusen, Houtz, Ringenbach, and Lash (1971); and Speedie, Treffinger, and Feldhusen (1973). Ray's research was supported by the United States Air Force, and the latter two works were supported by grants from the U.S. Department of Health, Education, and Welfare, Office of Education. The fact of support from these sources suggests that there is widespread, continued concern for advancement in the measurement of problem solving, and the articles themselves reflect the magnitude of the efforts engaged by these reviewers of problem solving tasks. In surveying existing tasks for the USMES evaluation, we did not attempt to duplicate their substantial efforts. Rather, we concurred with most of the

views expressed by these researchers, and we acknowledged their reports as the framework upon which we built additional review and critique.

Desirable Characteristics of Tests for Complex Problem Solving

Ray argued that "Progress in any field of research depends on many things in addition to the ideas produced by the individual research men, among them the availability of measuring instruments and of standard materials and techniques" (1955, p. 134). In order to keep his review of tasks for use in human problem solving within reasonable limits, he considered only the more "complex" tasks. Complexity referred to the amount of work required of the subject for problem solution: multiple responses were necessary; multiple hypotheses were possible; trial and error attempts at solution included more than one or two trials; and more time than that needed for a single brief act was required. Owing to their great number, Ray "arbitrarily" eliminated mathematical problems. He further discounted from citation in his review those problems i which the experimenter could see only the solution, the end product, and not the subject's solving processes at work. In a discouraging report, Ray commented that one

> "is struck with the fact that many of these tasks (included in his article) have been of the parlor puzzle or game variety rather than being tasks constructed for experimental use, which is even more true in the problems omitted than in those included....

> Few of the tasks described herein permit more than a two-category scoring system, success or failure, although some investigators have shown considerable ingenuity here. Reliability has been mentioned only twice and validity never by the writers of the reports" (Ray, 1955, p. 148).

Ray observed that the plea for more theory reiterated in articles and chapters summarizing the state of the field of problem solving had produced no great results. Offering "a more modest sounding plea," Ray suggested that

> "What this field needs are dimensionable independent variables, predictions of their effects on dimensionable dependent variables, and tasks especially designed to measure those effects. Once this is achieved, we can start to work on hypothetical constructs, perhaps using specific hypotheses as first order intervening variables and processes such as hypothesis formation by induction or deduction as second order variables" (Ray, 1955, p. 148).

Like Ray, John (1957) was particularly concerned that problem solving tests be developed which would allow direct observation of the problem solving process. Additionally, he proposed that the desired tests would impose a minimum of information, structure and external constraints in the directions given to the subject, and at the same time, such tests should be maximally free of special skills and content.

Cronbach (1955) stressed that the problems be meaningful for the student so that the testing situation does not degenerate into an exercise which makes no demands on the higher mental processes. Keisler (1969) agreed that problems must be realistic and interesting

for the student, so that he will accept it as something he would like to solve. Otherwise, interference to effective problem solving occurs; the student feels no desire to solve the problem for its own sake. His extrinsic motivation is to find an answer acceptable to the teacher or the examiner, or to avoid unfavorable consequences. Interested in the development of tests for classroom use, Keisler offered several additional criteria for the tests of problem solving.

The criteria for problem solving tests proffered by the above authors and other writers were summerized by Speedie, Treffinger, and Feldhusen (1973) in the list which follows:

1. Tasks selected for problem-solving tests should be complex; i.e., they should not be merely simple exercises, but rather problems in which there are a large number of steps from an initial state to a final state, or a reasonably large number of attributes.

2. Performance on the test should be minimally related to previous learning which could differentiate individuals at the time of the test.

3. The problems should command the attention and interest of the subject so as to insure an adequate level of motivation for optimum performance.

4. The test should yield a variety of continuous measures concerning the outcomes of problem solving, the processes, and the intellectual skills involved.

5. The test should contain a minimum number of constraints on the types of problem-solving behavior the individual may engage in.

6. The test should demonstrate both reliability and validity.

7. The test should be practical for group administration.
(Speedie, Treffinger, and Feldhusen, 1973, pp. 26-27)

Keisler's (1969) version of the seventh criterion listed above was that a test of problem solving should be amenable to group presentation "if at all possible." Having confronted the challenge of developing tests of component intellectual skills in problem solving and of developing techniques to examine the processes individuals engage in as they pursue the solution of a problem, the evaluation team noted that it seemed possible to achieve the former kind of assessment in a group administered test, but we suggest that it is virtually impossible to study the complex processes in group tests. We subscribed to the other criteria summarized by Speedie and his associates as the criteria we would attempt to meet in continued new instrument development efforts.

Limitations to Existing Measures

In their extensive search of problem solving measures, Feldhusen, Houtz, Ringenbach, and Lash (1971) found that four categories wer useful for classifying the great variety of tasks reported in the literature: (1) "puzzle-insight" problems, (2) "process" problems, (3) "component tasks," and (4) "real-life-relevant" tasks. Speedie, Treffinger, and Feldhusen (1973) critiqued a number of tests from each of these four classes, but they reserved serious doubts that

any of these tasks met their criteria for an ideal test of problem solving.

A. Puzzle-insight Problems

The puzzle-insight problems include many of the tools discussed by Ray (1955) which were used in most of the classical investigations of problem solving. Feldhusen, et al. (1971) listed the most famous of these: Maier's (1945) "two string" and "hatrack" problems; Duncker's (1945) "box problem;" Luchins' (1942) water jar problems; Katona's (1940) matchsticks problems, and anagrams (Johnson, 1966). But as Speedie, et al. (1973) noted, the puzzle-insight problems fail most of criteria for an ideal problem-solving test. Basically artifical intellectual games in which the initial conditions and final goals are stated precisely, and only a severely limited number of routes to solution are possible, these puzzle-insight problems are not sufficiently complex and they impose numerous constraints upon the strategies usable for solution or the solution itself. The puzzle-insight problems are highly affected by previous learning in that they tend to have a relatively simple if unusual solution which is easily remembered once it has been encountered. Most are strictly end-product measures scored by time to solution, number of mistakes, number of unacceptable answers, or number of hints necessary for solution. None of these provide much information about the processes engaged in, or even about the component intellectual skills involved. With all of these shortcomings, the

puzzle-insight problems generally require individual administration too. These tasks were quickly dismissed as having little or no validity for evaluating the success of USMES in meeting its developers' objectives for complex problem solving.

B. Process Problems

The second class of problems, the process problems, included tasks whose structure enables the recording of all of a subject's decisions throughout the course of the problem. These responses can be studied, and processes and strategies can be inferred. Among the process problems critiqued by Speedie, et al. (1973) were: switchlight problems (John, 1957; Tyler, 1958; Davis, 1966; Davis, Manske, and Train, 1963); the verbal maze problems of Hayes (1965); simulated problem situations (Glaser, Damrin, and Gardner, 1954; Rimoldi, 1960; McGuire and Babbott, 1967; Streufert, Kliger, Castore, and Driver, 1967; Nattress, 1970); and concept identification problems (Bruner, Goodnow and Austin, 1956; Clark, 1971; Bourne, Ekstrand, and Dominowski, 1971). We would agree with Speedie, et al. who concluded that "those measures in the process category are best qualified as ideal problem solving tests. That is, they should yield the maximum amount of information about the problem-solving process and reflect the utilization of a sizable number of human abilities and skills" (1973, p. 35). Yet we feel that each of the process problems listed above is beset with one or more serious deficiencies: limited complexity; the artificial, uninteresting

nature of the tasks; or especially unsatisfactory reliability and validity evidence.

Even Speedie's (1973) laudable efforts failed the criterion of adequate reliability. With Treffinger and Feldhusen, Speedie developed multiple forms for three group administration tasks -- simulated problem situations, verbal mazes, and concept identification tasks -- to measure problem solving processes. However, test-retest reliabilities were found to be quite low, and alternate forms reliabilities were essentially zero for most of the problems.

C. Component Problems

The third group of tasks Feldhusen, et al. (1971) labeled "component problems." The primary concern reflected in this type of test was the measurement of component skills and abilities involved in problem solving. Included in this category were the battery of tests which Guilford (1967) used to establish his Structure-of-Intellect model of intelligence, the Torrance Tests of Creative Thinking (1966), and Unfinished Stories (Lundsteen and Michael, 1966). Ramirez's (1971) model classroom tasks could also be considered a components problem.

Speedie, et al. criticized the Guilford battery as measures of complex problem solving because "In none of these tests was there a provision for measuring the strategies of problem solution or provision for multiple solutions, asking relevant questions about the problem, clarifying the goal. or defining the problem" (1973, p. 19). Thus the Guilford tests did not include measures

for many possible facets of efficient problem solving.

Charged with the development of a test of creative problem solving, Miles (1968) had faulted the Guilford battery for the same kinds of reasons. He felt that he could not pursue research on creative behavior with either the Guilford battery or the Torrence Tests because "these tests lack external validity, i.e., they are frequently unlike any common problem solving situation due to their brevity, testlike characteristics, artificiality, and lack of occupational or subject matter specificity" (Miles, 1968, p. 7). Certainly similar criticisms about these tests would also be voiced by the USMES developers.

The newest "components" styled test which held promise as a measure which might have been used in the USMES evaluation was the Purdue Elementary Problem Solving Inventory, the PEPSI (Feldhusen, Houtz, and Ringenbach, 1972). The twelve components synthesized from the problem-solving literature as the basis for the PEPSI resembled closely the nature and sequence of components of problem solving enumerated by the USMES developers. And, since the PEPSI was designed to measure the problem solving abilities of socioeconomically disadvantaged children in grades one to six, its developers placed a premium on constructing a test with real-life tasks which would appeal to children's interests. However, upon inspecting the test materials themselves, the evaluation team and our advisory board concurred that the PEPSI was not

appropriate for the USMES evaluation. The complexity of tasks was limited. Feldhusen, et al. (1972) observed that the test "appears too easy for sixth-graders" (p. 899), yet many USMES students are sixth through eighth-graders emanating from middle and upper middle socioeconomic backgrounds. The ceiling effect of the PEPSI test would have been most serious with these USMES students. Other reservations precluding our use of the PEPSI are noted in the next chapter of this volume.

Indeed, Speedie, Treffinger and Feldhusen seemed to have discounted any component measure as an ideal test of problem solving when they said that

> "The component problems, while they appear to qualify by most of the criteria are not usable due to the fact that they are based on pre-existing theories concerning the skills involved in problem solving behavior" (1973, p. 35).

The evaluation team did not agree with this assessment of component problems. Perhaps they could not achieve "ideal" status, but, designed with care, context-dependent (not content-dependent) component problems may yield data from which we could infer about processes. Construct validation efforts following the suggestions of Cronbach and Meehl (1955) would be necessary to support the theoretical base of such a test and to validate the test itself.

D. Real-life-relevant Problems

The final category of problem solving tests identified by Feldhusen, et al. (1971) consists of those tests which emphasize

"real-life" problem situations constructed to highten students interest in the problems. Speedie, et al. (1973) illustrated several tests of this type: Crutchfield and Covington's (1965) The Old Black House, The Man in the Pit, and The Missing Jewel, tests used by Saarni (1973) to study the correlation between Piagetian developmental level and children's problem-solving performance; Miles (1968) Creative Design Test, and Treffinger's (1970) Fighting on the Playground.

Speedie, et al.(1973, p. 30) noted that these "real-life-relevant" problems are similar to the component type in the scores they yield, but the "real-life-relevant" tasks place greater emphasis on motivating subjects to perform as well as possible by solving relevant problems. Speedie's observation is helpful because it is at this point that the scheme offered by Feldhusen, et al. (1971) for classifying problem tasks is difficult to use. "Real-life-relevant" tasks tend to be component problems, but th could be process problems with the designation "simulated problems," such as the simulation exercises in medicine developed by McGuire and Babbott (1967). Miles (1968) Creative Design Test is scored for measures of feasibility, fluency, flexibility, and originality, like the Torrance Tests of Creative Thinking, the TTCT. Yet, Miles' test is designated a "real-life-relevant"test, while the TTCT is a "components" test.

Depending on its underlying design then, a "real-life-relevant" test may be criticized for the limitations attributed to process problems or to component problems. Furthermore, the USMES developers would argue, and we would concur, that even the "real-life-relevant" problems listed above are artificial, sometimes gimmicky, and somewhat limited as measures of complex problem solving.

The evaluators could locate no more recent tests than those herein cited either in published journal articles or in the Annual Programs for AERA Conventions through 1976. A computer search of ERIC documents also failed to produce a new test which might measure complex problem solving.

Problem Solving Measures Used Previously in the Evaluation of USMES

Our review of the theoretical bases of problem solving and our critique of existing tests were on-going efforts of this USMES Evaluation Project Director and her staff. Yet the demands for a variety of evaluative feedback upon which to base curricular revisions had prevented the Project Director from offering a written account of these efforts until now. Additionally, we had to respond to the immediate need shared by the National Science Foundation as the sponsor of USMES and by the program developers for proof of concept of USMES. We had to apply the best available techniques in the evaluation programs, while acknowledging their limitations. To complement this chapter's review of existing problem solving

assessment techniques, those measures used in previous USMES evaluation work are discussed in this section.

The "Notebook Problem" was the first test of problem solving used in the evaluation of USMES during 1971-72, when no independent evaluation program had been engaged. Working as a consultant to the USMES development staff, Dr. Bernard J. Shapiro conceived the "Notebook Problem," an intriguing, different, individually administered test. Shapiro supervised the scoring, analysis and reporting of the data from the Notebook Problem (Shapiro, May, 1973). Students were asked to select one of three or more notebooks for their class and give reasons for their choice. Scored for measurability and "level of warrant," the reasons offered by USMES students were significantly more objective and testable than those offered by control students. Yet in the view of test administrators, the task was limited and highly artificial and uninteresting to many students who took the test. Worse for our purposes, the person who scored the tests noted that in most cases, the children seemed to make a "snap" decision and then struggle to offer "reasons" for their impulsive choice.

When the present evaluation team assumed responsibility for the independent evaluation of USMES for the 1973-74 academic year, the Playground Problem was applied as a measure of problem solving. The following year, both the Playground Problem and its parallel form, the Picnic Problem, were used as "real-life-relevant" problem tasks to compare the performance of USMES and control classes.

The conceptual bases for these problems reflected John Dewey's (1910) conceptualization of the problem solving process, whose "five logically distinct steps" permeated much of the literature about USMES prepared by the USMES Central Staff. The Playground Problem required that the students develop a plan for a playground which would serve children in their school and/or neighborhood. A catalog of equipment, cost data, and measuring instruments were given to the students along with the information that they could spend up to $2000.

The pre-test, post-test control group design used in the evaluation necessitated that a parallel form for the Playground Problem be developed, since retest results from such a unique test would be affected by memory factors. To answer this need, the Picnic Problem was developed. This test challenged students to develop plans for a class picnic. The students were provided with a photograph of various foods available to them and a map drawn to scale which included the locations of their school and three park areas as possible sites for the picnic. Along with measuring instruments, the students were given cost data and the information that they could spend up to $25. They were to assume that 25 students would be going on the picnic, and that a school bus would be provided for their transportation, free of charge.

Both problems are accompanied by administrator's manuals for presentation of the tests to groups of five children. It should be

noted that skilled administration of these tests is difficult and yet critical for reliable results. (See Appendices E and G for these manuals.)

Scoring is also rigorous. The scoring protocols developed for the tests offer both cognitive and affective assessments. The cognitive scores provide indices of the students' abilities to identify, measure, calculate, and record data on factors which they think are salient to the solution of the problems. The behavioral assessments include ratings on motivation to accept the problem, commitment to task. efficiency of manpower, and the nature of group leadership. Additionally, the protocol for the Playground Problem afforded an assessment of the students' product: their drawing of the play area design. (See Appendices F and H for the scoring protocols.)

Neither the Playground Problem nor the Picnic Problem satisfied the developers' concern that these tests meet all the criteria for "realness." The tests were simulated problems whose solutions would not have immediate, practical effects on students' lives. Nevertheless, data shown in Chapter VI of the 1974-75 USMES evaluation report (Shann, et al., December, 1975) indicated that the vast majority of students tested with the Playground and Picnic tasks were motivated to accept the problems. In that sense, we can say the tasks were meaningful to the students.

Another of the developers' criteria for "realness" is that real challenges are "big" enough to require many phases of class activity for any effective solution. The Playground and Picnic Problems did not meet this criterion. In the interest of observing reasonably large samples of children we had to abbreviate test times to approximately one hour.

Despite these limitations, the Playground and Picnic Problems have other important features in common with USMES-styled, real problems: they have no "right" solutions; they have no clear boundaries; they require students to use their own ideas for solving the problems; and they elicit group efforts toward the solutions to the problems. These assets prompted our use of the Playground and Picnic Problems over other available measures in the 1974-75 USMES Evaluation Program. While content validation of the tests as simulated measures of life-like, complex problem solving was established, insufficient control over the administration of the tests limited the judgements we could make about the effectiveness of the Playground and Picnic Problems as measures of complex problem solving. Neither the Playground Problem nor the Picnic Problem satisfied the program developers' concerns that these tests meet all of their criteria for "realness." Therefore, rigorous investigation of these tests' reliability and statistical validity did not seem to be warranted.

The Challenge to the USMES Evaluators

The primary objective of the USMES curriculum is the enhancement of elementary school students' abilities in real, complex, problem solving. Accordingly, the primary responsibility of the USMES evaluation project was the investigation of whether USMES was achieving that goal. Yet this determination of "proof of concept" was difficult and challenging for the evaluators, because the "state of the art" of measuring the problem solving abilities and processes of children was itself so limited. We added to our evaluation efforts a second thrust -- new instrument development for the measurement of complex problem solving in elementary-school-aged children.

Our consideration of the theoretical positions on problem solving, our study of the limited empirical evidence for the theories, and our perceptions of the most promising lines of test development efforts directed us to consider not one but two approaches to the measurement of problem solving in children. The approaches should complement one another in answering different kinds of research questions. Both approaches would tap the individual child's performance in problem solving, but the paper-and-pencil components approach, called the Test of Problem Solving Skills, or TOPSS, would measure the more limited component intellectual skills which each USMES child should acquire, while the process approach, called PROFILES, would be an observational/interview technique designed to

study the processes in which individual children engage as they attempt to solve complex problems.

The TOPSS approach grew out of our consideration of the component models of problem solving offered by Dewey (1910) and others as discussed in Chapter II. These models held conceptual appeal for us, and a paper-and-pencil test of components was a practical way to assess the problem solving abilities of large numbers of children. Quite reasonably, the USMES developers argued against a components approach because "it is inherently limited and because skill on the parts does not insure successful achievement on the whole." As Glaser, Damrin, and Gardner (1954) noted, elaborate performance measures usually possess the advantage of greater validity, but they are quite costly both in terms of equipment and man-hours of testing time, whereas the multiple-choice paper-and-pencil tests of proficiency generally achieve economy at the expense of validity. Nevertheless, we proceeded with the development of TOPSS because we felt that it was certainly worthwhile information to learn that USMES could achieve even the development of the component skills, especially in children as young as fourth graders.

The USMES developers urged us to develop a wholistic approach to measuring problem solving on actual and complex tasks, chosen by the students, whose solutions would affect those students' lives. If the evaluation of problem solving were limited to the paper-and-

pencil test of components, the USMES developers argued, then USMES might appear to be no more successful than other problem-solving-oriented curricula which use contrived and/or limited problems. The PROFILES approach was designed to meet the evaluators' concern for a valid assessment of USMES, one which they felt should examine the children's performance on real complex problems which have all the attributes of actual USMES challenges. The only course open to us which would meet these concerns was the periodic, trained scrutiny and skillful interview of individual children at work on their USMES units. This is the essence of PROFILES, a technique which would enable the study and reflection on two other important issues: (1) At what average age are children developmentally ready to internalize abstractions and formulate generalizations necessary to the process of complex problem solving (or how early can this be taught); and (2) What are the processes which children pursue as they attempt to solve real, complex problems? Thus the PROFILES approach also responds to the evidence for developmental stages in intellectual abilities and to the suggestions of theorists who claim that problem solving is best viewed as an information processing system.

The development of the paper-and-pencil test of components called TOPSS is detailed in Chapter IV, while the PROFILES technique of observation and interview is discussed in Chapter V.

CHAPTER IV

THE DEVELOPMENT OF A PAPER-AND-PENCIL TEST OF PROBLEM SOLVING SKILLS (TOPSS)

Introduction

The development of methods to assess problem solving skills in elementary school children was one of the major tasks of the 1974-75 USMES evaluation project. Initially, we hoped that we could locate an already existent standardized test which could be used to measure the problem solving abilities of large numbers of USMES and control children. Specifically, we were searching for a group-administered, paper-and-pencil test, suitable for pupils of the elementary school level. Such a test would also need to be based upon a conceptual view of problem solving which coincided in nature and spirit with the character of the USMES activities.

We were aware from the beginning of several limiting factors which would have to be considered in the selection of a standardized test suitable for this particular population. First, are elementary school children developmentally ready to form the concepts necessary for an understanding of the components of scientific problem solving? Developmental studies suggest that children pass through somewhat discrete stages of readiness which gradually lead to the ability to think abstractly and to use logical rules to guide actions. Further it has been observed that on the average, children may be ready to

use the rules of adult logic around the age of eleven (Inhelder, 1960). We felt that any testing of children's ability to grasp such abstracted components of problem solving as hypothesis generation and verification would have to reflect serious consideration of Piagetian developmental patterns.

Second, we were aware of the limitations inherent in any test which depends upon verbal skills for both input and output modalities. As children develop their ability to utilize rules and abstractions for planning and evaluating their actions, it is probable that they are able to sort according to a rule long before they are able to verbalize or explain the principle inherent in that rule. For example, children may be able to sort objects by category long before they are able to attach a verbal label to that category. However, in asking children to respond to a paper-and-pencil test, we would be requiring them to demonstrate both their understanding of abstract concepts plus their ability to verbalize these abstractions. A failure to respond correctly to a verbal item might indicate a lack of understanding of a concept, or it might indicate the inability to deal with the concept verbally, or both.

These first and second considerations led to a related problem -- that of wide variability in the abilities of students across a wide range of ages and grade levels. In view of normal development/maturational patterns, it is unlikely that a single test could be developed which would yield meaningful scores from pupils of the

second through the eighth grades. Questions which might differentiate high and low problem solving abilities among second and third graders would undoubtably be much too simple to reveal differences in seventh and eighth graders. Items which would be suited to the reading level of the older children would be too difficult for the younger grade levels to handle.

The most fruitful approach, it appeared, would be to narrow the age range of our population to grades four to six, for example, and to locate items geared to that level, both in terms of their degree of abstraction and in terms of their reading level. The utility of this approach would be enhanced by the fact that most USMES users are in grades four to six. A series of problem solving tests, graduated in difficulty level, which would yield comparable scores for students across the elementary school grades simply did not exist, nor was it feasible or realistic for the evaluation team to attempt its development.

Having carefully examined these factors, we began our search for a problem solving test with a realistic view of the limitations inherent in this method of assessment.

The Search for Appropriate Items

It had been suggested that the Purdue Elementary Problem Solving Inventory, the "PEPSI" (Feldhusen, et al., 1972), developed at Purdue University, might meet our assessment needs. The PEPSI uses line

drawings on a film strip to present selected aspects of problem solving. Students respond to tape recorded questions by choosing one of four possible answers. Other modes of presentation have also been developed for the same items. The test was used in conjunction with a problem solving curriculum designed for disadvantaged elementary school children, and therefore, its developers were specifically interested in minimizing the reading load.

Feldhusen conceptualized problem solving as a process involving "several different kinds of abilities," or components (Feldhusen et al., 1972, p. 24), a conceptualization which seemed quite compatible with that of USMES. However, when we reviewed the test items closely we concluded that many of them were either ambiguous or highly dependent upon skill in visual discrimination. Since no item analysis was reported in the literature of the test, we could not depend upon the items to discriminate properly, if at all. Our advisory board supported this critique and recommended that we continue to review the literature in an effort to locate a suitable standardized test. Since we had exhausted the literature on problem solving, we decided to consider science achievement tests as possible sources of appropriate items.

The term "problem solving" appears frequently in both mathematics and science education. As commonly used in reference to mathematical problems, especially "word problems," the concept is far too limited to be relevant to the scope of USMES activities and challenges. In science education, however, the objectives listed for teaching "the

scientific method" or the "process of science" correspond closely to the list of components of problem solving which were developed at the request of the USMES evaluation team by the USMES development staff. To illustrate this match, a set of general objectives from Klopfer (Klopfer, 1971, p. 562) is compared with USMES components in Figure 4.1.

In reviewing available science tests, we hoped to find items which were designed to test a child's understanding of the "process of science." In addition, we needed items which did not rely heavily on knowledge of science "facts," but which were drawn instead from USMES-like, "real-life" experiences. The majority of science tests we reviewed did not deal specifically with the "processes of science;" those which did were written for students of the ninth grade and above. Appendix D lists those tests which we reviewed.

One standardized test did meet the following essential criteria: the test (1) measured understanding of the "processes of science;" (2) drew from real-life experiences; and (3) was written for elementary school students. This test was the science subtest of the Sequential Tests of Educational Progress, or STEP Tests (ETS, 1958), a series of coordinated achievement tests covering a variety of academic areas for grade levels four to fourteen. The STEP Tests were designed to measure the "broad outcomes of general education, rather than the relatively narrow results of any specific subject matter course" (ETS, 1958, p. 5). For our purposes it was fortunate

Objectives for Process of Scientific Inquiry (Klopfer, 1971)	Goals for Individual Children in Problem Solving (USMES, 1975)
°Recognition of a Problem	°The child will identify and define the problem.
°Formulation of a working hypothesis	°The child will decide what information and investigations are needed in order to find some solution to the problem.
°Design of appropriate procedures for performing experiments	°The child will determine what needs to be done first. °The child will decide what is the best way to obtain the information needed. °The child will detect flaws in data gathering procedures or errors in the data itself.
°Processing of experimental data °Presentation of data in the form of functional relationships. °Interpretation of experimental data and observation.	°The child will organize, analyze, and interpret the data.
°Evaluation of a hypothesis under test in light of data obtained.	°The child will suggest some solution to the problem based on the data.
°Formulation of generalizations warranted by relationships found	°The child will suggest ways to implement the solution.

Figure 4.1. -- Mapping to illustrate the correspondence between Klopfer's (1971) objectives regarding the processes of scientific inquiry and the USMES developers' goals for individual children using USMES.

that their focus was "on skills in solving new problems on the basis of information learned" (ETS, 1958, p. 5). The science subtest was intended to examine each of the following specific skills:

1. Define problems
2. Suggest hypotheses
3. Select procedures
4. Draw conclusions
5. Evaluate critically
6. Reason quantitatively

Our examination of the actual items for the test's lowest level (IV: Grades four-six) proved encouraging. The context-dependent items are presented in sets of five to eight multiple choice questions. Each set of questions centers around a story in which several aspects of a situation are examined, hopefully a situation which is interesting and familiar to the children. The use of stories based upon such experiences as feeding guinea pigs and growing plants from seeds offered the possibility for testing students' process skills relatively independent of their success in the acquisition of facts.

Favorable reviews by Palmer Johnson and Julian Stanley in Buros' Mental Measurement Yearbook (5th ed., 1959) supported our view that selected items from the STEP Tests might play a useful role in our evaluation of problem solving. Both reviewers paid particular note to the test's concern for "everyday life and interests" (Johnson, 1959, p. 802), "commonly experienced by the age group" (Stanley, 1959, p. 803). Johnson distinguished the test as "unique" in its attempt

"to test understanding of the scientific method" (Johnson, 1959, p. 802).

Items from Forms 4A and 4B of the science subtest of the STEP Test were therefore considered in the construction of a new test appropriate for the evaluation of USMES. Combining the two forms, there were, in all, 120 questions to review. Each of the 120 questions was scrutinized closely. Despite published claims and reviews to the contrary, we felt that a large number of the items measured purely factual information, rather than a component of the problem solving method outlined by the STES-Science constructors. Those sets (groups of interrelated questions) which contained two or more "factual" questions were eliminated immediately. Next, each of the multiple choice items was reviewed for poor distractors, as determined by inspection rather than by item analysis. If more than half of the questions in any set contained poor distractors, the entire set was eliminated. Further, many topics found in Form A were repeated in Form B (e.g., "gardening"); in such cases, one of the sets had to be eliminated.

Using the methods described above, we discounted several sets of items. For our final choices, we selected those sets of items which appeared to offer the widest representation among the eight components of problem solving enumerated by the USMES developers as goals for individual children using USMES. These goals are presented in Appendix C. Five sets of items, encompassing thirty questions, were selected for inclusion in a new test of problem solving.

These ETS-developed STEP-Science items were to be used under licensing agreement between the Project Director of the USMES Evaluation and the Educational Testing Service. Additional items for the new test were developed by the evaluation team. The thirty ETS-developed items constitute Part I of the problem solving test shown in Appendix E, the team-developed items constitute Part II. Our team's item development efforts are detailed in the next section.

The Construction of New Items

One of the most important goals in the creation of a new paper-and-pencil test was to measure those aspects of problem solving which were deemed to be most critical to USMES by the USMES developers themselves. But from the beginning, one of the most difficult tasks for the evaluators was to determine an acceptable theoretic definition for "problem solving" which could be compared with what was being actualized by the USMES project. As we have already pointed out, the term "problem solving" has a myriad of meanings. We already knew that USMES purportedly delt with only "real" problems. However, "real" refers only to the context of the problem. Further definition was needed to delineate the processes which guided problem solving within the context of a "real" USMES challenge.

The USMES staff was reluctant to define problem solving as a series of sequential components. Their position, that problem solving did not occur in a pre-determined sequence, seemed totally justifiable. Yet to aid our measurement efforts, they were willing to identify

twenty components of "Problem Solving/Decision Making" which were relevant aspects of the USMES experience. These components were listed in Appendix A.

Could these components provide a useful outline of a test of problem solving skills for individual children? Careful scrutiny suggested that they could not, as we pointed out earlier. The components as listed are not goals for individuals during the course of one, or even two, USMES units. Instead, one finds listed here several of the possible components of the scientific process each of which may be experienced by some individual members of a total group during the course of an USMES unit. It is not expected that every child will have experience with or mastery of each and every component listed. Therefore, it did not make sense to base an individual test of skills and abilities on the complete set of components outlined in Appendix A. When we requested a list of the components of problem solving which could be regarded as objectives for each child as an individual (as opposed to group member). within each unit challenge, regardless of context, the USMES development staff responded with the list of eight components presented in Appendix C.

Using these eight components of problem solving as a framework, we began to develop an item pool. A preliminary set of guidelines helped bring the necessary limitations for item writing into focus:

a. The questions would employ multiple-choice format. The multiple-choice style was selected for versatility,

TABLE 4.2 (cont'd)

Item Number	Option #1	Option #2	Option #3	Option #4	Omits
25	37	6	7	49*	6
26	82*	5	5	3	6
27	4	7	55*	28	6
28	11	22	51*	6	8
29	16	31*	30	14	9
30	11	14	42*	23	11
		Part II			
41	4	85*	3	6	1
42	16	7	74*	4	0
43	9	8	58*	24	1
44	2	5	75*	17	0
45	13	9	70*	7	1
46	4	13	75*	8	0
47	18*	11	57	14	0
48	14	75*	6	4	1
49	5	65*	11	18	0
50	30	7	57*	5	0
51	25	9	31*	34	1
52	16	32*	36	16	1
53	15	20	20	45*	0
54	39	14	31	16*	0
55	26	13	30	30*	2
56	17	27	47*	8	0
57	14	16	22	48*	0
58	13	11	67*	9	2
59	8	10	72*	9	1
60	3	8	82*	6	1
61	19	61*	8	11	1
62	11	8	26	53*	2

*Correct response

We were further persuaded to select these three challenges because each provided a single problem focus, easily understood, yet one which might allow us to tap all eight components of problem solving from recognition of the problem to its resolution.

An introductory paragraph was drafted for each of the problems. Eight sample items were created for "Pedestrian Crossing." Each of the eight items was designed to correspond with a specific USMES component. These sample items were to be used by the item writers as models.

Then we enlisted several classes of undergraduate and graduate students in the field of education to write items. Approximately fifty students, representing a variety of experiences and skills in education were involved in the creation of the item pool. Each group of item writers was given a two-hour briefing which included instructions and information on the following topics:

1. The nature of USMES and USMES challenges.
2. The evaluation of problem solving abilities.
3. The components approach to the concept of problem solving and the USMES components.
4. Rules and guidelines for writing multiple choice questions.

The sample items were distributed to the item writers along with USMES manuals for the selected challenges. After the sample items were discussed and critiqued, the students began writing their own items, working in small groups or individually. Each writer was asked

to develop eight items for each of the three problems. The eight items were supposed to correspond to the eight components of problem solving which had been presented and discussed during the introductory lecture. Final items were collected approximately three weeks later.

The task of editing these items came next. Many items had to be eliminated because they contained faulty logic or because they were merely repetitious of the original sample items. In the first phase of editing and rewriting, nine sets were developed, three for each of the three problems. These sets of eight questions each, were based upon the "best" sets of items from the original pool. Editing consisted of replacing particularly "bad" items with better items from other sets, improving the quality of distractors, and scrutinizing the newly created sets for faulty logic and ambiguities.

These nine sets of questions were duplicated and submitted to several measurement, philosophy, and reading specialists for their reactions. Critiques of the item sets took the form of conferences in which the style and content of the test, as well as individual items, were discussed in some detail. Advice from Dr. Elizabeth Reynolds, a reading specialist, was particularly helpful for shaping our revisions.

Several problems remained to be solved. The most difficult dilemma lay in the logical structure for the sets of items. As originally constructed, the items were not independent, i.e., the answer to one question was frequently stated in the following question. This format insured that students would not be penalized because they

followed the path of incorrect distractors. On the other hand, this format made it possible for the student to infer a correct answer to each question by referring ahead to the indicated correct path of questions moving from beginning to end of the problem set.

Two possible solutions to this dilemma seemed feasible. First, the test could be physically constructed and administered in such a way as to make it difficult for students to go back and change answers. Or, the test could be revised so that the answers to questions were not so clearly spelled out in other questions. At this point the decision was made to pursue the latter alternative and rewrite. We finally arrived at one "best" possible set of items for each of the three problems by reducing the number of sets from nine to three. Further revisions were made following the administration and interpretation of the pilot testing.

The three sets of questions which comprise the final draft incorporate several aspects of each problem. The correct answers depend upon reasoning as much as possible. Answers to questions are not as blatently spelled out in other questions as they had been in earlier drafts of the items. A limitation in the final set was that skills could not be represented equally among the item because some components were more difficult to measure in the form of multiple choice questions than were others. The complete test of thirty ETS-developed items and twenty-two team-developed items was entitled the "Test of Problem Solving Skills," or TOPSS; it appears in Appendix 1.

Pilot Testing the Test

A. Sample

In an effort to evaluate TOPSS, we administered it in June, 1975 to students from working class families in Arlington and Watertown, Massachusetts, two predominantly white, multi-ethnic, urban/suburban communities adjacent to Boston. These sites were chosen because they could provide us with both USMES and non-USMES classes, in grades four, five, and six, and because their teachers and administrators were willing to participate in this research project. Furthermore, our field staff in these communities were available to administer the tests. Pressed to complete the testing program for the 1974-75 USMES evaluation project itself, other USMES schools were not asked to participate in this additional pilot testing. Instead, we had to seek sixth-grade students from a non-USMES community, Fairfield, California, where teachers had agreed to administer the test themselves.

Much less densely populated than the two Massachusetts communities, Fairfield is a working class community located about midway between San Francisco and Sacramento, California. Afro-American, Mexican-American, and Oriental-American students were heavily represented in the Fairfield group. The children of both white and minority enlisted servicemen assigned to Travis Airforce Base were also heavily represented in the Fairfield sample, The distribution of the 398 students who were involved in the pilot testing is given by community and by grade level in Table 4.1. In all, seventeen classes were represented.

TABLE 4.1

Distribution of the Sample
for the Pilot Test

Community	Grade 4	Grade 5	Grade 6
Arlington, MA	57	77 (25 USMES)	0
Watertown, MA	53	69	75 (26 USMES)
Fairfield, CA	0	0	67
Total	110	146	142

Before the new problem solving test is adopted for wide-spread use, additional pilot testing with larger samples of wider geographical distribution might be considered. The USMES evaluation team would like to have pilot tested the new test more extensively, but our resources and entree into schools had been taxed heavily with a very extensive USMES evaluation program and with large scale pilot testing of another instrument to measure attitudes.

B. Administration

The new paper-and-pencil problem solving test was administered by our field staff/observers in Arlington and Watertown, Massachusetts, and by the classroom teachers themselves in Fairfield, California. A problem resulted in the California administration. While we could deliver the test materials to the observers in Massachusetts, giving them an opportunity to review the instructions and receive clarification for their questions, we had to mail the test materials to the California teachers, and any further clarification had to be offered by telephone. A misunderstanding occurred in this latter case, and the classroom rating forms asking teachers to rank-order their students according to overall-real problem solving ability were misapplied. As a result, that part of the generated data could not be used.

The instructions for the STEP-Test items and USMES items (cf. Part I and Part II respectively of Appendix I) are a slightly modified version of the instructions for the original STEP Test. Those few changes which we made related primarily to the USMES items.

We had decided that children taking the test should not be encouraged to go back and check their answers; this direction was intended to prevent perceptive children from finding the answer to a previous question in material given later on in the test. Therefore, those passages in the STEP text which reminded the children to go back and check over their answers had to be eliminated from the instructions for the USMES items. Otherwise, our new text's instructions were almost identical to those of the original STEP Test.

Part I of the test contained the thirty ETS developed items numbered sequentially "1" through "30." However, one might note that the twenty-two team-developed items constituting Part II were numbered "41" through "62." The Digitek answer sheets which we used for the testing could be scored by subtest only if the subtest responses were located on specified sections of the answer sheets. Thus, the item numbers "31" through "40" were omitted deliberately.

C. Teachers' Ratings of Students' Problem Solving Abilities

While the children were taking the test, their teachers were asked to fill out a rating form designed to measure the problem solving ability of individual students on a comparative basis. We created this form in an attempt to examine the construct validity of the test. The teachers were directed to rate their students on a scale of 1 to "n" (where "n" = the number of students in the class) as to which student is the best problem solver in the class, which is second best, and so forth. Because "problem solving" is so ambiguous,

we gave the following directions hoping to standardize their responses to some extent:

> "In the right-hand column put a 1 next to the name of the student who copes most effectively with problems which arise in daily activities. Put a 2 next to the name..." (cf. Appendix J)

This rating scale was used to obtain a rough estimate of the teacher's perception of the problem solving ability of each student.

Technical Information

Three technical aspects of measurement must be considered in judging the effectiveness of a test. First, does each item discriminate properly? Does the person who is a good problem solver, as defined by the total score on this test, get the item correct? Next, is the test reliable? That is, if the student takes the test over again, will he maintain approximately the same relative position in the group? And finally, but most importantly, is the test valid? Does the test really measure problem solving ability, or is it measuring something else, perhaps science content or test wiseness? The next sections are devoted to these issues.

A. Item Analysis

Tables 4.2, 4.3, and 4.4 contain the item analysis data. Table 4.2 gives the percentage of students replying to each option. In Table 4.3, columns 2 and 3 give the point-biserial correlation of each item with its subtest score and with the total test score. Because of the nature of the data, biserial coefficients seemed justifiable. These coefficients are given in Table 4.4.

TABLE 4.2

Percentage Distribution of Students (N = 398) for Options to Each Item on the New Paper-and-Pencil Test of Problem Solving Skills (TOPSS)

Item Number	Option #1	Option #2	Option #3	Option #4	Omits
		Part I			
1	54*	12	13	20	1
2	24	54*	14	7	0
3	25	9	20	46*	0
4	10	15	59*	15	1
5	10	9	68*	11	1
6	68*	18	8	6	1
7	17	16	14	52*	1
8	5	6	83	4	2
9	8	68*	6	16	1
10	70*	13	11	5	1
11	4	3	6	85*	2
12	3	21	13	61*	1
13	18	44*	16	20	2
14	6	14	74*	5	1
15	4	65*	28	3	1
16	24	51*	19	5	1
17	5	3	89*	1	1
18	18	17	57*	7	1
19	31	44*	9	15	2
20	34	18	29*	17	2
21	11	75*	4	6	2
22	43*	18	23	13	3
23	13*	31	36	16	4
24	18	5	7	66*	5

*Correct response

TABLE 4.2 (cont'd)

Item Number	Option #1	Option #2	Option #3	Option #4	Omits
25	37	6	7	49*	6
26	82*	5	5	3	6
27	4	7	55*	28	6
28	11	22	51*	6	8
29	16	31*	30	14	9
30	11	14	42*	23	11
		Part II			
41	4	85*	3	6	1
42	16	7	74*	4	0
43	9	8	58*	24	1
44	2	5	75*	17	0
45	13	9	70*	7	1
46	4	13	75*	8	0
47	18*	11	57	14	0
48	14	75*	6	4	1
49	5	65*	11	18	0
50	30	7	57*	5	0
51	25	9	31*	34	1
52	16	32*	36	16	1
53	15	20	20	45*	0
54	39	14	31	16*	0
55	26	13	30	30*	2
56	17	27	47*	8	0
57	14	16	22	48*	0
58	13	11	67*	9	2
59	8	10	72*	9	1
60	3	8	82*	6	1
61	19	61*	8	11	1
62	11	8	26	53*	2

*Correct response

TABLE 4.3

Point-Biserial Item Correlations (PB)
with Sub-Test (ST) and with
Total Test (TT) Scores

Item Number	PB-ST[1]	PB-TT[1]	PB-ST[2,3]	PB-TT[3]
		Part I		
1	.25	.23		.23
2	.24	.18		.19
3	.20	.15		.14
4	.49	.52		.52
5	.44	.42		.43
6	.34	.27		.30
7	.39	.36		.37
8	.55	.53		.54
9	.48	.51		.50
10	.52	.50		.50
11	.41	.39		.40
12	.38	.38		.38
13	.33	.30		.31
14	.39	.36		.36
15	.36	.29		.31
16	.39	.38		.38
17	.44	.41		.42
18	.33	.3[illegible]		.31
19	.45	.44		.44
20	.33	.33		.34

[1] Based on all items

[2] These remain unchanged for Part I

[3] Items 50-55 deleted

TABLE 4.3 (cont'd)

Item Number	PB-ST[1]	PB-TT[1]	PB-ST[2,3]	PB-TT[3]
21	.31	.31		.31
22	.44	.41		.42
23	.05	.00		.01
24	.47	.47		.47
25	.32	.32		.32
26	.46	.43		.44
27	.38	.37		.37
28	.40	.36		.37
29	.21	.15		.16
30	.43	.41		.41
		Part II		
41	.38	.33	.42	.34
42	.41	.40	.44	.41
43	.33	.26	.32	.25
44	.50	.45	.51	.44
45	.45	.42	.48	.42
46	.51	.49	.55	.49
47	.24	.24	.27	.24
48	.50	.47	.52	.47
49	.40	.36	.41	.34
50	.37	.32	-	-
51	.21	.13	-	-
52	.16	.06	-	-
53	.35	.24	-	-

[1] Based on all items

[2] These remain unchanged for Part I

[3] Items 50-55 deleted

TABLE 4.3 (cont'd)

Item Number	PB-ST[1]	PB-TT[1]	PB-ST[2,3]	PB-TT[3]
54	-.04	-.09	-	-
55	.08	-.03	-	-
56	.39	.33	.41	.33
57	.47	.45	.48	.45
58	.44	.43	.46	.43
59	.55	.52	.58	.52
60	.49	.49	.52	.47
61	.43	.36	.46	.36
62	.37	.31	.40	.32

[1] Based on all items

[2] These remain unchanged for Part I

[3] Items 50-55 deleted

TABLE 4.4 (cont'd)

Item Number	BS-ST[1]	BS-TT[1]	BS-ST[2,3]	BS-TT[3]
21	.42	.42		.43
22	.56	.52		.53
23	.07	.00		.02
24	.61	.61		.61
25	.40	.40		.40
26	.67	.62		.63
27	.48	.46		.47
28	.51	.45		.46
29	.28	.19		.21
30	.55	.51		.51
		Part II		
41	.58	.50	.63	.51
42	.55	.54	.60	.55
43	.41	.32	.41	.31
44	.68	.61	.69	.60
45	.59	.55	.63	.55
46	.70	.66	.75	.67
47	.36	.35	.39	.36
48	.68	.64	.71	.64
49	.52	.44	.53	.43
50	.46	.40	-	-
51	.27	.17	-	-
52	.21	.08	-	-
53	.44	.30	-	-

[1] Based on all items

[2] These remain unchanged for Part I

[3] Items 50-55 deleted

TABLE 4.4 (cont'd)

Item Number	BS-ST[1]	BS-TT[1]	BS-ST[2,3]	BS-TT[3]
54	-.06	-.14	-	-
55	.10	-.04	-	-
56	.50	.42	.51	.41
57	.59	.57	.60	.56
58	.57	.55	.60	.55
59	.73	.69	.77	.69
60	.70	.67	.75	.68
61	.55	.46	.58	.46
62	.46	.39	.50	.40

[1]Based on all items

[2]These remain unchanged for Part I

[3]Items 50-55 deleted

Both the biserial and point-biserial correlation coefficients are an indication of the ability of an item to discriminate properly. A coefficient of +1.00 indicates that those students who are good problem solvers (as defined by a high score on this test) got this item correct. As the index approaches 0.00, the ability of the item to discriminate properly grows weaker, and at 0.00 it does not discriminate between good and poor problem solvers on the test as a whole.

Although the data in Table 4.3 are interpreted easily without explanation, a few comments are offered.

(1) The item correlations for the test developed specifically for USMES (Part II) are not very different from the correlations on the STEP Test items (Part I). Considering the time, staff and money available to ETS, the USMES evaluation staff is pleased overall with these results for Part II.

(2) One set of items in Part II is quite weak. Those items, numbered 50-55, constitute the set designed for the "Lunch Lines" problem. It appeared that these items probably were adding little, if anything, to the overall test. Therefore, it was decided to drop items 50-55 from the test and to reanalyze the data. If the test could be shortened, while the technical aspects of the test were retained or improved, we would shorten the test.

Columns 4 and 5 of Table 4.3 contain the data obtained from the second analysis. Since no items were added or deleted from Part I (1-30) the correlations between each item and the subtest score remain the same as the correlations obtained from the first analysis. Inspection of all other correlations indicates however, that although the differences are slight, in most cases the correlations are higher for the shorter test. The same is true of the biserial coefficients given in Table 4.4. This was not sufficient evidence to decide whether to drop the six items from the final version of the test, however. The effect of the deletion upon the reliability and validity had also to be studied.

Reliability

Since test-retest or parallel forms methods of determining reliability were not feasible, internal consistency reliability estimates were obtained for Part I, Part II and the total test using Hoyt's estimate of reliability. This information, shown in Table 4.5 was used to answer two question: how reliable was the original 52-item test, and what was the effect of dropping six items on the test reliability?

While certainly not as high as one would expect from an achievement test which deals with a readily defined body of knowledge, the reliability coefficients for TOPSS, particularly the shortened version of TOPSS, are very respectable for a test of manageable length which deals with "real-life" problem solving, where such a clear definition

TABLE 4.5

Reliability Estimates for the Original TOPSS and the Shortened TOPSS

	Part I		Part II		Total	
	Original Test	Snortened Test	Original Test	Shortened Test	Original Test	Shortened Test
n	398	398	398	398	398	398
r_{tt}	.79	.79	.69	.74	.84	.86
S_e	2.36	2.36	1.98	1.61	3.13	2.90
$\bar{x}$	17.25	17.25	12.35	10.25	29.60	27.50
SD	5.18	5.18	3.64	3.25	7.79	7.73
# of items	30	30	22	16	52	46
highest score	28	28	20	16	48	44
lowest score	2	2	2	2	10	8

is lacking. One should also remember that the lower reliabilities for the parts of the test (.79 and .74 respectively for the Parts I and II, versus .86 for the total) are affected by the smaller number of items on the parts versus the whole.

Probably of greater interest to the potential user of this test, however, should be the standard error of measurement (S_e), which is an indicator of the amount of variability one can expect in an individual's true score. Again, we feel that the standard errors (2.36 and 1.6. for the parts, 2.90 for the total) fall within an acceptable range for a test 44 items long. Table 4.5 indicates that deleting six items from Part II raised the reliability of Part II from .69 to .74, and the reliability of the entire test was maintained at approximately the same level, .84 versus .86. In light of these results, not only do we conclude that the test in its original form is reliable, but we have additional information encouraging the deletion of items 50-55.

Validity

Establishing the validity of a test is always difficult, except for those tests where content or face validity is deemed sufficient. Unfortunately, content or face validity is not satisfactory for this test, since one cannot tell by looking at an item whether or not it is tapping the kinds of skills which USMES purports to teach.

Rather, problem solving must be considered a construct and therefore the test should be validated through construct validation

procedures as suggested by Cronbach and Meehl (1955). These procedures involve looking for other behaviors which should correlate with the scores on the test. If predicted results are supported by the observed correlations, one has not proved that the test is measuring the construct "real-life" problem solving, but the results lend support to that supposition.

We hypothesized that if this test was measuring "real-life" problem solving, the following should be noted:

(1) There should be growth in the problem solving skills measured by the test from grade four to grade six, with grade six students achieving higher scores than students in grade four.

(2) Given teachers' rankings of their students' abilities in problem solving, there should be a high positive correlation between the teacher's rankings and the students' scores on the test.

(3) Part I of the test, which consists of items from the STEP Tests which purport to measure problem solving processes should correlate well with scores on Part III of the test which purports to measure components of problem solving in USMES type problem situations.

(4) USMES classes which had experience with a unit which served as the basis for one of the problem scenarios in Part II of TOPSS should score higher on the test

than other USMES students who, in turn, should score higher than non-USMES students.

Information relevant to these hypothesis is offered below:

(1) Hypothesis #1. We predicted that there would be growth in problem solving skills as indicated by a higher mean score on the test, across grade levels, from grade four to grade six. Cross-sectional evidence supporting this hypothesis was obtained. The mean for grade four students was 26.5; for grade five -- 29.43; and for grade 6 -- 33.2.

(2) Hypothesis #2. We predicted there would be a high positive correlation between teachers' rankings of the students' abilities in real problem solving and their students' scores on the test. Table 4.6 gives the correlations between teachers' rankings and Part I, Part II and the total score.

Two things should be noted for this data. First, correlations are not reported for classes #16 and #17, as the teachers did not follow the directions properly and their rankings were not usable. Second, correlations were computed by class, since teacher rankings were offered by class, and a rank ordering of students' abilities was not possible.

These correlations between teacher rankings and students' TOPSS scores are offered in Table 4.6 as evidence for the construct validity of TOPSS. The correlations across classes are quite variable -- a few disappointingly low, one strikingly high at .91. The median value was .68. We feel these data do offer promise for the validity of TOPSS.

TABLE 4.6

Spearman Rank-Order Correlations Between Teacher Ratings and Students' TOPSS Scores, by Class

Class[a]	n	Part I Scores and Teacher Ranks	Part II Scores and Teacher Ranks	Total Scores and Teacher Ranks
1	14	.231	.495[1] (.3715)[2]	.5595[1] (.5077)[2]
2	18	.585	.699 (.7173)	.576 (.7132)
3	20	.647	.546 (.6060)	.725 (.8300)
4	24	.738	.4818 (.4196)	.8192 (.7027)
5	23	.781	.865 (.8993)	.8959 (.8765)
6	21	.516	.703 (.7633)	.6221 (.6702)
7	17	.309	.632 (.6986)	.483 (.4902)
8	12	.727	.350 (.3637)	.5595 (.4546)
9	16	.738	.712 (.6412)	.779 (.7500)
10	16	.818	.772 (.8000)	.835 (.8442)
11	20	.412	.682 (.7412)	.447 (.5977)
12	20	.913	.605 (.7031)	.879 (.8963)
13	23	.705	.770 (.7977)	.777 (.7962)
14	26	.631	.791 (.6404)	.7224 (.6971)
15	23	.33	.556 (.6216)	.330 (.5050)

[1]The correlation between teacher ranking and the original 52 item test.

[2]The correlation between teacher ranking and the test with items 50-55 deleted.

[a]Rankings for classes 16 and 17 could not be computed because teachers did not apply the form correctly.

As to the few low correlations, several factors may have suppressed these values besides test invalidity. In a very homogeneous group of students, it would be difficult for a teacher to rank order the group. Further, correlations based on small groups tend to be low; sample size affects the observed correlation. The lowest correlation in Table 4.6 ($p = .23$) was obtained for the smallest class ($n = 14$). Finally, the assumption underlying the presentation of these correlational data as evidence of a construct validity is that teachers who had worked closely with the students over the period of a year or more were in a good position to judge the real problem solving abilities of their students. Perhaps a few of the "expert judges" were not so perceptive, and their ratings offered a poor criterion.

(3) Hypothesis #3. This hypothesis dealt with the correlation between parts of the test. We reasoned that if the STEP Test purports to measure problem solving and has been validated by ETS for that purpose, then a high positive correlation between Part I and Part II should help to establish that the ETS items and the team items are measuring much the same thing. Table 4.7 presents the intercorrelations for the parts of TOPSS. We feel that the correlations in Table 4.7 are sufficiently high to offer additional support to our claim for the validity of TOPSS.

(4) Hypothesis #4. The relative performance of USMES and non-USMES classes as predicted in this hypothesis could not be tested adequately because of the limitations to sampling. The

TABLE 4.7

Correlations Between the Parts of TOPSS

	Part I	Part II	Total
Part I	—		
Part II	.623 (.665)	—	
Total	.934 (.949)	.861 (.866)	—

Note.-- The correlations based on the original version of TOPSS are given without parenthesis; the correlations based on the shortened version of TOPSS, with six items deleted, are contained in the parenthesis.

test of this important hypothesis remains for future research.

Despite our inability to test the last hypothesis adequately, we feel that results of the pilot study of TOPSS reported in this chapter offer promising evidence for the effectiveness of this new test as a reliable, valid and practical measure of the problem solving skills of elementary school students.

Reflections on the Measurability of Components of Complex Problem Solving

The feasability and desirability of assessing problem solving skills by means of a paper-and-pencil test were discussed in Chapter III of this report. For large scale evaluation of USMES, we proceeded on the assumption that if the skills of problem solving were to be examined through traditional testing methods, they must be broken down into identifiable components which can be represented as behavior objectives for individual students. In the interest of future test development along similar lines, we would like to conclude this chapter by examining each of the eight components in Appendix C which were identified by the USMES developers as goals for individual children in problem solving, and which were used as the basis of our TOPSS test. Thus, we offer our perspectives on the degree to which each component skill lends itself to multiple choice testing.

Objective #1: The child will identify and define the problem.

In the most "real" sense, the identification and definition of a problem is a highly personal reaction. Situations which present themselves as a problem to one person may not be problematic at all to another. A situation is not a problem if (1) a person knows exactly how to respond to it, or if (2) a person sees nothing amiss about the situation in its original presentation. It is quite difficult to cite a situation which would be viewed clearly as a problem to all readers. Moreover, the variety of pertinent and subtle factors which make up a "real" problem situation does not lend itself to the level of abstraction required in the writing of a multiple choice question. Finally, it is difficult to declare any answer as wrong with justification.

Objective #2: The child will decide what information and investigations are needed in order to find some solution to the problem.

This component again borders on the ideosyncratic. The kinds and amounts of information required to resolve a problem depend on the manner and degree to which the child sees the situation as problematic. Furthermore, it is difficult to limit the paths toward solution -- and their accompaning sets of information and investigations -- to the singular. The distractors arranged among the multiple choice options were either obviously not plausible to even the poor problem solver, or were valid answers. Any discrimination between poor problem solvers and strong ones was difficult to obtain.

Objective #3: The child will determine what needs to be done first.

Questions on this component were limited to the most basic sequence of procedural activities. In a "real" situation, the most fruitful course of action would depend heavily on the values of the individual and the context in which the problem resides. The examiner could not validly order them for the student. One could only presume that a gathering of facts must precede a plunging into the "work" of any problem. Yet, questions asked the student to identify this basic sequence.

Objective #4: The child will decide what is the best way to obtain information needed.

Some methods of measurement yield information which is more appropriate to a given problem than others, and some methods of measurement provide more accurate results than others. Skills in information gathering and measurement techniques can be investigated fairly well with multiple choice questions.

Objective #5: The child will detect flaws in data-gathering procedures or errors in the data itself.

As the preceding component, this skill is based upon experience with the activities of data collection. The potential for testing this skill with a multiple choice test is, as above, relatively high.

Objective #6: The child will organize, analyze and interpret data.

If the skills called upon in this component are related to the area of mathematics, then the knowledge of appropriate techniques is quite testable. However, there are other types of "real" problems in which the artistry involved in organizing and interpreting the data is too idiosyncratic to be explored by multiple choice questions.

Objective #7: The child will suggest some solution to the problem based on the data.

In order to write items for this component, we first had to supply some specific data on a given problem and then offer a limited interpretation on that data. On this basis, we asked the student to select a "best" solution. Again, the abstracted artificiality of this test required the student to do his reasoning without the benefit of an actual testing ground or the knowledge of "real" criteria of success, criteria present in real-life situations.

Objective #8: The child will suggest ways to implement the solution.

The political knowledge and power a child has available to implement his solutions to problems are much too limited to provide us with the bases for many interesting questions in this area. A review of USMES logs reveals that when applied this element of problem solving usually consisted of a presentation of the findings to the appropriate authorities.

These and other limitations to the paper-and-pencil testing of components of problem solving underscored the need to develop a process approach for observing and questioning the child as he/she pursues the solution of a complex problem. Designed to meet this need, the PROFILES technique is described in the next chapter.

CHAPTER V

THE DEVELOPMENT OF PROFILES: AN INTERVIEW/OBSERVATION TECHNIQUE TO ASSESS PROBLEM SOLVING PROCESSES IN CHILDREN

Rationale for Profiles

Although USMES unit challenges have been thoroughly documented for the purposes of curriculum development, there has been no previous attempt to examine the individual child's experiences and thought processes as he participates in a challenge from the formulation of the problem to its final resolution. And, from the point of view of this researcher in educational psychology, there was little evidence that the abilities to solve complex problems were not a function of developmental stages; there was noteworthy evidence to the contrary, as noted in Chapter II. PROFILES is an assessment technique newly designed to address these issues. In PROFILES, attention is focused on the child's grasp of the process of problem solving, rather than on the discrete activities in which he may participate along the way.

With this measuring technique, a limited number of children should be selected and periodically interviewed and observed on a regular basis, while they are involved in a classroom USMES challenge. (This technique could also be used to assess the processes students employ with non-USMES, group problems or projects which are complex, long-term, and designed to be meaningful to the students.) In each periodic

review, attention should be focused on the child's individual perception of the problem solving process and on such related factors as (1) his ability to explain his current activities in terms of a larger plan of action, (2) his understanding of the steps which preceded his current work, and (3) his ability to forsee the consequences and purposes of future activities.

Because decision making and planning are crucial aspects of the problem solving process, attention in PROFILES is also focused on the child's perception of the source of the decision making, whether it be the teacher, the group, or himself. We are interested in determining the extent to which the child experiences that he himself is directing the processes in which he is involved.

Observational information collected by the observer prior to each interview puts the child's responses into context and provides the criteria necessary for evaluating the extent of his understanding of the processes in which he is involved. A complete set of data for each child consists of a series of observations and interviews in which the child has been given the opportunity to express his accumulated understanding of the USMES (or other complex project) experience.

Research in child development, particularly research based on the theories of Piaget, has lead to important questions, as yet unresolved, regarding the young child's ability to internalize abstractions and to formulate generalizations. These questions have particular relevance to USMES. It is the contention of the developers that experience with a series of "real" problems results in the development

of useful problem solving skills in young children. If this is the case, USMES may provide evidence that young children can master the conceptual logic inherent in the problem solving process. One of the aims of PROFILES is to search for such evidence.

Development of the PROFILES Interview Schedule

The central question of our inquiry was: Can elementary school children acquire an understanding of the problem solving process which is transferable to other dilemmas of "real" life, through the discovery experiences provided by USMES challenges?

As we began to search for means to examine this question, certain limitations presented themselves immediately. For example, we were unable to observe, and therefore assess, the children's successes and failures in dealing with problems outside of the institutional environment of the school. In addition, we could not reasonably expect students to work through non-USMES, USMES-like "real" problems simply for the purposes of observation unless we limited the problems to the extent that they are no longer "real" in length or in depth. Furthermore, we could not measure the children's ability to work through verbal abstractions of "real" problems without bringing into play the less relevant variables of the verbal ability and the intelligence of the child. Finally, we could not draw conclusions from performances on limited, selected aspects of problem solving for there is no single aspect that is known to be indicative of a generalized ability to solve problems.

Given these methodological limitations, we first decided to explore the potential for examining children's actual experiences with USMES challenges by means of a modified case study approach. The case study model, as ordinarily utilized, proved to have limited utility for our purposes. It is designed to describe unique deviations from a norm or to make a diagnosis -- purposes outside our stated objectives. However, the concept of a case study provided a starting point, the final result of which was the PROFILES method for collecting information about individual students through sequential interviews and observations over an extended period of time.

We began with a set of categories based on the developers objectives for USMES and derived from a classical analysis of problem solving (Dewey. 1910), interfaced with descriptions of USMES challenges. (See Appendix K for this initial outline of categories.) The questions which were to be used by those observers who would explore the children's experiences and examine these aspects of problem solving evolved slowly through a series of trials and refinements. A set of tentative interview questions was developed prior to the initial classroom visits.

Next, USMES teachers in the Greater Boston area were contacted. and three observers began to visit classes in which two children, randomly selected, were observed and interviewed. The lack of functioning USMES classes was a impediment which we had not forseen and which severely reduced the number of trials which we were able to carry out during the instrument development phase. Through the

efforts of three trained observers, we had planned to study with periodic observation and interview a total of 42 children in 21 classes over a three-month interval. Yet, only ten local classes had operative USMES units between January and May, 1975. Of these, some classes had only two or three sessions to observe over that period of several months.

In piloting PROFILES during the first few visits to each class, the observers concentrated on the effectiveness of the interview questions, making notes on which questions elicited the most illuminating responses, and which questions needed to be modified or eliminated. On the basis of observer debriefings between their visits to classes, questions were rewritten and categories were frequently adjusted to reflect the realities of the classroom. (The final set of problem solving behavior categories to be observed with PROFILES is shown in Appendix L.)

As the questions began to take shape, observers shifted their concentration from the questions they asked to the method by which they received the childrens' responses. At first the observers attempted to take verbatim notes of responses, but this procedure proved to be both difficult and unreliable. The decision was then made to tape record the childrens' interviews, a method which proved to be more satisfactory.

The final forms of the interview questions which were generated are shown in Appendices M and N. These questions, consolidated into the format of a branched interview schedule, became the standard

starting point for all discussions with the children. The questions are flexible enough to be used with any USMES unit. At the same time, they are sufficiently structured to provide standard common elements which remain constant from one interview to another, whether they are conducted by a single interviewer or by several different interviewers.

Only those questions which were understood by every child who participated in the pilot study were included in the final draft. Two forms of the interview schedule were developed, one for the second and third grades (Appendix M) and one for the fourth through sixth grades (Appendix N). The vocabulary of the original set proved to be too difficult for younger children; they could not understand terms such as "involvement," "decision," "suggestions," and "problem." Some vocabulary was changed, and some questions were reworded for ease of understanding. We paid careful attention to ascertain that both forms of the questions tap the same general perceptions and understandings of the problem solving process.

Importance of Observer Training and Monitoring

The PROFILES method requires rigorous and intensive training of the observers. Although an observation form is provided for the observers, along with a "script" for the child interviews, a thorough grasp of problem solving processes is essential if useful and significant data are to be collected. It is impossible to develop totally structured interviews, due to the mixed nature of the various USMES units. Therefore, the success of each interview

must depend upon the observer's ability to frame probing questions which are both appropriate to the specific situation and relevant to the entire process. Observers must, in some instances, make their own judgements and decisions as to which line of questioning they might pursue most profitably. Unless each observer can identify skillfully responses and behaviors which might be indicative of problem solving, irrelevant questions may be asked, and relevant ones may be omitted from the interview.

In addition to careful scrutiny of the problem solving process, it is important that all the observers and interviewers pursue training together, or at least that they begin data collection from a common frame of reference. Otherwise, the interreliability ratings between observers will be dangerously low. Most importantly, continual monitoring of observers is necessary in order to make certain that observers pursue uniform guidelines. Previous experience has shown that even a highly structured interview schedule will permit the element of "drift" to occur. Unless corrected, such a drift can invalidate a study. As documented by many researchers, it is extremely important to monitor observers in order to check the degree of drift which gradually emerges after the initial training sessions. Either additional training sessions should be held, or some other method should be employed whereby observers can be carefully monitored and informed of any deviations from the intended interview technique.

Drift can have very serious implications, invalidating the child interviews. Among the more serious emerging problems are promptings by the observer and his or her deviations from the standard procedure. Nevertheless, these difficulties can be controlled. Indeed, effective control is one of the most important factors to be considered in the training and continual monitoring of observers.

Procedures for Administration of PROFILES

For use in the evaluation of USMES, the PROFILES technique should be administered to approximately 25 to 30 USMES classes. Two children selected randomly from each class should be observed and and interviewed six to eight times over the course of an entire USMES unit. The time lapse between each interview should be approximately two weeks, although this will depend upon the length of the unit and the intensity with which it is taught.

The observer should observe each class for approximately 30 minutes, during which time he/she will become familiar with the class activities and will fill out the Observation Form shown in Appendix O. The observer must determine first which of the 11 categories of the problem solving process are applicable to the particular class under observation, then check those categories on the observation form, and finally fill out the questions falling under each of the selected categories.

After the 30 minute observation period, the observer should take each of the two children randomly selected to a quiet room and

interview him/her on tape using the appropriate "script" from the PROFILES Interview Guides provided to the observer. Observers should ask only those questions pertaining to the categories they checked during the 30 minute observation period. They should try to determine the child's understanding of the place his work assumes in the total group problem solving efforts. Observers should be thoroughly informed about the PROFILES technique during intensive training. An instruction manual has been prepared for the observers' administration of PROFILES (see Appendix P) and several examples of actual observations and interviews have been provided in Appendix Q.

Development of Scoring Protocol for PROFILES

Through the use of the tapes and observation forms, PROFILES scorers should be trained to analyze the results of the interviews. The observation forms provide the scorers with the classroom context, i.e., the actual events which occurred in the classroom on a particular day as the class engaged a USMES unit. Through this form, the observer will have noted those components of the problem solving process which were actually engaged and where these activities fit into the overall design of problem solving teaching as conceived by the USMES curriculum developers.

The tapes provide the scorers with the perceptions of the students as they were involved in that same set of activities. Through the individual interviews, the scorers should be able to assess the students' degree of understanding of their activities as

they relate to problem solving: (1) what relationship the day's activities had to problem solving, (2) whether the student can explain the particular activities in terms of the total problem solving process, (3) whether he can recall the steps which preceded his current work, and (4) if he can forsee the consequences and goals of his present activities.

The scoring protocol developed for this study uses a simple rating scale for the various questions relating to each of the eleven components of the problem solving process, plus a reliability check. (See Appendix R for the "Instructions for Using the Scoring Protocol.") The protocol has been developed to allow for many correlations during the analysis of the results.

The protocol was piloted in a trial scoring of several interviews. Two scorers, previously involved in scoring Playground and Picnic Problems in the 1974-75 USMES evaluation, were trained by one of the people involved in the development of the Profile Interviews. Two interviews were scored simultaneously by all three people, in order to demonstrate the protocol to the two new scorers. A third interview was then scored individually by each person and concurrence was reached on 85% of the scored questions. The 15% discrepancy was due mainly to an ambiguous wording of some of the questions. These have since been clarified and, at this point, the scoring protocol seems to be appropriate and unambiguous.

The evaluators are generally satisfied with PROFILES as an instrument to complement the Test of Problem Solving Skills for the

evaluation of the problem solving in USMES. The ability of the PROFILES technique to examine the individual child's experience of the USMES units from the formulation of a problem to its final resolution and to estimate his understanding of the process of problem solving as a whole offers a new and significant technique for the assessment of complex problem solving processes. We hope that both TOPSS and PROFILES can benefit the research development, and evaluation of curricula designed to teach complex problem solving.

REFERENCES

American Association for the Advancement of Science. Science--A process approach. Competency measures, part five. (AAAS misc. pub. 66-23) Washington, D.C.: AAAS, 1966.

American Association for the Advancement of Science. Science--A process approach. Competency measures, part six. (AAAS misc. pub. 66-26) Washington, D.C.: AAAS, 1966.

American Association for the Advancement of Science. Science--A process approach. Competency measures, part seven. (AAAS misc. pub. 66-29) Washington, D.C.: AAAS, 1966.

Ausubel, D.P., & Robinson, F.G. School learning: An introduction to educational psychology. New York: Holt, Rinehart & Winston, 1969.

Bayles, E.E. Democratic educational theory. New York: Harper & Row, 1960.

Beilin, H. Developmental determinants of word and nonsense anagram solution. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 523-527.

Bourne, L.E., Ekstrand, B.R., & Dominowski, R.L. The psychology of thinking. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1971.

Brandwein, P.F., Watson, F.G., & Blackwood, P.E. Teaching high school science: A book of methods. New York: Harcourt, Brace & World, 1958.

Bruner, J.S. The process of education. Cambridge, Mass.: Harvard University Press, 1960.

Bruner, J.S. The act of discovery. Harvard Educational Review, 1961, 31, 21-32.

Bruner, J.S., Goodnow, J., & Austin, G.A. A study of thinking. New York: Wiley, 1956.

Bunderson, C.V. Transfer of mental abilities at different stages of practice in the solution of concept problems. (ETS RB 67-20) Princeton, N.J.: Educational Testing Service, 1967.

Buros, O.K. The fifth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1959.

Cambridge Conference on the Correlation of Science and Mathematics in the Schools. Goals for the correlation of elementary science and mathematics. Boston: Houghton Mifflin Co., 1969.

Cohen, M., & Nagel. E. Logic and the scientific method. New York: Harcourt, Brace, 1934.

Conant, J. Science and common sense. New Haven, Conn.: Yale University Press, 1961.

Cronbach, L.J. The meaning of problems. In J.M. Seidman (Ed.), Readings in educational psychology. Boston: Houghton Mifflin Co., 1955.

Cronbach, L.J., & Meehl, P.E. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 281-302.

Crutchfield, R.S., & Covington, M.V. Programmed instruction and creativity. Programmed Instruction, 1965, 4(1-2), 8-10.

Davis, G.A. Current status of research and theory in human problem solving. Psychological Bulletin, 1966, 66(1), 36-54.

Davis, G.A., Manske, M.E., & Train, A.J. Trial and error versus "insightful" problem solving: Effects of distraction, additional response alternatives, and longer response alternatives, and longer response chains. Journal of Experimental Psychology, 1968, 76, 337-340.

Dewey, J. How we think. Boston: Heath, 1910.

Duncan, C.P. Recent research on human problem solving. Psychological Bulletin, 1959, 56, 397-429.

Duncker, K. On problem solving. Psychological Monographs, 1945, 58, (whole no. 270).

Education Development Center. The USMES guide. Newton, Mass.: May, 1974.

Educational Testing Service. Sequential tests of educational progress -- Science, forms A and B, all levels. Princeton, N.J.: ETS, 1958.

Eimas, P.D. Information processing in problem solving as a function of developmental level and stimulus saliency. Presented at the meetings of the Society for Research in Child Development, Santa Monica, California, March, 1969. [Available from the author, Walter S. Hunter Laboratory of Psychology, Brown University, Providence, Rhode Island 02912.]

Feldhusen, J.F., Houtz, J.C., & Ringenbach, S.E. The Purdue elementary problem-solving inventory. Psychological Reports, 1972, 31, 891-901.

Feldhusen, J.F., Houtz, J.C., Ringenbach, S., & Lash, A. Review of problem solving measurement tasks. Paper presented at the annual meeting of the American Psychological Association, Washington, D.C., September, 1971.

Flavell, J.H. The developmental psychology of Jean Piaget. Princeton, N.J.: Van Nostrand, 1963.

Foster, M., & Martin, M. Probability confirmation and simplicity. New York: Odyssey Press, 1966.

Gagne, R.M. Problem solving and thinking. Annual Review of Psychology, 1959, 10, 147-172.

Gagne, R.M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Gagne, R.M. Problem solving. In A.W. Melton (Ed.), Categories of human learning. New York: Academic Press, 1964.

Gagne, R.M. Learning hierarchies. Educational Psychologist, 1968, 6, 1-9.

Gagne, R.M. The conditions of learning. New York: Holt, Rinehart, & Winston, 1970.

Getzels, J.W. Creative thinking, problem solving, and instruction. The sixty-third yearbook of the national society for the study of education, part I. Chicago: University of Chicago Press,

Glaser, R., Damarin, D.E., & Garner, R.M. The tab item: A technique for the measurement of proficiency in diagnostic problem-solving tasks. Educational and Psychological Measurement, 1954, 14, 283-293.

Goodlad, J.I. School curricular reform in the United States. New York: Fund for the Advancement of Education, 1964.

Guilford, J.P. Intellectual factors in productive thinking. In M.J. Aschner & C. Bish (Eds.), Productive thinking in education. Washington, D.C.: National Education Association, 1965.

Guilford, J.P. The nature of human intelligence. New York: McGraw-Hill, Inc., 1967.

Guilford, J.P., & Hoepfner, R. The analysis of intelligence. New York: McGraw-Hill, 1971.

Gurnee, H. Maze learning in the collective situation. Journal of Psychology, 1937, 3, 437-443.

Gurnee, H. Group learning. Psychological Monographs, 76, 1962, (whole no. 532).

Harootunian, B., & Tate, M.W. The relationship of certain selected variables to problem solving ability. Journal of Educational Psychology, 1960, 51, 331-336.

Hempel, C. Philosophy of natural science. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1966.

Hermanovicz, Henry. Problem solving as a teaching method. Educational Leadership, 1961, 18, 299-306.

Hudgins, B.B. Effects of group experience on individual problem solving, Journal of Educational Psychology, 1960, 51(1), 37-42.

Hudgins, B.B. Problem solving in the classroom. New York: The Macmillan Co., 1966.

Hullfish, H.G., & Smith, P.G. Reflective thinking: The method of education. New York: Dodd Mead & Co., 1961.

Inhelder, B., & Matalon, B. The study of problem solving and thinking. In P. Mussen (Ed.), Handbook of research methods in child development. New York: John Wiley, 1960.

Inhelder, B., & Piaget, J. The growth of logical thinking from childhood to adolescence. New York: Basic Books, 1958.

John, E.R. Contributions to the study of the problem solving process. Psychological Monographs, 1957, 71, (18, whole no. 447).

Johnson, D.M. The psychology of thought and judgment. New York: Harper & Row, 1955.

Johnson, D.M. Solution of anagrams. Psychological Bulletin, 1966, 66, 371-384.

Johnson, D.M. Systematic introduction to the psychology of thinking. New York: Harper & Row, 1972.

Johnson, P. Review of the sequential tests of educational progress. In O.K. Buros, Fifth mental measurements yearbook, 1959.

Katona, G. Organizing and memorizing: Studies in the psychology of learning and teaching. New York: Columbia University Press, 1940.

Keisler, E.R. Teaching children to solve problems: A research goal. Journal of Research and Development in Education, 1969, 3, 3-14.

Kleinmuntz, B. (Ed.) Problem solving: Research, method and theory. New York: John Wiley, 1966.

Klopfer, L.E. Evaluation of learning in science. In B.S. Bloom, J.T. Hastings, & G.F. Madaus, Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971. Pp. 559-641.

Klugman, S.F. Cooperative versus individual efficiency in problem solving. Journal of Educational Psychology. 1944, 35, 91-100.

Lewin, K. Group decision and social change. In E.E. Maccoby, et. al. (Eds.), Readings in social psychology (3rd ed.). New York: Holt, Rinehart, & Winston, 1958.

Lorge, I., Fox, D., Davitz, J., & Brenner, M. A survey of studies contrasting the quality of group performance and individual performance 1920-1957. Psychological Bulletin, 1958, 55, 337-372.

Luchins, A.S. Mechanization in problem solving: The effect of einstellung. Psychological Monographs, 1942, 54, (whole no. 248).

Lundsteen, S.W., & Michael, W.B. Validation of three tests of cognitive style in verbalization for the third and sixth grades. Educational and Psychological Measurement, 1966, 26(2), 449-461.

Maier, N.R.F. Reasoning in humans: III. The mechanisms of equivalent stimuli and reasoning. Journal of Experimental Psychology, 1945, 35, 349-360.

Marshall, J.S., & Burkman, E. Current trends in science education. New York: Center for Applied Research in Education, 1966.

Martin, M. Concepts of science education. Glenview, Ill.: Scott, Foresman, 1972.

McGuire, C.H., & Babbott, D. Simulation technique in the measurement of problem solving skills. Journal of Educational Measurement, 4(1), 1-10.

Merrifield, J.R., Guilford, J.P., Christensen, P.R., & Frick, J.W. The role of intellectual factors in problem solving. Psychological Monographs, 1962, 76, (10, whole no. 529).

Miles, D.T. Development of a test for an experimental research program in creative problem solving. Final Report Southern Illinois University, 1968.

Miller, G.A., Galanter, E. & Pribram, K.H. Plans and the structure of behavior. New York: Holt and Co., 1960.

Mouly, G.J. Psychology for effective teaching. (2nd ed.) New York: Holt, Rinehart, & Winston, 1968.

Nagel, E. The structure of science. New York: Harcourt, Brace, 1961.

National Science Teachers Association. Theory into action in science curriculum development. Washington, D.C.: NSTA, 1964.

National Science Teachers Association. Position statement on curriculum. Washington, D.C.: NSTA, 1963.

National Society for the Study of Education. 59th Yearbook. Part I. Rethinking science education. Chicago: University of Chicago Press, 1960.

Nattress, L.W. The branching-type written simulation exercise. Paper presented at the annual meeting of the American Psychological Association, September, 1970.

Neimark, E.D., & Lewis, N. The development of logical problem-solving strategies. Child Development, 1967, 38, 107-117.

Neimark, E.D., & Lewis, N. Development of logical problem solving: A one-year retest. Child Development, 1968, 39, 515-526.

Newell, A., Shaw, J.C., & Simon, H.A. Elements of a theory of human problem solving. Psychological Review, 1958, 65, 151-166.

Newell, A., Shaw, J.C., & Simon, H.A. The processes of creative thinking. In H.E. Bruber & M. Wertheimer (Eds.), Contemporary approaches to creative thinking. New York: Atherton Press, 1962.

Newell, A., & Simon, H.A. Human problem solving. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972.

Perlmutter, H.V., & de Montmollin, G. Group learning of nonsense syllables. Journal of Abnormal Social Psychology, 1952, 47, 762-769.

Piaget, Jean. Judgement and reasoning in the child. New York: Harcourt, Brace & World, 1948.

Piaget, J. The construction of reality in the child. New York: Basic Books, Inc., 1954.

Piaget, J., & Inhelder, B. The psychology of the child. New York: Basic Books, 1969.

Polya, G. How to solve it. Princeton, N.J.: Princeton University Press, 1948.

Ramirez, J.V. Effects of tutorial experiences on the problem-solving behavior of sixth-graders. California Journal of Educational Research, 1971, 22(2), 80-90.

Ray, W.S. Complex tasks for use in human problem solving research. Psychological Bulletin, 1955, 52, 134-149.

Rimoldi, J.J.A. A technique for the study of problem solving. Educational and Psychological Measurement, 1960, 20, 449-460.

Saarni, C.I. Piagetian operations and field independence as factors in children's problem-solving performance. Child Development, 1973, 44, 338-345.

Schwab, J.J. What do scientists do? Behavioral Science, 1960, 5, 1-27.

Schwab, J.J. The concept of the structure of a discipline. Educational Record, 1962, 43, 197-205.

Schwab, J.J. The teaching of science as enquiry. In J.J. Schwab & P.F. Brandwein, The teaching of science. Cambridge, Mass.: Harvard University Press, 1962. Pp. 1-103.

Schwab, J.J. Problems, topics and issues. In S. Elam (Ed.), Education and the structure of knowledge. Chicago: Rand McNally, 1964. Pp. 4-43.

Schwab, J.J. Structure of the disciplines: Meanings and significances. In G.W. Ford, & L. Pugno (Eds.), The structure of knowledge and the curriculum. Chicago: Rand McNally, 1964. Pp. 6-30.

Schwab, J.J. The practical: A language for curriculum. School Review, 1969, 78, 1-23.

Schwab, J.J., & Brandwein, P.F. The teaching of science. Cambridge, Mass.: Harvard University Press, 1962.

Shann, M.H. An evaluation of unified science and mathematics for elementary schools during the 1973-74 school year. Report to the National Science Foundation, Materials and Instruction Development Section. Boston University, August, 1975.

Shann, M.H., Reali, N.C., Bender, H., et al. Student effects of an interdisciplinary curriculum for real problem solving: The 1974-75 USMES evaluation. Report to the National Science Foundation Materials and Instruction Development Section. Boston University, December, 1975.

Shapiro, B.J. The notebook problem: A report on observations of problem solving activity in USMES and control classrooms. Report prepared for the Unified Science and Mathematics for Elementary Schools Project of the Education Development Center under a grant from the National Science Foundation. Newton, Mass.: EDC, May 1973.

Shulman, L.S., & Tamir, P. Research on teaching in the natural sciences. In R.M.W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Silberman, C.E. Crisis in the classroom. New York: Random House, 1970.

Simon, H.A., & Newell, A. Human problem solving: the state of the theory in 1970. American Psychologist, 1971, 26, 145-159.

Simon, R.I. Encoding effects on complex problem solving. Journal of Experimental Psychology, 1970, 83, 227-231.

Speedie, S.M., Treffinger, D.J., & Feldhusen, J.F. Teaching problem-solving skills: Development of an instructional model based on human abilities related to efficient problem solving. Report to the U.S. Department of Health, Education, and Welfare, Office of Education, Bureau of Research. Purdue University, August, 1973.

Stanley, J. Review of the sequential tests of educational progress. In O.K. Buros, Fifth mental measurements yearbook, 1959.

Stevenson, H.W., Hale, G.A., Klein, R.E., & Miller, L.K. Interrelationships and correlations in children's learning and problem solving. Society for Research in Child Development Monographs, 1968, 33(7), (whole no. 123).

Streufert, S., Kliger, S.C., Castore, C.H., & Driver, M.J. A tactical and negotiations game for the analysis of decision intergration across decision areas. Psychological Reports, 1967, 20, 155-157.

Taylor, D.W., & Faust, W.L. Twenty questions: Efficiency in problem solving as a function of size of groups. Journal of Experimental Psychology, 1952, 44, 360-368.

Torrance, E.P. Torrance tests of creative thinking. Princeton, N.J.: Personnel Press, 1966.

Travers, R.M.W. Essentials of learning. Rev. Ed. New York: The Macmillan Co., 1967.

Treffinger, D.J. Solving problems #1 and #2. Mimeographed test used with fifth and sixth grade students. Purdue University, 1970.

Tyler, R.B. Expectancy for eventual success as a factor in problem solving behavior. Journal of Educational Psychology, 1958, 49, 166-172.

USMES Central Staff. Bicycle transportation: Teacher's resource book. Preliminary ed. Newton, Mass.: Education Development Center, 1974.

USMES News. Newsletter for Unified Science and Mathematics for Elementary Schools. Newton, Mass.: October, 1975, 3(1).

Wallas, G. The art of thought. New York: Harcourt, Brace, and World, 1926.

Weir, M.W. Developmental changes in problem-solving strategies. Psychological Review, 1964, 71(6), 473-490.

Werdelin, I. A synthesis of two factor analyses of problem solving in mathematics. Didakometry, 1966, 8.

Wertheimer, M. Productive thinking. 2nd ed. New York: Harper & Row, 1959.

APPENDIX A

Component Skills of Problem Solving/Decision Making Enumerated by the USMES Developers

- Identifying and defining the problem; being able to distinguish it from related but secondary problems
- Determining important aspects of problem and forming groups to work on these aspects
- Deciding on information and investigations needed and determining priorities
- Deciding upon efficient ways to carry out investigations
- Formulating possible solutions (making hypothese) [capable of being tested]*
- Obtaining information from a variety of sources
- Distinguishing facts from opinions, relevant from irrelevant information, and reliable from unreliable sources
- Detecting simple errors; identifying unsupported assumptions or generalizations
- Deciding upon the best manner to represent data
- Using data and graphs to test hypotheses and draw inferences
- Evaluating procedures used for data collection and analysis
- Determining the best way to collect survey and measurement data

*Brackets indicate a skill to be included when appropriate for a particular unit.

APPENDIX A (Cont'd)

- Considering practicality of suggested solutions
- Considering that a problem may have different solutions depending on the values applied
- [Deciding upon the most effective way of presenting proposals to authorities]*
- Utilizing different methods of group decision making
- Trying out various suggestions and evaluating the results
- Applying process learned to other real problems
- Deciding on generalizations that might hold true under similar conditions
- Making suitable simple mathematical models of real situations and refining them

*Brackets indicate a skill to be included when appropriate for a particular unit.

APPENDIX B

Affective Goals Enumerated by the USMES Developers

- Appreciating the importance of the many facets of problem solving
- Developing self-reliance, curiosity and initiative
- Making value judgements
  - Recognizing differences in values according to age, experience, occupation, income and interests (culture, race, religion, ethnic background)
  - Recognizing that facts alone do not determine decisions, that problematic situations have no set answers
  - Recognizing core values of daily living: fair play and justice, free speech, opportunity for decision making, opportunity for self-respect, choice, right to privacy, acceptance of the life styles of the community, group identity
- Accepting responsibility for work being done
- Participating in decision making relevant to their lives
- Learning to work cooperatively in large and small groups; recognizing the values of cooperation among individuals, group work and division of labor
- Respecting the views, thoughts and feelings of others

APPENDIX C

Cognitive Goals for USMES: Component Skills of the Problem Solving Process to be Mastered by Each USMES Child

- The child will identify and define the problem.
- The child will decide what information and investigations are needed in order to find some solution to the problem.
- The child will determine what needs to be done first.
- The child will decide what is the best way to obtain the information needed.
- The child will detect flaws in data gathering procedures or errors in the data itself.
- The child will organize, analyze, and interpret the data.
- The child will suggest some solution to the problem based on the data.
- The child will suggest ways to implement the solution.

APPENDIX D

List of Science Tests Reviewed for Possible Use in the USMES Evaluation

American Association for the Advancement of Science. Science -- A process approach. Competency measures, part five. (AAAS misc. pub. 66-23) Washington, D.C.: AAAS, 1966.

American Association for the Advancement of Science. Science -- A process approach. Competency measures, part six. (AAAS misc. pub. 66-26) Washington, D.C.: AAAS, 1966.

American Association for the Advancement of Science. Science -- A process approach. Competency measures, part seven. (AAAS misc. pub. 66-29) Washington, D.C.: AAAS, 1966.

Biological Sciences Curriculum Study. BSCS Processes of Science Test, Form A. Boulder, Colo.: University of Colorado, 1965.

Biological Sciences Curriculum Study. Final Examination, Form J. (Rev.) Boulder, Colo.: University of Colorado, 1965.

Borman, I.M., & Sanders, M.W. Borman-Sanders Elementary Science Test, grades 5-8. Emporia, Kans.: Kansas State Teachers C College, 1964.

Boyer, P.A., & Gordon, H.C. General Science Objective Test. Chicago: Lyons & Carnahan, 1959.

Burmester, M.A. A test of aspects of scientific thinking. East Lansing, Mich.: Michigan State University, 1951. (Also available from the author, Dept. of Natural Science, Michigan State University)

Carrier, E.O., Geis, F., Klopfer, L.E., & Shoresman, P.B. Test On Understanding Science, Form Ew. Urbana, Ill.: Authors, 1966. (Available from P.B. Shoresman, School of Education, University of Illinois, Urbana, Ill.)

Chemical Education Material Study (CHEMS). CHEM Study achievement tests. Berkeley, Calif.: University of California, 1963.

Cooley, W.W., & Klopfer, L.E. Test on Understanding Science, Form W. Princeton, N.J.: Educational Testing Service, 1961.

Cross, D., & Sanders, M.W. Emporia General Science Test. Emporia, Kans.: Kansas State Teachers College, 1964.

Dressel, P.L., & Nelson, C.H. Questions and problems in science: Test items folio no. 1. Princton, N.J.: Educational Testing Service, 1956.

Educational Testing Service. Cooperative Biology Test. Form Y. Princeton, N.J.: ETS, 1948.

Educational Testing Service. Cooperative Science Tests, Advanced General Science, Form B. Princeton, N.J.: ETS, 1962.

Educational Testing Service. Cooperative Science Tests, General Science, Form B. Princeton, N.J.: ETS, 1962.

Educational Testing Service. Cooperative Science Tests, Chemistry, Form B. Princeton, N.J.: ETS, 1963.

Educational Testing Service. Cooperative Science Tests, Physics, Form B. Princeton, N.J.: ETS, 1963.

Educational Testing Service. Sequential Tests of Educational Progress -- Science, Forms A and B, all levels. Princeton, N.J.: ETS, 1958.

Klopfer, L.E. Word Association Study, Form Ez. Urbana, Ill.: Elementary-school Science Project, University of Illinois, February, 1964.

Kruglag, H., & Wall, C.N. Laboratory performance tests for general physics. Kalamazoo, Mich.: Western Michigan University, 1959.

Learning Research and Development Center, University of Pittsburgh. Individually Prescribed Instruction Program: Science, Level A, Inventory of Prerequisite Skills. Unpublished test, University of Pittsburgh, 1969.

Learning Research and Development Center, University of Pittsburgh. Individually Prescribed Instruction Program, Level A Placement Test, Part II. Unpublished test, University of Pittsburgh, 1969.

Welch, W.W. Welch Science Process Inventory, Form D. Cambridge. Mass.: Author, 1966. (Available from the author, 330 Burton Hall, University of Minnesota, Minneapolis, Minn. 55455)

APPENDIX E

ADMINISTRATOR'S MANUAL
for
THE PLAYGROUND PROBLEM

A Measure of Problem Solving Ability for
Use in the Evaluation of USMES

Prepared by

The USMES Evaluation Staff
Boston University

Mary H. Shann, Ph.D.
USMES Evaluation Project Director

TO THE OBSERVER:

This Manual and the accompanying materials consist of the following:

1. Instructions to guide you in the administration of the Playground Problem
2. A catalog of playground equipment
3. A form on which to record your observations of the children's behaviors
4. A cassette tape for recording various segments of the sessions.

GENERAL INSTRUCTIONS

The problem solving behaviors of elementary school children constitute one of the most important areas for evaluation of the USMES program. The Playground Problem is to be used as one means of assessing the success of the USMES program in reaching its goals. This test is designed to enable the observer to collect data on both verbal and non-verbal behaviors involved in problem solving.

The Playground Problem should be administered to designated USMES classes and control classes. Five children are to be selected randomly from each USMES class and similarly from each control class in the evaluation sample. The test is to be given to each group of five children rather than to individuals.

Each group of children should be taken to an open area near the school and asked to plan a playground. The materials the children are to use in solving the problem, the instructions you are to give them, and the role you are to play as an observer will be explained in detail shortly.

We are interested in assessing the degree of cooperation and self- or group-motivated interest the children demonstrate during the entire problem solving period and the follow-up question period. We are equally interested in the degree to which the children employ practical considerations in solving the problem.

Our analysis of the Playground Problem test results will be based on three kinds or records: (a) a tape recording of the children's verbal presentation during the follow-up question period; (b) your observations of the children's behaviors as recorded on the observation form accompanying this Manual; and (c) a layout of the proposed playground which the children will be asked to draw on a large sheet of paper.

In general, your role as an observer will be to organize the test session, to instruct the children on what to do, and to observe and record their behavior. Specific instructions for administration of the Playground Problem are given in the following sections of this Manual.

ORGANIZATION

1. Selection of Children

A random sample of five children should be picked from each control class and each USMES class in your school. In the past, children have not always been picked randomly, and this is not acceptable. When children are picked on the basis of good academic performance on the one hand, or on the basis of "getting rid of the troublemaker" on the other, the entire session will have to be disregarded.

It would be best for you to pick the children yourself, but the teacher can also make the selections if correct procedures are used. The easiest appropriate method is to write the names of each child on a piece of paper, throw each piece in a hat, and then select five.

2. When to Administer the Playground Problem Test

This can be a critical factor. Oftentimes, children are more restless and less attentive at certain times of the day, and especially at certain times of the year--for example, the day before Christman Vacation.

Try to run your test sessions at approximately the same time of day--that includes the control classes as well as the USMES classes. The recommended time of day is as close to the beginning of the day as possible. Avoid extremely cold or rainy days, since the Playground Problem is to be administered outside.

Second, do not run your test sessions on the day before or after vacation periods, or on the days when special school events are to take place. In the past, some sessions have had to be discounted because of confounding factors of this nature.

In all of these considerations, use your own good judgement. A test administered under somewhat less than ideal conditions is probably better than no test returns at all for a class.

3. Where to Administer the Playground Problem Test

In preparation for the test, you should locate a suitable open area near the school. An empty lot would be ideal. However, if one is not available, a playing field or clear black topped area would be appropriate. This area should be the same for all groups of children in the same schools on your sample list, both USMES groups and control groups.

4. Materials to Accompany Test Administration

Prior to the testing session, you will need to gather together the following items:

Observation Equipment

Observation form

Tape recorder and blank cassette

Watch

Tools (in a cardboard box)

50 foot tape measure
Yard stick
Ball of string
Large piece of paper
Tri-wall (to use as hard surface for drawing plan)
Felt tip pens
Pencils
12" rulers
Catalog of playground equipment
Scrap paper
Scissors

INSTRUCTIONS TO THE CHILDREN

Soon after arriving at the open area, you should give the children the following instructions and you should record them on tape:

> "Let's suppose this area was going to be made into a new playground for the children in your school." (Indicate clearly the limits of the area). "How would you plan this playground?"
>
> "Here is a catalog of playground equipment which could be bought. If you had $2,000 to spend, which equipment would you choose?"
>
> "Please work together to decide which equipment should be bought. Draw a plan of the playground on this piece of paper showing where the equipment would be placed."
>
> "You have forty minutes to work together to make your plan. Here are some things you may use if you want to." (Hand one child the box containing the tape measure, pencils, etc.) "Remember, you can spend up to $2,000 on equipment."

DO NOT GIVE THE CHILDREN ANY SUGGESTIONS AS TO WHAT OTHER CONSIDERATIONS THEY SHOULD KEEP IN MIND. In the past, some test results have had to be invalidated because of suggestions and clues which observers had given to the children in the instructions. The instructions should be as similar as possible for the USMES groups and for the control groups. Any evidence of intentional or unintentional bias unfortunately results in invalidation of the test session.

Let the children know that they will have forty minutes to figure out their plan and draw it on paper. Tell them that at the end of this period, you will ask them questions about their plan, and that their answers will be recorded on tape (more about taping later).

OBSERVATIONS

During the forty minute problem solving period, stay in the area in view of the children. You can repeat the instructions, if necessary. However, you should not participate in the problem solution by answering other questions or

suggesting possible strategies. It is up to the children to decide whether or not to use the measuring equipment. Do not demand that any particular child help out in planning the playground if he or she does not want to.

After thirty minutes of the problem solving period have expired, tell the children that they have ten minutes to complete drawing their plan if they have not already done so.

During the forty minute problem solving period, the observer should make notes on the observation form describing the children's activities. Please write clearly. Each activity should be noted under the appropriate category heading. These notes should be specific and numbered sequentially. For example, under the heading "Measuring" the observer might note:

"5. Two kids measured the width of the lot with the 50' tape." The number "5" indicates that this is the fifth note the observer has made on the observation form. The next note might be:

"6. One child recorded the width of the lot as 45 feet." This observation would be placed under the heading "Recording Data."

You will have received intensive training in the use of this observation form at the Observers' Training Workshop.

PREPARATION FOR TAPING

After the forty minute problem solving period is completed, you should call the children together to prepare for tape recording the ten minute question period.

Children are often shy or giggly when they first speak into a microphone. Inaudiable responses make our work of analysis very difficult. To get around this problem, please ask each child to recite a sentence into the microphone. such as: "This is our plan," or "My name is ..." Tell the children that they must speak one at a time, and ask them to speak slowly and clearly.

Play the tape back to the children. This will give them some chance to get used to recording their voices, and it will give you a chance to see how well their voices are being picked up. (Note: this part of the recording is not important to us and can be erased).

When the entire session is over, we would like to have the following recordings returned to us:

Part 1: the instructions as you gave them originally to the children

Part 2: the ten minute question period given after the thirty minute problem solving period and after the practice taping.

QUESTION PERIOD

This period during which the children explain their plan and outline their reasoning should be tape recorded in its entirety. The children's presentation may be up to ten minutes long. You should record the data and group at the beginning of each question period taping. If you wish, you may take the children back into the school to make the recording.

It is very important to remember that the questions you ask the children and the procedures you use in soliciting their answers MUST be as similar as possible for the USMES groups and for the control groups. Again, any evidence of bias may invalidate the results.

Although you may have to use your imagination and various strategies to encourage the children to respond or to explain what they mean in greater detail, use the following "script" as a guide to the specific questions you should ask. It is very helpful, we are sure you know, if you show interest and enthusiasm in what the children have done. Remind the children to speak slowly and clearly so that other people can understand what they have said later. Do not rush the children but rather gently encourage them to say what they want.

FIRST QUESTION SERIES (Directed to the entire group*)

-- "How did you do?"

-- "Was it fun?"

SECOND QUESTION SERIES (Directed to the entire group*)

-- "Explain your playground plan."

-- "Why did you decide to but (4) pieces of equipment?"

-- "Do you know how much the equipment you have chosen will cost?"

-- "Why did you decide to put the swings over here? The slide over here?"

-- "What kinds of information did you need to help you make your decisions?"

THIRD QUESTION SERIES (Directed first to the entire group, and then to each child in turn who has not yet responded)

-- "Were there any other important factors you had to consider in making your decisions?"

-- "Is there anything anyone would like to say before we finish?"

While it may be necessary to structure the children's report by asking questions, you as the observer should not suggest rationale to the children by means of your questioning. For example, if there has been no mention of safety factors or indications that the issue of safety has been taken into consideration, the observer should not bring it up during the tape recording.

The playground problem does not have one solution. However, in the playground problem, a certain approach to problem solving is valued. An excellent response to the playground problem would include:

1. Measurement or calculation of available space.
2. Meaningful use of measuring equipment
3. Careful consideration of types of playground equipment chosen.
4. Comparisons between size of equipment as listed in catalog and space available on playground area.
5. Consideration of budget limitations.
6. Accuracy in drawing lay-out of proposed playground.
7. Consideration of human elements such as safety and aesthetic appeal.
8. Logical and clear presentation of rationale.

- -

* When the question is directed to the entire group make sure that everyone talks who wants to, not only the "spokesman" for the group. Be sure they talk one at a time so that it is easy to understand what is being said.

However, particularly on the pre-test, the children may not respond in this manner. This in itself is interesting and important data and should not be interpreted as resulting from the format of the problem.

After the testing session is over, review the tape on your own. If you think any part of the conversation will be difficult for us to understand, please make a note of what was said and attach it to the observation form. Please be sure to return to us all tapings, observation sheets, scrap papers the students wrote on, and the playground layouts. The pre-test results should be sent to us soon after they have been completed. The Playground Manual and Catalog should be retained by you after administration of the pre-tests. They should be used again for administration of the posttests. Upon completion of the post-tests, please return to us the Manual and Catalog along with the testing results for the post-test.

Instructions for administration of this Playground Problem will have been reviewed in detail at your Observers' Training Workshop. However, if you have any further questions when you are ready to administer the test, please call the USMES Evaluation Team, collect, at (617) 353-3312.

Dr. Mary H. Shann

APPENDIX F

Scoring Protocol for the Playground Problem: A Manual for Rating and Coding Students' Performance On a Test of Complex Problem Solving

Prepared by

Mary H. Shann, Ph.D.

USMES Evaluation Project Director

Boston University

1974

Section I.--IDENTIFICATION (Columns 1-20)

I.D. code records the teacher grade level, unit and other descriptive information related to reliability and validity issues.

Column 1: identifies form of the problem-solving test.

6 = Playground

7 = Picnic

Column 2: identifies time of testing.

1 = Pre-test

2 = Post-test

Column 3: identifies treatment.

1 = USMES

2 = Control

Columns 4,5: identify teacher.

(See master list for teacher codes)

Column 6,7,8: identify grade level.

(See master list for grade level codes)

In columns 9 and 10 enter the unit code as follows:

Advertising	01
Bicycle Transportation	02
Burglar Alarm Design (now called Protecting Property), (may also be called Security by some teachers)	03
Classroom Design	04
Classroom Management	05

Community Gardening	06
Consumer Research	07
Describing People	08
Designing for Human Proportions	09
Design Lab Design	10
Dice Design	11
Eating in School	12
Getting in Shape	13
Getting There (formerly Finding Your Way, Getting From Place to Place)	14
Growing Plants	15
Lunch Lines	16
Making School Safer	17
Manufacturing	18
Mass Communications (formerly Mass Media)	19
Nature Trails	20
Orientation (formerly Student Migration)	21
Pedestrian Crossings	22
Planning Special Occasions	23
Play Area Design and Use	24
School Rules (formerly School Rules and Decision Making)	25
School Supplies (formerly Managing and Conserving School Resources), (or Recycling)	26
School Zoo (formerly Outgrowth of Animal Behavior, and Ecosystems which are no longer units)	27
Soft Drink Design	28
Sound in the Environment (formerly Outgrowth of Music which is no longer a separate unit)	29

Traffic Flow	30
Using Free Time (formerly Designing Indoor/Outdoor Games)	31
Using Free Time After School (After School Activities)	32
Ways to Learn	33
Weather Predictions	34

										Blank
1	2	2	4	5	6	7	8	9	10	11

Column 11: Leave Blank

Based on your review of the audio tape and observer's notes, indicate whether you think any of the following factors may render this testing session invalid. Code your response 0 = No, 1 = Yes in the appropriate column.

Problem	Column
Biased selection of students	12
Prompting by observer	13
Prior student experience with this test	14
Inclement weather (0 for picnic problem)	15
Noisy testing environment	16
Outside interference/interruptions	17
Observer deviated from standard procedure	18
Blank	19-20

Section II.--BEHAVIORAL ASPECTS (Columns 21-24)

There are four factors which are considered in this segment. The scoring of this group shall proceed as follows:

Factor: 1

Motivation: to accept the problem and attempt to solve the problem.

Scoring:
- 0 No one accepts problem or trys to solve problem.
- 1 1 Student accepts/trys to solve problem.
- 2 2 Students accepts/trys to solve problem.
- 3 3 Students accepts/trys to solve problem.
- 4 4 Students accepts/trys to solve problem.
- 5 5 Students accepts/trys to solve problem.

Enter the proper score in column 21.

Factor: 2

Commitment to task: the level of itensity of the group to continue working toward a solution.

Scoring:
- 0 No effort.
- 1 Disinterested, fooling around, little input.
- 2 Some positive input (one or two interested in problem and working with little progress).
- 3 Group is interested but efforts are not organized, and time is being wasted.
- 4 Group is interested, working and not wasting time or effort.

Enter proper score in column 22.

Factor: 3

Organization: allocation of responsibilities for efficiency of manpower.

Scoring:
- 0 No effort.
- 1 Unplanned, haphazard, or chaotic (students do their own thing - do not allocate item or all work on the same thing).
- 2 No all students involved (either by choice or flat). Some are working on problem some are not - may be arguing among each other.

3 Students have allocated some tasks - may have some working on same item; or possibly 1 may not be involved.

4 Tasks are allocated and students working efficiently-however students may have trouble with their item and seek help.

5 Tasks allocated and all are working productively.

Enter proper score in column 23.

Factor: 4

Structure: Group leadership

Scoring: 0 None

1 Autocratic--one person dominates who does not listen to other students' ideas.

2 Minority Leadership--one or two persons listen to others and then lead or direct.

3 Plurality--general agreement of several members leads to direction and leadership; most contributions are recognized and evaluated.

4 Democratic--all students contribute; no one's suggestions are ignored or ridiculed. One spokesman may arise but sources of ideas/efforts are recognized.

Enter proper score in column 24.

21	22	23	24

Section III.--COGNITIVE ASPECTS (Columns 25-56)

Data for this section can be derived primarily from the observer form and the tapes. It will be necessary to read the observer form and listen to the tapes to bridge any apparent gaps or vague statements found in either the form or the tape.

The cognitive aspects shall include variables considered in solving the problem and the level or method of measuring the variables. The implementation of the measurement in terms of calculation and the recording of the data will be collected and encoded.

A total of 10 variables can be accommodated by the scoring protocol. For each variable, its identification, measurement, calculation and recording will be scored.

IIIA. Factor: COST OF EQUIPMENT

Identification:

Scoring: 0 No

1 Yes

Enter in column 25.

Measurement:

Scoring 0 No measurement done.

1 Vague or very general estimates.

2 Estimations by imprecise methods or by eyeballing. It does not provide enough information to arrive at a decision.

3 Useful information which can be used to arrive at a decision but the data should be more accurate or precise.

4 Precise measurement or clearly appropriate data that can lead to solution.

Enter in column 26.

Calculations:

Scoring: 0 No calculations.

1 Vague or very general calculations that do little quantification.

2 Calculations are imprecise or guesses are arrived at by trial and error and are not sufficient to provide necessary data to arrive at a solution.

3 Useful calculations which can be used to arrive at a solution. It may not be accurate or have considered totals or balances. It should be more precise.

4 Calculations are appropriate, precise and can lead to a solution.

Enter in column 27.

Recording:

Scoring 0 No records.

1 Very general or imprecise records.

2 Adequate records.

Enter in column 28.

25	26	27	28

IIIB. Factor: SIZE OF EQUIPMENT VS. SIZE OF CHILDREN (i.e., larger scale equipment for older children; smaller scale equipment for younger children)

Identification:

Scoring: 0 No

1 Yes

Enter in column 29.

Measurement:

Scoring: 0 No measurement.

1 Vague or general estimates, i.e., big equipment for big kids.

2 Express need to know proportion of big and small kids in their school.

Enter in column 30.

Calculations:

Scoring: 0 No calculations.

1 General or arbitrary assignment of equipment for size of children i.e., for example "lets get half big equipment; half small."

2 More careful estimates on how many big and small kids attend their school and selections of equipment reflects distribution of size of students.

Enter in column 31.

Recording:

Scoring: 0 No records.

1 Very general or imprecise records.

Enter in column 32.

29	30	31	32

IIIC. Factor: SIZE OF EQUIPMENT VS. AREA AVAILABLE (e.g., a swing will use 100 sq. feet and we have 1000 sq. feet all together to use.)

Identification:

Scoring: 0 No

1 Yes

Enter in column 33.

Measurement:

Scoring: 0 No measurement done.

1 Vague or very general estimates.

2 Estimations by imprecise methods or by eyeballing. It does does not provide enough information to arrive at a decision.

3 Useful information which can be used to arrive at a decision but the data should be more accurate or precise.

4 Precise measurement or clearly appropriate data that can lead to solution.

Enter in column 34.

Calculations:

Scoring: 0 No calculations.

1 Vague or very general calculations that do little quantification.

2 Calculations are imprecise or guesses are arrived at by trial and error and are not sufficient to provide necessary data to arrive at a solution.

3 Useful calculations which can be used to arrive at a solution. It may not be accurate or have considered totals or balances. It should be more precise.

4 Calculations are appropriate, precise and can lead to a solution.

Enter in column 35.

Recording:

Scoring: 0 No records.

1 Very general or imprecise records.

2 Adequate records.

Enter in column 36.

33	34	35	36

IIID. Factor: CAPACITY OF EQUIPMENT
(e.g., 4 kids can use a swing set with four seats; more kids can use a big jungle jim.)

Identification:

Scoring: 0 No

1 Yes

Enter in column 37.

Measurement:

Scoring: 0 No measurement.

1 Vague or general estimates; i.e., big stuff can be used by more kids.

2 Express need to know specific number of children who can use each piece of equipment at one time.

Enter in column 38.

Calculations:

Scoring: 0 No calculation.

1 General estimates of capacity (e.g., most of the kids in a class could use something at the same time).

2 Precise figures on capacity (e.g., altogether, the equipment we choose will handle 25 kids at one time).

Enter in column 39.

Recording:

Scoring: 0 No records.

1 Very general or imprecise records.

2 Adequate records.

Enter in column 40.

37	38	39	40

IIIE. Factor: DURABILITY OF EQUIPMENT
(i.e., stronger, lasts longer)

Identification:

Scoring: 0 No

1 Yes

Enter in column 41.

Measurement:

Scoring: 0 No measurement.

1 Vague statements, i.e., its better.

2 General/precise, i.e., stronger, lasts longer.

Enter in column 42.

Calculations:

Scoring: 0 No calculations.

1 Calculations in a general or vague sense.

Enter in column 43.

Recording:

Enter 0 in column 44.

41	42	43	44

IIIF. Factor: PLACEMENT OF EQUIPMENT FOR SAFETY CONSIDERATIONS

Identification:

Scoring: 0 No

1 Yes

Enter in column 45.

Measurement:

Scoring: 0 No measurement.

1 General or vague statements of more or less safety.

2 More precise measures of safety, i.e., more distance so kids do not run into the other stuff.

Enter in column 46.

Calculations:

Scoring: 0 No calculations.

1 Vague as to placement, i.e., that close enough.

2 Some concept of calculation, i.e., about 6 ft. or the like.

Enter in column 47.

Recording:

0 No records.

1 Records.

Enter in column 48.

45	46	47	48

111C. Factor: PLACEMENT OF EQUIPMENT FOR EFFICIENT UTILIZATION OF AREA

Identification:

Scoring: 0 No

1 Yes

Enter in column 49.

Measurement:

Scoring: 0 No measurement.

1 Vague or general statements, i.e., it fits.

2 More precise statements of placement based on size or shape of equipment or terrain.

Enter in column 50.

Calculations:

Scoring: 0 No calculations.

1 General or vague calculation based on placement and practical considerations, e.g., putting it there leaves us with more space for playing ball.

Enter in column 51.

Recording:

Scoring: 0 No records.

1 Very general or vague records.

Enter in column 52.

49	50	51	52

IIIH. Factor: OTHER CONSIDERATIONS

Column:

53 Number of additional factors mentioned.

54 "Fun" mentioned as consideration (0=no, 1=yes).

55 "Appeal of equipment for all ages" mentioned as consideration (0=no, 1=yes).

56 Blank

Enter in column 53-56.

53	54	55	56

Section IV.--PRODUCT ASPECTS (Columns 57-60)

Evaluation of four product aspects shall be based on the students' drawing of their playground design.

The Product - Plan

Scale:

Scoring:
- 0 No scale.
- 1 Approximate scale that indicated relative size of equipment; representations of distances are reasonable.
- 2 Scale is precise or is coded.

Enter in Column 57.

Labels:

Scoring:
- 0 No labels.
- 1 Labels are present and appropriate to equipment.

Enter in Column 58.

Landmarks:

Scoring:
- 0 No landmarks.
- 1 Landmarks are present.
- 2 Landmarks are present, appropriate and/or coded, i.e., enduring and relevant to playground area.

Enter in Column 59.

Area:

Scoring:
- 0 No area limitations.
- 1 Area is defined.

Enter in Column 60.

57	58	59	60

APPENDIX G

ADMINISTRATOR'S MANUAL
for
THE PICNIC PROBLEM

A Measure of Problem Solving Ability for
Use in the Evaluation of USMES

Prepared by

The USMES Evaluation Staff
Boston University

Mary H. Shann, Ph.D.

USMES Evaluation Project Director

TO THE OBSERVER:

This Manual and the accompanying materials consist of the following:

1. General instructions to help guide you in the implementation of the picnic problem.

2. Observation sheets upon which all of your observations and notes should be made.

3. Park Map and Photograph of Picnic Foods for use by the children.

4. Cassette Tape for recording various segments of the session.

GENERAL INSTRUCTIONS

The problem solving behaviors of elementary school children constitute one of the most important areas for evaluation of the USMES program. The Picnic Problem is to be used as one means of assessing the success of the USMES program in reaching its goals. This test is designed to enable the observer to collect data on both verbal and non-verbal behaviors involved in problem solving.

The Picnic Problem should be administered to designated USMES classes and control classes. Five children are to be selected randomly from each USMES class and similarly from each control class in the evaluation sample. The test is to be given to each group of five children rather than to individuals.

Each group of children should be brought to a separate room if possible, or some other quiet location, where they are to be given a common problem to be solved, in this case, the Picnic Problem. The materials the children are to use in solving the problem, the instructions you are to give them, and the role you are to play as an observer will be explained in detail shortly.

We are interested in assessing the degree of co-operation and self or group-motivated interest the children demonstrated during the entire problem-solving period and the follow-up question period. We are equally interested in the degree to which the children employ practical considerations in solving the problem.

Our analysis of the Picnic Problem Test results will be based on three kinds of records: (a) a tape recording of the children's verbal presentation during the follow-up question period; (b) your observations of the children's behaviors as recorded on the observation form accompanying this Manual; and (c) the pieces of scrap paper on which the children recorded measurements or made calculations.

Your role as an observer will be to organize the test session, to instruct the children on what to do, and to observe and record their behavior. Specific instructions for administration of the Picnic Problem are given in the following sections of this Manual.

ORGANIZATION

1. Selection of Children

A random sample of five children should be picked from each control class and each USMES class in your school. In the past, children have not always been picked randomly, and this is not acceptable. When children are picked on the basis of good academic performance on the one hand, or on the basis of "getting rid of the troublemaker" on the other, the entire session will have to be disregarded.

It would be best for you to pick the children yourself, but the teacher can also make the selections if the correct procedures are used. The easiest appropriate method is to write the names of each child of a piece of paper, throw each piece in a hat, and then select five.

2. When to administer the Picnic Problem

This can be a critical factor. Oftentimes, children are more restless and less attentive at certain times of the day, and especially at different times of the year--for example, the day before Christmas vacation.

Try to run your test sessions at approximately the same time of day--that includes the control classes as well as the USMES classes. The recommended time of day is as close to the beginning of the day as possible.

Secondly do not run your sessions on the day before or after vacation periods, or on the days when special school events are to take place. In the

past some sessions have had to be discounted bacause of confounding factors of this nature.

In all of these considerations, use your own good judgement. A test administered under somewhat less than ideal conditions is probably better then no test returns at all for a class.

3. Where to administer the Picnic Problem

The instructions given to the children, the actual problem solving period and the follow-up question period should all take place in the same area and it should be the same area for all groups of children (i.e. both USMES and control groups).

The ideal location for the sessions would be a quiet room where there is minimal possibility for distractions.

4. Materials to Accompany Test Administration

Prior to the testing session, you will need to gather together the following items:

Observation Equipment

Observation form
Tape recorder and blank cassette
Watch

Tools (In a cardboard box)

Yard stick
12" rulers
Ball of string
Scissors
Pencils
Scrap paper
50 foot tape measure

Other Materials

Map of parks
Photograph of food

INSTRUCTIONS TO THE CHILDREN

After the children are in the test area and you have their attention, you should give the children the following instructions and you should record your instructions on tape.

"You have been picked to take part in a game to see how well you can plan a picnic, as a group." (Try to get the children's enthusiasm and interest by asking them a few questions about their own experiences, if they went on any picnics last summer, etc).

"Let's suppose that you are asked to plan a picnic for 25 children and that you will have $50 to spend."

"Let's suppose that none of the parks allows Bar-B-Qing, but that you can order food for your picnic from a food service which has stands at the picnic areas in each park." (Point out the picnic areas on the map). "You must place your order 2 (two) days ahead of time so that they will have enough food on hand."

"Here is a picture showing the foods you may order and the price of each item: Hamburgers are 50¢ each; hotdogs are 30¢ each, soda is 20¢ a can; potato chips are 10¢ a bag; and ice cream cones or ice cream sandwiches are 20¢ each."

"This map shows the areas you can choose for the picnic. Each park charges admission." (Review the map of the picnic areas with the children. Point out the admission charges per person for each park, and explain the various symbols on the map). For example, "This symbol indicates a playground, and here are the playgrounds in each park." (Do likewise for all the other symbols). "Notice that the map is drawn to scale, and 1" on the map equals 10 miles."

"Your transportation will be provided via school bus free of charge. You may spend from 10:00a.m. to 4:00 p.m., from the time you must board the bus until the time you must be back at the school."

"Please work together to decide where you would choose to go for this picnic, and what foods you would buy."

"You have forty minutes to work together to make your plan. Here are some things you may use if you want. (Hand one child the box containing the rulers, pencils, etc.) "Remember, you can spend up to $50 and that your time is from 10:00a.m. to 4:00p.m. including time spent traveling in the bus."

DO NOT GIVE THE CHILDREN ANY SUGGESTIONS AS TO WHAT OTHER CONSIDERATIONS THEY SHOULD KEEP IN MIND. In prior years, some of the test results had to be invalidated because of suggestions or clues given to the children. Any evidence of intentional or unintentional bias unfortunately results in invalidation of the test session. The instructions should be as similar as possible for

USMES groups and for control groups.

Let the children know that they will have forty minutes to figure out their plan. Tell them that at the end of this period, you will ask them questions about their plan, and that their answers will be recorded on tape (more about taping later).

OBSERVATIONS

During the forty minute problem solving period, stay in the area in view of the children. You can repeat the instructions, if necessary. However, you should not participate in the problem solution by answering other questions or suggesting possible strategies. It is up to the children to decide whether or not to use the measuring equipment. Do not demand that any particular child help out in planning the picnic if he or she does not want to.

After thirty minutes of the problem solving period have expired, tell the children that they have ten minutes to complete their plan if they have not already done so.

During the forty minute problem solving period, the observer should make notes on the observation form describing the children's activities. Please write clearly. Each activity should be noted under the appropriate category heading. These notes should be specific and numbered sequentially. For example, under the heading "Measuring" the observer might note:

"4. Two kids measured the distance to each park with string.'

"5. Two kids measured the string distances against a ruler."

The numbers "4" and "5" indicate that these are the fourth and fifth notes the observer has made on the observation form. The next note might be:

"6. One child converted string lengths to distances in miles."

This observation would be placed under the heading "Calculating."

You will have received intensive training in the use of the observation form for the Picnic Problem at the Observers' Training Workshop.

PREPARATION FOR TAPING

After the forty minute problem solving period is completed, you should call the children together to prepare for tape recording the ten minute question period.

Children are often shy or giggly when they first speak into a microphone. Inaudible responses make our work of analysis very difficult. To get around this problem, please ask each child to recite a sentence into the microphone, such as: "This is our plan," or "My name is..." Tell the children that they must speak one at a time, and ask them to speak slowly and clearly.

Play the tape back to the children. This will give them some chance to get used to recording their voices, and it will give you a chance to see how well their voices are being picked up. (Note: this part of the recording is not important to us and can be erased).

When the entire session is over, we would like to have the following recordings returned to us:

Part 1: the instructions as you gave them originally to the children.

Part 2: the ten minute question period given after the forty minute problem solving period and after the practice taping.

QUESTION PERIOD

This period during which the children explain their plan and outline their reasoning should be tape recorded in its entirety. The children's presentation may be up to ten minutes long. You should record the date and the group at the beginning of each question period taping.

It is very important to remember that the questions you ask the children and the procedures you use in soliciting their answers MUST be as similar as possible for the USMES groups and for the control groups. Again, any evidence of bias may invalidate the results.

Although you may have to use your imagination and various strategies to encourage the children to respond, or to explain what they mean in greater detail, use the following "script" as a guide to the specific questions you should ask. It is very helpful, we are sure you know, if you show interest and enthusiasm in what the children have done. Remind the children to speak slowly and clearly so that other people can understand what they have said later. Do not rush the children but rather gently encourage them to say what they want.

FIRST QUESTION SERIES (Directed to the entire group*)

-- "How did you do?"

-- "Was it fun?"

SECOND QUESTION SERIES (Directed to the entire group*)

-- "Explain your plans for the picnic."

-- "Which park did you choose? Why?"

-- "Which foods did you choose? Why?"

-- "Do you know how much the picnic will cost?"

-- "What kinds of information did you need to help you make your decisions?"

THIRD QUESTION SERIES (Directed first to the entire group, and then to each child in turn who has not yet responded)

-- "Were there any other important factors you had to consider in making your decisions?"

-- "Is there anything anyone would like to say before we finish?"

While it may be necessary to structure the children's report by asking questions, you as the observer should not suggest rationale to the children by means of your questioning. For example, if there has been no mention of distance factors or indications that the traveling time has been taken into consideration, the observer should not bring it up during the tape recording.

---

* When the question is directed to the entire group make sure that everyone talks who wants to, not only the "spokesman" for the group. Be sure they talk one at a time so that it is easy to understand what is being said.

The Picnic Problem does not have one solution. However, in the Picnic Problem, a certain approach to problem solving is valued. An excellent response to the Picnic Problem would include:

1. Measurement and calculation of the distances to each park.
2. Meaningful, efficient use of measuring equipment.
3. Careful consideration of the advantages of each park.
4. Consideration of reasonable quantities and the variety of foods chosen.
5. Weighing the admission costs to parks against the costs of the foods desired.
6. Consideration of budget limitations
7. Consideration of human elements such as taste preferences and activity preferences
8. Logical and clear presentation of rationale.

However, particularly on the pre-test, the children may not respond in this manner. This in itself is interesting and important data and should not be interpreted as resulting from the format of the problem.

After the testing session is over, review the tape on your own. If you think any part of the conversation will be difficult for us to understand, please make a note of what was said and attach it to the observation form. Please besure to return to us all tapings, observation sheets, and scrap papers the students wrote on. The pre-test results should be sent to us soon after they have been completed. The Picnic Problem Manual, map and photograph should be retained by you after administration of the pretests. They should be used again for administration of the posttests. Upon completion of the posttests, please return to us the Manual, map and photograph along with the testing results for the posttest.

Instructions for administration of this Picnic Problem will have been reviewed in detail at your Observers' Training Workshop. However, if you have any further questions when you are ready to administer the test, please call the USMES Evaluation Team, collect, at (617) 353-3312.

Dr. Mary H. Shann

APPENDIX H

Scoring Protocol for the Picnic Problem: A Manual for Rating and Coding Students' Performance On a Test of Complex Problem Solving

Prepared by

Mary H. Shann, Ph.D.

USMES Evaluation Project Director

Boston University

Section I. -- IDENTIFICATION (Columns 1-20)

I.D. code records the teacher, grade level, unit and other descriptive information related to reliability and validity issues.

Column 1: identifies form of the problem-solving test.

6 = Playground

7 = Picnic

Column 2: identifies time of testing.

1 = Pre-test

2 = Post-test

Column 3: identifies treatment.

1 = USMES

2 = Control

Columns 4,5: identify teacher.

(See master list for teacher codes.)

Columns 6,7,8: identify grade level.

(See master list for grade level codes.)

In columns 9 and 10 enter the unit code as follows:

Advertising	01
Bicycle Transportation	02
Burglar Alarm Design (now called Protecting Property) (may also be called Security by some teachers)	03
Classroom Design	04
Classroom Management	05

Traffic Flow	30
Using Free Time (formerly Designing Indoor/ Outdoor Games)	31
Using Free Time After School (After School Activities)	32
Ways to Learn	33
Weather Predictions	34

										Blank
1	2	3	4	5	6	7	8	9	10	11

Column 11: Leave Blank

Based on your review of the audio tape and observer's notes, indicate whether you think any of the following factors may render this testing session invalid. Code your response 0 = No, 1 = Yes in the appropriate column.

Problem	Column
Biased selection of students	12
Prompting by observer	13
Prior student experience with this test	14
Inclement weather (0 for picnic problem)	15
Noisy testing environment	16
Outside interference/interruptions	17
Observer deviated from standard procedure	18
Blank	19-20

Section II. -- BEHAVIORAL ASPECTS (Columns 21-24)

There are four factors which are considered in this segment. The scoring of this group shall proceed as follows:

Factor: 1

Motivation: to accept the problem and attempt to solve the problem.

Scoring:

0 No one accepts problem or trys to solve problem.

1 1 Student accepts/trys to solve problem.

2 2 Students accepts/trys to solve problem.

3 3 Students accepts/trys to solve problem.

4 4 Students accepts/trys to solve problem.

5 5 Students accepts/trys to solve problem.

Enter the proper score in column 21.

Factor: 2

Commitment to task: the level of intensity of the group to continue working toward a solution.

Scoring:

0 No effort.

1 Disinterested, fooling around, little input.

2 Some positive input (one or two interested in problem and working with little progress).

3 Group is interested but efforts are not organized, and time is being wasted.

4 Group is interested, working and not wasting time or effort.

Enter proper score in column 22.

Factor: 3

Organization: allocation of responsibilities for efficiency of manpower.

Scoring:

0 No effort.

1 Unplanned, haphazard, or chaotic (students do their own thing - do not allocate item or all work on the same thing).

2 Not all students involved (either by choice or flat). Some are working on problem, some are not - may be arguing among each other.

3 Students have allocated some tasks - may have some working on same item; or possibly 1 may not be involved.

4 Tasks are allocated and students working efficiently -- however, students may have trouble with their item and seek help.

5 Tasks allocated and all are working productively.

Enter proper score in column 23.

Factor: 4

Structure: Group leadership

Scoring:

0 None

1 Autocratic -- one person dominates who does not listen to other students' ideas.

2 Minority Leadership -- one or two persons listen to others and then lead or direct.

3 Plurality -- general agreement of several members leads to direction and leadership; most contributions are recognized and evaluated.

4 Democratic -- all students contribute; no one's suggestions are ignored or ridiculed. One spokesman may arise but sources of ideas/efforts are recognized.

Enter proper score in column 24

21	22	23	24

Section III. -- COGNITIVE ASPECTS (Columns 25-68)

Data for this section can be derived primarily from the observer form and the tapes. It will be necessary to read the observer form and listen to the tapes to bridge any apparent gaps or vague statements found in either the form or the tape.

The cognitive aspects shall include variables considered in solving the problem and the level or method of measuring the variables. The implementation of the measurement in terms of calculation and the recording of the data will be collected and encoded.

A total of 13 variables can be accomodated by the scoring protocol. For each variable, its identification. measurement, calculation and recording will be scored.

III A Factor: COST OF ADMISSION

Identification:

Scoring
0 No
1 Yes

Enter in Column 25.

Measurement:

Scoring
0 No measurement.
1 Vague or very general estimates of the cost of admission at each park.
2 Estimations by imprecise methods or by eyeballing. It does not provide enough information to arrive at a decision.
3 Useful information which can be used to help select park but data should be more accurate or precise.
4 Precise measurement of cost of admission for whole class at each park.

Enter in Column 26.

Calculations:

Scoring
0 No calculations.
1 Vague or very general calculations of cost of admission to each park.
2 Calculations are imprecise or guesses used as an estimate of cost. This is not sufficient to provide necessary data to arrive at a solution.
3 Useful calculations which can be used to arrive at solution, but the data should be more accurate or precise.
4 Calculations are appropriate and precise. Correct calculation of price of admission to each park for entire class.

Enter in Column 27.

Recording:

Scoring
0 No records.
1 Very general or imprecise records.
2 Adequate records.

Enter in Column 28.

25	26	27	28

III B Factor: COST OF FOOD

Identification:

Scoring 0 No

1 Yes

Enter in Column 29.

Measurement:

Scoring 0 No measurement done.

1 Vague or very general estimates of cost of food per person or for entire class.

2 Estimations by imprecise methods or by eyeballing. No attempt to plan menu. It does not provide enough information to arrive at a decision.

3 Useful information which can be used to help select food but data should be more accurate or precise. There is an attempt to plan menu for the class.

4 Precise measurement of cost of food for the entire class is made, staying within budget limitations. A menu is planned.

Enter in Column 30.

Calculations:

Scoring 0 No calculations.

1 Vague or very general calculations of cost of food per person or for entire class.

2 Calculations are imprecise or guesses used as an estimate of cost. No considerations of menu for each person or for entire class. This is not sufficient to provide necessary data to arrive at a solution.

3 Useful calculations which can be used to arrive at solution, but the data should be more accurate or precise. Consideration of menu takes place.

4 Calculations are appropriate and precise. Correct calculation of cost of food for entire class. Menu well planned out, which can lead to a solution.

Enter in Column 31.

Recording:

Scoring 0 No records.

1 Very general or imprecise records.

2 Adequate records.

Enter in Column 32.

29	30	31	32

III C Factor: COST OF FOOD vs. COST OF ADMISSION

Identification:

Scoring 0 No

1 Yes

Enter in Column 33.

Measurement:

Scoring 0 No measurement.

1 Vague or general awareness that cost of food and cost of admission must not exceed the $50. budget limit.

2 Estimates relationship of cost of food to admission cost by imprecise methods or by eyeballing it does not provide enough information to arrive at a decision.

3 Useful information which can be used to judge what proportion of money should be allocated to the food and to the cost of admission respectively, but the data should be more accurate or precise.

4 Precise measurement of relationship between cost of food and cost of admission. Allocates certain proportion of $50. to food and certain proportion to admission fee.

Enter in Column 34.

Calculations:

Scoring 0 No calculations.

1 Vague or very general calculations that do little quantification.

2 Calculations are imprecise or guesses used as an estimate of cost of food and admission. Little awareness of relationship between cost of food and cost of admission. This is not sufficient to provide necessary data to arrive at a solution.

3 Useful calculations which can be used to arrive at solution, but the data should be more accurate or precise. Is aware that certain proportion of money should be allocated to food and a certain proportion to admission.

4 Calculations are appropriate and precise. Correct calculation of both food costs and admission costs. keeping within a budget of $50.

Enter in Column 35.

Recording:

Scoring 0 No records.

1 Very general or imprecise records.

2 Adequate records.

Enter in Column 36.

33	34	35	36

III D Factor: TIME AVAILABLE FOR PICNIC

Identification:

Scoring	0	No
	1	Yes

Enter in Column 37.

Measurement:

Scoring	0	No measurement.
	11	Vague or very general awareness of time limit.
	2	Acknowledges time limitation of 6 hours, including travel time and time at park, and makes plan according to this time limit.

Enter in Column 38.

Calculations:

Scoring	0	No calculations.
	1	Vague or very general calculations involving travel time to each park. General awareness of time limitation as a consideration in choosing a park.
	2	More precise calculations of relative times to get to each park, and then relating travel time to time limitation of 6 hours.

Enter in Column 39.

Recording:

Scoring	0	No records.
	1	Very general or imprecise records.
	2	Adequate records.

Enter in Column 40.

37	38	39	40

III E Factor: TRAVEL TIME vs. PLAYTIME

Identification:

Scoring	0	No
	1	Yes

Enter in Column 41.

Measurement:

Scoring	0	No measurement.
	1	Vague or very general awareness that both time factors should be taken into consideration in choosing park.
	2	More precise measurements of travel time to each park and judging what proportion of time should be spent traveling and what proportion of time should be spent for playing in in the park.

Enter in Column 42.

Calculations:

Scoring	0	No calculations.
	1	Vague or general estimates of relative travel times to each park by eyeballing or guessing and then consideration and general estimation of time left over for play at each park.
	2	More precise calculations of relative travel times to each park, and time left over for play at each park.

Enter in Column 43.

Recording:

Scoring	0	No records.
	1	Very general or imprecise records.
	2	Adequate records.

Enter in Column 44.

41	42	43	44

III F Factor: CONSIDERATION OF FOOD COST, ADMISSION COST AND TIME RELATIONSHIPS

Identification:

Scoring 0 No

1 Yes

Enter in Column 45.

Measurement:

Scoring 0 No measurement.

1 Vague or general statements regarding the relationship of the 3 factors, which are used to help lead to a solution.

2 More precise statements and/or estimates of the relationship of the three factors, which can help lead to a solution.

Enter in Column 46.

Calculations:

Scoring 0 No calculations.

1 Very general estimates of the relationship of food costs, admission costs and time. Weighing of the pros and cons of different alternatives occurs.

2 More precise calculations of different alternative solutions (regarding selection of food and a specific park), recognition of the relationship of the 3 factors, and selection of one alternative (e.g., calculates travel time, and amount of money left for food at each of the 3 parks).

Enter in Column 47.

Recording:

Scoring 0 No records.

1 Very general or imprecise records.

2 Adequate records.

Enter in Column 48.

45	46	47	48

III G Factor: DISTANCES TO PARKS

Identification:

Scoring 0 No

1 Yes

Enter in Column 49.

Measurement:

Scoring 0 No measurement.

1 Awareness that distance to each park should be taken into consideration. Vague or very general estimates are made.

2 Estimates by imprecise methods or by eyeballing.

3 Useful information which can be used to arrive at a decision. Recognition of the use of the map scale, but measurement should be more accurate or precise.

4 Precise measurement of distance to each park made, and recognition that travel time within the park to particular facilities should be included in the total distance to each park.

Enter in Column 50.

Calculations:

Scoring 0 No calculations.

1 Vague or very general calculations that do little quatification (e.g. Forest Valley Park looks twice as far away as Pine Hill Park).

2 Calculations are imprecise or guessing occurs and are not sufficient to provide necessary data to arrive at a solution, (e.g., Pine Hill Park looks about 30 miles away).

3 Useful calculations using the map scale which can be used to arrive at a solution. It may not be accurate or have considered distances to be traveled within the park to the facilities in to the total distance to be traveled to each park.

4 Calculations are appropriate, precise and can lead to a solution.

Enter in Column 51.

Recording:

Scoring 0 No records.

1 Very general or imprecise records.

2 Adequate records.

Enter in Column 52.

49	50	51	52

III H Factor: SIZE OF FACILITIES

Identification:

Scoring	0	No
	1	Yes

Enter in Column 53.

Measurement:

Scoring	0	No measurement.
	1	Vague or general estimates (i.e., Greehill Park is much bigger than Pine Hill Park).
	2	More precise measures of the size of each park (i.e., using map scales to roughly measure the area of each park).

Enter in Column 54.

Calculations:

Scoring	0	No calculations.
	1	General estimates of the size of each park, mainly by eyeballing.
	2	More careful calculations, using the map scale to figure out the approximate areas of each park.

Enter in Column 55.

Recording:

Scoring	0	No records.
	1	Very general or imprecise records.
	2	Adequate records.

Enter in Column 56.

53	54	55	56

III I Factor: PLAY EQUIPMENT (Brought along for children's use at playground, e.g., baseballs and bats)

Identification:

Scoring 0 No

1 Yes

Enter in Column 57.

Measurement:

Scoring 0 No measurement.

1 Vague or general estimates of type of amount of equipment that should be brought to park.

2 Express need to know specific number of different pieces of equipment to be brought to park, taking in to consideration the number of children who would be using each particular piece of equipment.

Enter in Column 58.

Calculations:

Scoring 0 No calculations.

1 General or arbitrary assignment of equipment for children participating in the picnic.

2 More careful estimates, with selection of equipment reflecting individual child preferences, abilities and whether or not the amount of equipment brought along is in proportion to the number of children utilizing it.

Enter in Column 59.

Recording:

Scoring 0 No records.

1 Very general or imprecise records.

2 Adequate records.

Enter in Column 60.

57	58	59	60

III J Factor: SAFETY CONSIDERATIONS FOR TRIP

Identification:

Scoring
0 No
1 Yes

Enter in Column 61.

Measurement:

Scoring
0 No measurement.
1 General or vague considerations of safety precautions, and more or less safety of each park, (e.g., in the large park, there is a greater possibility of someone getting lost).
2 More precise safety measures taken, (e.g., specific assignment of adults for supervision on the bus and at the park.

Enter in Column 62.

Calculations:

Scoring
0 No calculations.
1 Vague or general references to safety precautions that should be taken.
2 More careful or precise calculations made in order to have a safe trip, (e.g., number of supervisors needed).

Enter in Column 63.

Recording:

Scoring
0 No records.
1 Records.

Enter in Column 64.

61	62	63	64

III K Factor: OTHER CONSIDERATIONS

Column:

65 Number of additional factors mentioned.

66 "Fun" mentioned as consideration, (0=no, 1=yes).

67 Blank.

68 Blank.

Enter in Column 65-68.

65	66	67	68

APPENDIX I

Test of Problem Solving Skills[a]

"TOPSS"

A Paper-and-Pencil Test Designed

for Group Administration to

Elementary School Children

[a]Part I of TOPSS consists of 30 items selected from the STEP-Test, Level IV (ETS, 1958) to be used for research purposes under licensing agreement between ETS and Professor Mary H. Shann, Project Director of the USMES Evaluation. Part II of TOPSS includes 22 items developed by the USMES Evaluation Team. The deletion of items numbered 50-55 is recommended in future use of this test.

PROBLEM TEST I

DIRECTIONS FOR ADMINISTERING

Note: Instructions which are to be read aloud to students are in capital type. Instructions printed in regular type are intended only for the examiner.

In these Directions and in the test booklets the students are instructed to make their marks heavy and black. This is desirable for IBM answer sheets. When the students are assembled in the examination room and seated, say:

THE TESTING PERIOD HAS BEGUN. THERE SHOULD BE NO TALKING AMONG YOU UNTIL AFTER YOU HAVE BEEN DISMISSED.

WE SHALL NOW PASS OUT TEST MATERIALS. DO NOT OPEN YOUR BOOKLET OR TURN IT OVER UNTIL YOU ARE TOLD TO DO SO.

Distribute booklets, answer sheets, and pencils. Have students fill in the following identification information on their answer sheets: (1) fill in the school, city, instructor, grade, and test name (Problem Test I) in the blank spaces, (2) fill in the name, grade, birthdate and sex in the grids.

Then say:

LOOK AT THE GENERAL DIRECTIONS. READ THESE DIRECTIONS SILENTLY WHILE I READ THEM ALOUD.

GENERAL DIRECTIONS

THIS IS A TEST OF SOME OF THE UNDERSTANDINGS, SKILLS, AND ABILITIES YOU HAVE BEEN DEVELOPING EVER SINCE YOU FIRST ENTERED SCHOOL. YOU SHOULD TAKE THE TEST IN THE SAME WAY THAT YOU WOULD WORK ON ANY NEW AND INTERESTING ASSIGNMENT. HERE ARE A FEW SUGGESTIONS WHICH WILL HELP YOU TO EARN YOUR BEST SCORE:

1. MAKE SURE YOU UNDERSTAND THE TEST DIRECTIONS BEFORE YOU BEGIN WORKING. YOU MAY ASK QUESTIONS ABOUT ANY PART OF THE DIRECTIONS YOU DO NOT UNDERSTAND.
2. YOU WILL MAKE YOUR BEST SCORE BY ANSWERING EVERY QUESTION BECAUSE YOUR SCORE IS THE NUMBER OF CORRECT ANSWERS YOU MARK. THEREFORE, YOU SHOULD WORK CAREFULLY BUT NOT SPEND TOO MUCH TIME ON ANY ONE QUESTION. IF A QUESTION SEEMS TO BE TOO DIFFICULT, MAKE THE MOST CAREFUL GUESS YOU CAN, RATHER THAN WASTE TIME PUZZLING OVER IT.
3. IF YOU FINISH BEFORE TIME IS CALLED, GO BACK AND SPEND MORE TIME ON THOSE QUESTIONS ABOUT WHICH YOU WERE MOST DOUBTFUL.

ARE THERE ANY QUESTIONS?

Answer any legitimate questions. Stay within the meaning and, as far as possible, use the vocabulary of the printed directions.

YOU WILL FIND DIRECTIONS FOR THE TEST. LOOK AT THEM AND READ THESE DIRECTIONS SILENTLY WHILE I READ THEM ALOUD.

PROBLEM TEST I

DIRECTIONS

EACH OF THE QUESTIONS OR INCOMPLETE STATEMENTS IN THIS TEST IS FOLLOWED BY FOUR SUGGESTED ANSWERS. YOU ARE TO DECIDE WHICH ONE OF THESE ANSWERS YOU SHOULD CHOOSE. YOU MUST MARK ALL OF YOUR ANSWERS ON THE SEPARATE ANSWER SHEET YOU HAVE BEEN GIVEN; THIS TEST BOOKLET SHOULD NOT BE MARKED IN ANY WAY. YOU MUST MARK YOUR ANSWER SHEET BY BLACKENING THE SPACE HAVING THE SAME LETTER AS THE ANSWER YOU HAVE CHOSEN.

FOR EXAMPLE:

0 WHICH ONE OF THE FOLLOWING IS AN ANIMAL?
A BED B DOG C CHAIR D BOX

SINCE A DOG IS AN ANIMAL, YOU SHOULD CHOOSE THE ANSWER LETTERED B. ON YOUR ANSWER SHEET, YOU WOULD FIRST FIND THE ROW OF SPACES NUMBERED THE SAME AS THE QUESTION--IN THE EXAMPLE ABOVE, IT IS 0. THEN YOU WOULD BLACKEN THE SPACE IN THIS ROW WHICH HAS THE SAME LETTER AS THE ANSWER YOU HAVE CHOSEN.

MAKE YOUR ANSWER MARKS HEAVY AND BLACK. MARK ONLY ONE ANSWER FOR EACH QUESTION. IF YOU CHANGE YOUR MIND ABOUT AN ANSWER, BE SURE TO ERASE THE FIRST MARK COMPLETELY.

THE EXAMPLE HAS BEEN GIVEN TO YOU SO THAT YOU WILL KNOW HOW TO MARK YOUR ANSWER SHEETS. THE QUESTIONS ON THE INSIDE OF THE TEST ARE NOT JUST LIKE THE EXAMPLE; BUT EACH ONE DOES PRESENT FOUR CHOICES, AND YOU MUST CHOOSE YOUR ANSWER FROM AMONG THEM.

ARE THERE ANY QUESTIONS?

Answer any legimate questions. Stay within the meaning and, as far as possible use the vocabulary of the printed directions.

WHEN I SAY "BEGIN," TURN TO p.1 AND START WORKING. READY? BEGIN!

Examiner and proctors (if any) should move quietly about the room to see that each student is working on the proper pages of his test booklet and that he is marking his answers correctly in the proper section of the answer sheet.

At the end of exactly 35 minutes, say:

STOP! EVEN IF YOU HAVE NOT FINISHED, YOU MUST STOP AND LAY DOWN YOUR PENCIL.

Collect answer sheets, test booklets, and other test materials and then dismiss the students.

At this time you should write down for the record a description of any unexpected variation from the normal testing procedure that may have occurred. Such incidents need to be in the record and considered when scores are interpreted.

PROBLEM TEST I

GENERAL DIRECTIONS

This is a test of some of the understandings, skills, and abilities you have been developing ever since you first entered school. You should take the test in the same way that you would work on any new and interesting assignment. Here are a few suggestions which will help you to earn your best score:

1. Make sure you understand the test directions before you begin working. You may ask questions about any part of the directions you do not understand.

2. You will make your best score by answering every question because your score is the number of correct answers you mark. Therefore, you should work carefully but not spend too much time on any one question. If a question seems to be too difficult, make the most careful guess you can, rather than waste time puzzling over it.

3. If you finish before time is called, go back and spend more time on those questions about which you were most doubtful.

Each of the questions or incomplete statements in this test is followed by four suggested answers. You are to decide which one of these answers you should choose.

You must mark all of your answers on the separate answer sheet you have been given; this test booklet should not be marked in any way. You must mark your answer sheet by blackening the space having the same letter as the answer you have chosen. For example:

0 Which one of the following is an animal?
- A Bed
- B Dog
- C Chair
- D Box

Since a dog is an animal, you should choose the answer lettered B. On your answer sheet, you first find the row of spaces numbered the same as the question--in the example above, it is 0. Then you would blacken the space in this row which has the same letter as the answer you have chosen. Example:

Make your answer marks heavy and black. Mark only one answer for each question. If you change your mind about an answer, be sure to erase the first mark completely.

PROBLEM TEST II

DIRECTIONS FOR ADMINISTERING

Note: Instructions which are to be read aloud to students are in capital type. Instructions printed in regular type are intended only for the examiner.

In these Directions and in the test booklets the students are instructed to make their marks heavy and black. This is desirable for IBM answer sheets. When the students are assembled in the examination room and seated, say:

THE TESTING PERIOD HAS BEGUN. THERE SHOULD BE NO TALKING AMONG YOU UNTIL AFTER YOU HAVE BEEN DISMISSED.

WE SHALL NOW PASS OUT TEST MATERIALS. DO NOT OPEN YOUR BOOKLET OR TURN IT OVER UNTIL YOU ARE TOLD TO DO SO.

Distribute booklets, answer sheets, and pencils. Have students fill in the following identification information on their answer sheets: (1) fill in the school, city, instructor, grade, and test name (Problem Test II) in the blank spaces, (2) fill in the name grade, birthdate and sex in the grids.

Then say:

LOOK AT THE GENERAL DIRECTIONS. READ THESE DIRECTIONS SILENTLY WHILE I READ THEM ALOUD.

GENERAL DIRECTIONS

THIS IS A TEST OF SOME OF THE UNDERSTANDINGS, SKILLS, AND ABILITIES YOU HAVE BEEN DEVELOPING EVER SINCE YOU FIRST ENTERED SCHOOL. YOU SHOULD TAKE THE TEST IN THE SAME WAY THAT YOU WOULD WORK ON ANY NEW AND INTERESTING ASSIGNMENT. HERE ARE A FEW SUGGESTIONS WHICH WILL HELP YOU TO EARN YOUR BEST SCORE:

1. MAKE SURE YOU UNDERSTAND THE TEST DIRECTIONS BEFORE YOU BEGIN WORKING. YOU MAY ASK QUESTIONS ABOUT ANY PART OF THE DIRECTIONS YOU DO NOT UNDERSTAND.
2. YOU WILL MAKE YOUR BEST SCORE BY ANSWERING EVERY QUESTION BECAUSE YOUR SCORE IS THE NUMBER OF CORRECT ANSWERS YOU MARK. THEREFORE, YOU SHOULD WORK CAREFULLY BUT NOT SPEND TOO MUCH TIME ON ANY ONE QUESTION. IF A QUESTION SEEMS TO BE TOO DIFFICULT, MAKE THE MOST CAREFUL GUESS YOU CAN, RATHER THAN WASTE TIME PUZZLING OVER IT.

Answer any legitimate questions. Stay within the meaning and, as far as possible, use the vocabulary of the printed directions.

YOU WILL FIND DIRECTIONS FOR THE TEST. LOOK AT THEM AND READ THESE DIRECTIONS SILENTLY WHILE I READ THEM ALOUD.

DIRECTIONS

EACH OF THE QUESTIONS OR INCOMPLETE STATEMENTS IN THIS TEST IS FOLLOWED BY YOUR SUGGESTED ANSWERS. YOU ARE TO DECIDE WHICH ONE OF THESE ANSWERS YOU SHOULD CHOOSE. YOU MUST MARK ALL OF YOUR ANSWERS ON THE SEPARATE ANSWER SHEET YOU HAVE BEEN GIVEN; THIS TEST BOOKLET SHOULD NOT BE MARKED IN ANY WAY.

PROBLEM TEST II

YOU MUST MARK YOUR ANSWER SHEET BY BLACKENING THE SPACE HAVING THE SAME LETTER AS THE ANSWER YOU HAVE CHOSEN.

FOR EXAMPLE:

0 WHICH ONE OF THE FOLLOWING IS AN ANIMAL?
A BED B DOG C CHAIR D BOX

SINCE A DOG IS AN ANIMAL, YOU SHOULD CHOOSE THE ANSWER LETTERED B. ON YOUR ANSWER SHEET, YOU WOULD FIRST FIND THE ROW OF SPACES NUMBERED THE SAME AS THE QUESTION--IN THE EXAMPLE ABOVE, IT IS 0. THEN YOU WOULD BLACKEN THE SPACE IN THIS ROW WHICH HAS THE SAME LETTER AS THE ANSWER YOU HAVE CHOSEN.

MAKE YOUR ANSWER MARKS HEAVY AND BLACK. MARK ONLY ONE ANSWER FOR EACH QUESTION. ONCE YOU GO ON TO THE NEXT QUESTION DO NOT GO BACK AND CHANGE YOUR ANSWER.

THE EXAMPLE HAS BEEN GIVEN TO YOU SO THAT YOU WILL KNOW HOW TO MARK YOUR ANSWER SHEETS. THE QUESTIONS ON THE INSIDE OF THE TEST ARE NOT JUST LIKE THE EXAMPLE; BUT EACH ONE DOES PRESENT FOUR CHOICES, AND YOU MUST CHOOSE YOUR ANSWER FROM AMONG THEM.

ARE THERE ANY QUESTIONS?

Answer any legitimate questions. Stay within the meaning and, as far as possible, use the vocabulary of the printed directions.

WHEN I SAY "BEGIN," TURN TO p. 1 AND START WORKING. REMEMBER, START AT NUMBER 41 ON THE ANSWER SHEET. READY? BEGIN!

Examiner and proctors (if any) should move quietly about the room to see that each student is working on the proper pages of his test booklet and that he is marking his answers correctly in the proper section of the answer sheet.

At the end of exactly 45 minutes, say:

STOP! EVEN IF YOU HAVE NOT FINISHED, YOU MUST STOP AND LAY DOWN YOUR PENCIL.

Collect answer sheets, test booklets, and other test materials and then dismiss the students.

At this time you should write down for the record a description of any unexpected variation from the normal testing procedure that may have occurred. Such incidents need to be in the record and considered when scores are interpreted. Please note on a piece of paper the time at which the first student finished the test, and the time at which the last student finished the test (if applicable).

PROBLEM TEST II

GENERAL DIRECTIONS

This is a test of some of the understandings, skills, and abilities you have been developing ever since you first entered school. You should take the test in the same way that you would work on any new and interesting assignment. Here are a few suggestions which will help you to earn your best score:

1. Make sure you understand the test directions before you begin working. You may ask questions about any part of the directions you do not understand.

2. You will make your best score by answering every question because your score is the number of correct answers you mark. Therefore, you should work carefully but not spend too much time on any one question. If a question seems to be too difficult, make the most careful guess you can, rather than waste time puzzling over it.

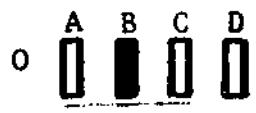
Each of the questions or incomplete statements in this test is followed by four suggested answers. You are to decide which one of these answers you should choose.

You must mark all of your answers on the separate answer sheet you have been given; this test booklet should not be marked in any way. You must mark your answer sheet by blackening the space having the same letter as the answer you have chosen.

For Example:

0 Which one of the following is an animal?
- A Bed
- B Dog
- C Chair
- D Box

Since a dog is an animal, you should choose the answer lettered B. On your answer sheet, you first find the row of spaces numbered the same as the question--in the example above, it is 0. Then you would blacken the space in this row which has the same letter as the answer you have chosen. Example:



Make your answers heavy and black. Mark only one answer for each question. Once you go on to the next question do not go back and change your answer.

BIKE TRANSPORATION

Lots of kids ride their bikes to Vista School. There have been some bike accidents this year. A boy rode over a curb and fell off his bike. One girl almost got hit by a car crossing a busy street.

The students in Mrs. Martin's class want to try to make it safer and easier for everyone to ride bikes to school.

The class talked about bike safety. Some kids said they rode on the sidewalk. Some said they rode on the left side of the street. Some rode on the right side. Many kids said they did not watch traffic signs.

41. If Mrs. Martin's class wants to help make bike riding safer, what do you think they need to find out first?

    A What kind of bikes kids have.
    B What are the traffic laws for bikes and for cars.
    C How many people ride on the left side of the street.
    D How many bike accidents there have been in the past year.

42. The class invited a policeman to explain bike safety rules. Later John and Jake went to the library to look up bike safety rules in some books. They found three safety rules which the policeman did not tell them. What do you think they should do?

    A Tell the policeman he missed three rules.
    B Use only the rules from the book.
    C Put all the rules together in one list.
    D Pick the rules they like the best.

43. John's group decided to teach bike safety rules to other kids. What should they do to find out how much people know about safety already?

    A Graph the number of kids who have had accidents.
    B Ask people if they ride bikes safely.
    C Make up a test of safety rules and give it to everyone.
    D Put up a chart of the safety rules and ask each person if he read it.

44. The group found out that many kids did not know the rules of bike safety. What should they do now?

A Let only the kids who know the rules ride to school.
B Suggest that everyone rides the bus because it is safer.
C Plan a bike safety program to teach everyone the rules.
D Send a copy of the rules home to all the parents.

45. Amy and Marty's group decided to make a record of all the accidents near the Vista School. The group recorded these facts from November to March:

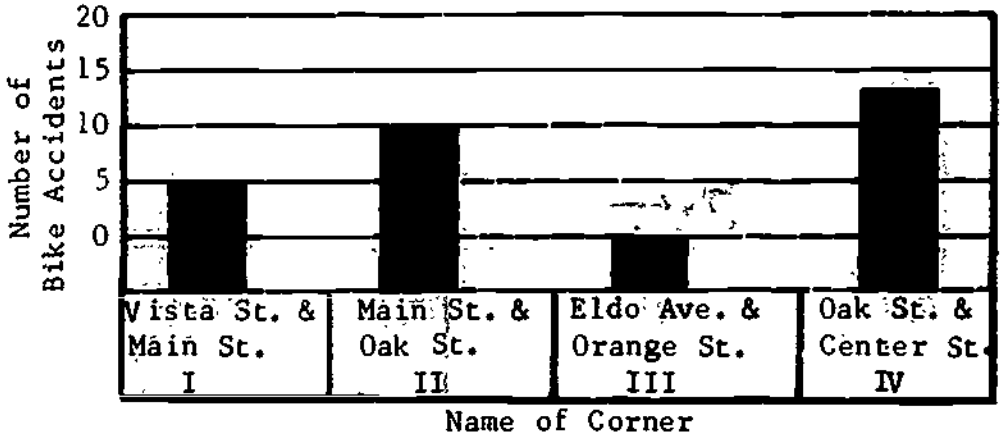
Accidents from Nov. 1, 1974 to March 1, 1975

Type of Accident	Number of Accidents
1. Bike riders hit by car.	2
2. Bike riders who hit curb and fell.	1
3. Bike riders hit by another bike.	2
4. Cars hit by trucks.	2
5. Cars hit by cars.	2
6. Bike riders hit by school bus.	1

Which facts on the chart would <u>not</u> help solve the problem of bike safety?

A 1 and 6.
B 2 and 4.
C 4 and 5.
D 2 and 5.

46. Marty made this graph showing how many bike accidents there were at four different corners in one year.



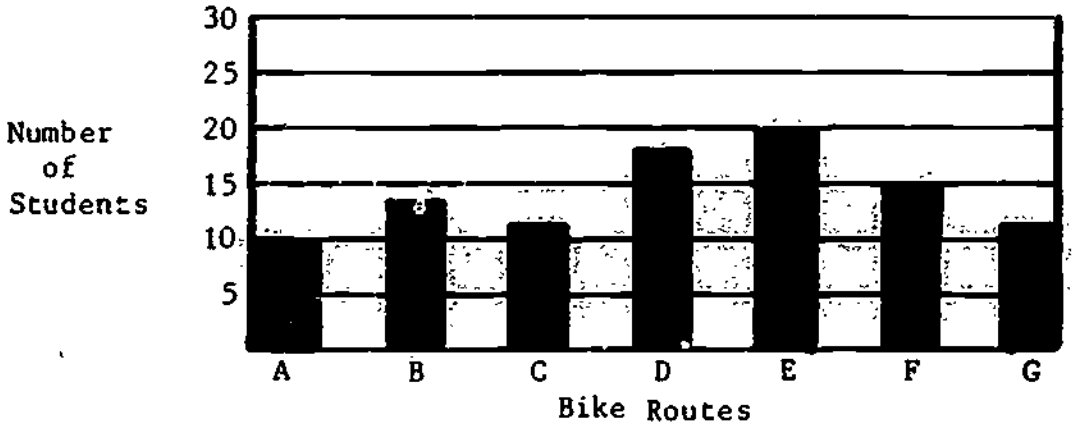
According to the graph, which fact is true?

A All four corners have the same number of accidents.
B Half as many accidents happened at Corner II as at Corner IV.
C The smallest number of accidents happened at the corner of Eldo Ave. and Orange St.
D 15 accidents happened at the corner of Main St. and Oak St.

47. After finding out about the number of accidents near the Vista School, Chris' group wants to map out safe bike routes. They ask groups of children to map out safe routes to school. They want to check their maps for accuracy. What is the best way?

A The children go out by car and retrace the routes checking their maps.
B The children ask Mrs. Martin to check their maps.
C Have a policeman come to the school and check the maps.
D Put the maps in an opaque projector and shine them on the wall when checking them.

48. Next Chris' group wants to find out which bicycle routes the kids will use and which ones they will not use. After questioning all the kids in the school they make a graph showing how many kids will take each route. Which three routes will be used most?



A BDE.
B DEF.
C BEF.
D BDF.

49. The kids found out that three routes would be used the most. What should they do to make these routes ready for bike riders?

A Ask the principal for permission to have the bike routes.
B Go to the police department for their advice on making these new routes safe.
C Send a note out to all bike riders and tell them these are the routes they must use.
D Put an article in the town newspaper telling all to use these bike routes.

LUNCH TIME

Lunch at the Smith School is a busy, noisy time. Lunch period is from 12 to 1 o'clock. Grades 1 to 6 come to the lunchroom at 12. The lunch line is very long.

50. Miss Walter's class decided that the trouble was that too many people were trying to get lunch at the same time. They would like to work on this problem. What do they need to find out?

    A How many kids buy lunch each week.
    B Can the lunchroom be made bigger.
    C Can a new lunch schedule be made up.
    D Do students want different kinds of lunches to be served.

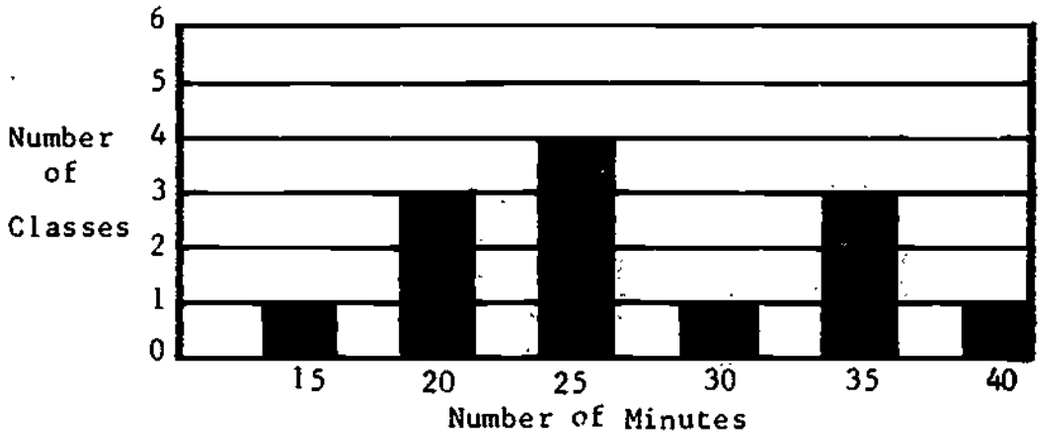
51. The class decided to work in small groups. Sally's group wanted to find out how long it took each class to have lunch. What is the best way for her group of six kids to do this?

    A Ask each teacher about how long it takes for her class to have lunch.
    B Have each student timed by one kid in the group.
    C Have each class timed by one kid in the group.
    D Ask all the students how long lunch takes and find the average.

52. What is the easiest way for Sally's group to record the time it takes each class to get and eat lunch?

    A Record the time the first person in each class left his classroom and the time the last person sat down to eat.
    B For each class, record the time the first person left the room and the time the last person left the lunchroom.
    C Time how long it takes each person to get his tray, sit down and finish eating.
    D Record the time the first kid in each class reached the cash register and the time he picked up his tray to leave.

53. Sally and Bart made this graph to show how long it took the classes to have lunch:



Most classes took how long?

A 15-20 minutes.
B Over 30 minutes.
C 25-30 minutes.
D 20-25 minutes.

54. When Jenny was timing Mr. Carter's sixth grade, she saw many of the kids cut in line. They got through fast and took only 15 minutes to have lunch. What do you think Sally's group ought to do?

A Tell Mr. Carter his students cut in line and time them again tomorrow.
B Leave Mr. Carter's class out when making their graph.
C Add 5 minutes to the time Mr. Carter's class took to have lunch.
D Use the 15 minute time for Mr. Carter's class when making their graph.

55. When Sally's group finished their project they found that most students could get and eat their lunch in 19 minutes. They want to ask for a new lunch schedule so that the lunchroom is not so noisy and crowded. Which schedule do you think is best?

A 4 lunch periods: 12:00, 12:15, 12:30, 12:45.
B 3 lunch periods: 12:00, 12:15, 12:30.
C 3 lunch periods: 12:00, 12:20, 12:40.
D 2 lunch periods: 12:00, 12:30.

A BUSY CORNER

Most people who go to the Raymond School have to cross the street at the corner of Broadway and Lincoln Avenue. Cars come speeding down the street. There is no stop light or stop sign at the corner. The crossing guard has trouble helping the kids get across in time. Some of the kids think it's scary to cross the street there.

56. In Mr. Newman's fifth grade, the class talked about trying to do something about the problem at the corner. What do you think would be most important to work on?

    A Finding another way for kids to get to school.
    B Asking for an extra crossing guard.
    C Figuring out how to stop the speeding cars.
    D Talking to the kids who are scared about crossing.

57. Molly made a list of the questions which people asked when the class talked about the problem of crossing the street at the corner. Which question do you think is the most important?

    A What time is the crossing guard at the corner each day?
    B What do the red, yellow and green lights mean?
    C Should we obey the traffic laws?
    D Can the corner be made safer?

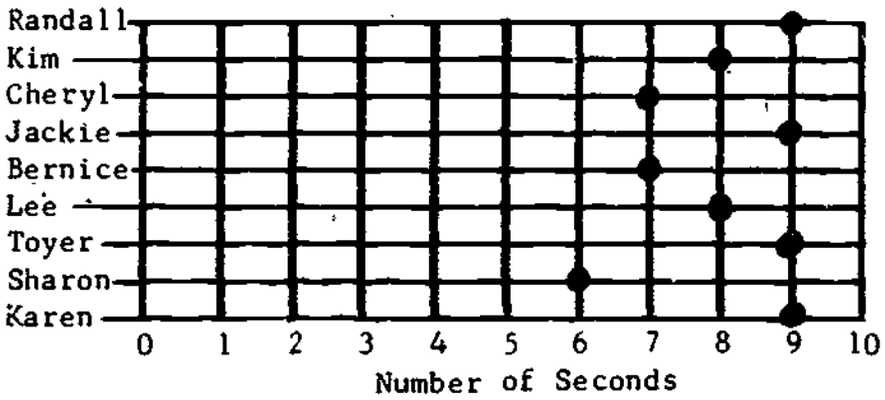
58. One group wants to measure how long it takes to cross the street. How would you do it?

    A Measure the distance from one side of the street to the other with a tape measure.
    B Time the cars going by with a stop watch.
    C Time kids going across with a stop watch.
    D Time kids going across by counting seconds.

59. Kim and Sharon decided to time Randall crossing the street. Kim says it took ten seconds, Sharon says it took eight seconds. They both used stop watches. What do you think Kim and Sharon should do now?

A Try again, using new stop watches.
B Try again, using the teacher's wrist watch.
C Try again, both starting when Randall steps off the curb and stopping when he steps on the curb.
D Try again, but this time ask Randall to run from curb to curb in eight seconds exactly.

60. Mark made this graph showing how long it took people to cross the street:



How long did it take Cheryl to cross the street?

A 6 seconds.
B 9 seconds.
C 7 seconds.
D 8 seconds.

61. After timing people from each grade and adults, too, the kids discovered that the average crossing time was 11 seconds. What would you suggest to make the crossing safe?

A Put up a walk light that stops cars for 11 seconds each time it goes on.
B Put up a walk light that stops cars for 15 seconds each time it goes on.
C Teach each person to cross the street in less than 11 seconds.
D Have the crossing guard stop traffic every 11 seconds.

62. Mr. Newman's class has written a report about their work on the corner. They have decided to suggest that a stop light or stop signs are needed. What do you think they should do now?

A Get permission to tell the other children in the Raymond School about their project.
B Ask Mr. Newman to give each student a grade for his work on the report.
C Ask the principal for materials to build a stop light.
D Invite the police chief to their class to listen to their report.

APPENDIX J

Problem-Solving Ability of Individual Students

On this form we would like you to rate the problem-solving ability of individual students. Students are to be rank ordered, from those showing most ability to solve problems to those showing the least ability.

In the left-hand column below list your class alphabetically.

In the right-hand column put a "1" next to the name of the student who copes most effectively with problems which arise in daily activities. Put a "2" next to the name of the student who is second-best in dealing with every-day problems. Continue numbering in this manner until each student is ranked.

We realize the difficulty of ranking in this way but your estimate of each student's ability to solve real problems will help us in determining the value of our test.

TEACHER: _____ SCHOOL: _____

STUDENT'S NAME	RANK

APPENDIX K

Initial Draft of Categories of Behaviors

to be Observed with PROFILES

Process Study - Draft Form

Date:_____ Student:_____

Visit Number:_____ Teacher:_____

Unit:_____

☐ 1. Defines problem.
[Does the problem satisfy the criteria for "real?" See Form II.]

☐ 2. Selects and defines individual sub-task.

☐ 3. Relates sub-task to total problem.

☐ 4. Relates contributions of others to total problem.

☐ 5. Acquires needed skills.

☐ 6. Acquires pertinent information.

☐ 7. Plans action.

☐ 8. Implements plan of action.

☐ 9. Encounters hang-ups, problems, and errors.

☐ 10. Reformulates plan of action.

☐ 11. Completes sub-task.

☐ 12. Relates results of sub-task to total problem.

☐ 13. Communicates findings to others.

GO TO STEP 2

☐ 14. Participates in final re-olution of problem.

☐ 15. Implements solution to the problem.

☐ 16. Measures success against all relevant aspects of the situation.

☐ 17. Accepts consequences of the solution.
[Does the problem satisfy the criteria for "real?"]

APPENDIX L

PROFILES of Problem Solvers Categories of Behaviors

to Observe in the Problem Solving Process

A. Define Problem (Challenge)

The child is able to recognize (and verbalize) that a challenge, (or a "real" problem) has been presented to the class which needs to be solved. He must realize that many questions must be answered, (and sub-tasks performed) in order to help him solve the problem.

B. Select and Define Sub-Task

The child is able to explain the work his group (or class) is involved in at the present time. He can verbalize why the group (or class) decided to undertake this particular task.

C. Relate Current Sub-Task to Problem

The child explains how the work he is presently involved in will help solve the total problem.

D. Relate Contributions of Others to Problem

This category only applies to classes in which separate groups have been formed. The child can explain what work other groups are involved in, and why their work will help to solve the total problem.

E. Acquire Skills and Pertinent Information

The child indicates that he has learned new skills or information that will help him in his work, and in solving the problem.

F. Plans Action

The child is able to explain how the class (or group) decided what work to do at the present time. He is also aware of who made these decisions.

G. Encounters Errors and Hang-Ups and Reformulates Plans

The child is aware that he has come across a problem or problems while doing work on a sub-task. He is able to explain what was done, or what will be done to rectify the situation.

H. Plan Next Steps

The child is aware of the next step or steps he will take to help solve the problem. This could relate to the next sub-task that will be undertaken by the child.

I. Organize, Analyze, and Interpret Data

The child can explain what he has found out as a result of his inquiries. He can explain any records or reports that he has drawn up.

J. Relate Results of Sub-Task to Total Problem

The child can explain how the results of his work on various sub-tasks help him solve the total problem.

K. Communicates Findings

The child has told other members of the class, school and/or community the results of his work on the problem.

APPENDIX M

PROFILES Interview Guide: Level I

for Second and Third Grades

A. Define Problem (Challenge) [Note: This category should be used during your first interview with each child, even if the problem was defined during a previous USMES class. Ask these questions again if the issue is brought up at a later class meeting.]

1. "Can you tell me what you've been doing today?" [If the child answers the question in terms of the sub-task he is working on, rather than in terms of the total problem. THEN add]: "Yes, but I was wondering if you could tell me a little bit more about [mention the title of USMES project] in general."

2. Why do you think it is a good idea for the class to do this work."

3. [Note: This question should only be asked if the challenge involves a "problem"].
"Do you really think you'll be able to do anything about this problem?"
"Why/Why not?"

4. "How did the class decide to do this work?"

5. "Can you tell me whose idea it was to do this work?"

B. Select and Define Sub-task [Note: For category B, only ask questions from Ba, or Bb. Section Ba, should be used if the whole class is working on the USMES challenge as one unit. Section Bb should be used if the class has split up into several groups which are working independently of one another.]

Ba. Select and Define Class Work

1. "Can you tell me what your class is working on right now?"

2. "How did you decide to do this work?" "Project?"

OR

Bb. Select and Define Group Work

1. "What are the different groups in your class?"

2. "What group are you working on now?"

3. "What is your group doing now?"

4. "How did the class pick the different groups?"

INTERVIEWS QUESTIONS - LEVEL I

5. [Note: This question should only be asked depending on the answer to part 4. If the child answers part 4 with an answer such as "the teacher picked the groups," then it is not pertinent to ask the following question.]
"Who came up with the ideas about which groups to pick?"

6. "How did you end up being in this group?"

7. "Why did you pick this group, and not one of the others?"

C. Relate Current Sub-task to Problem

1. "How do you think that the work you are doing today will help with [mention USMES unit title]?"

D. Relate Contributions of Others to Problem [Note: This section should only be used if section Bb is used. If section Ba is used, ignore this section.]

1. "Can you tell me what the other groups are doing?"

2. "Why do you think it is a good idea for them to do their work?"

E. Acquire Skills and Pertinent Information

1. "Did you have to learn anything new to do this work?"

2. "Did you know how to do this kind of work beforehand?"

3. [Note: only ask this question if applicable.]
"How did you learn how to do this (saw, measure, graph, etc.)?"

F. Plan Action [Note: As for category B, section Fa should be used if the whole class is working on the USMES challenge as one unit. Section Fb should be used if the class has split up into several groups which are working independently of one another.]

Fa. Class Plan of Action

1. "How did the class decide what to do today?"

2. "Who decided what to do today?"

3. "Why did you decide to do this work?"

INTERVIEW QUESTIONS - LEVEL I

Fb. Group Plan of Action

1. "How did your group decide what to do today?"

2. "Who decided what you were going to do today?"

3. "Why did you decide to do this work?"

[Note: When asking questions from this section, we have not included a wide variety of questions pertaining to actual decisions and the work done by the individual groups (or the class) as a whole, as the variation between different USMES units is too large to write such general questions. Please try to ask the children how they made decisions within their group, and why they made these decisions to do whatever they are doing, ONLY if applicable. Please do not prompt them too much.]

G. Encounters Errors and Hang-ups

1. "Did you have any problems?"

2. "What did you do about the problem(s)?"

H. Plan Next Steps

1. "What are you going to do next?"

2. "How are you going to do this?"

3. "How did you decide to do this next?"

I. Organize, Analyze and Interpret Data

1. "What did you find out so far?"

2. "Why do you think it was a good idea to find this out?"

3. "Did you make reports of what you found out?"

K. Communicates Findings

1. "Did you tell anyone else in the class what you found out?" [If the child answers 'no,' ask: "Will you?"]

2. [Note: only ask this question if the answer to part a is "yes."] "What did you tell them?"

3. "Did any of the other groups tell you what they found out?" [If the child answers 'yes,' ask: "What?"]

APPENDIX N

PROFILES Interview Guide: Level II

for Fourth through Sixth Grades

A. Define Problem (Challenge) [Note: This category should be used during your first interview with each child, even if the problem was defined during a previous USMES class. Ask these questions again if the issue is brought up at a later class meeting.]

1. "Can you tell me what you've been doing today?" [If the child answers the question in terms of the sub-task he is working on, rather than in terms of the total problem, THEN add]: "Yes, but I was wondering if you could tell me a little bit more about [mention the title of USMES project] in general."

2. "Why do you think it is important to work on this problem?" "Project?"

3. [Note: This question should only be asked if the challenge involves a "problem"].
"Do you really think you'll be able to do anything about this problem?" "Why/Why not?"

4. "How did the class decide to work on this problem?" "Project?"

5. "Can you tell me whose idea it was to work on this problem?" "Project?"

B. Select and Define Sub-task [Note: For category B, only ask questions from Ba, or Bb. Section Ba, should be used if the whole class is working on the USMES challenge as one unit. Section Bb should be used if the class has split up into several groups which are working independently of one another.]

Ba. Select and Define Class Work

1. "Can you tell me what your class is working on right now?"

2. "How did you decide to do this work?" "Project?"

OR

Bb. Select and Define Group Work

1. "What are the different groups in your class?"

2. "What group are you working on now?"

3. "What is your group doing now?"

4. "How did the class pick the different groups?"

INTERVIEW QUESTIONS - LEVEL II

5. [Note: This question should only be asked depending on the answer to part 4. If the child answers part 4 with an answer such as "the teacher picked the groups," then it is not pertinent to ask the following question.]
"Who came up with the suggestions for the different groups?"

6. "How did you get involved in this group?"

7. "Why did you get involved in this group?"

C. Relate Current Sub-task to Problem

1. "How do you think that the work you are doing today will help with [mention USMES unit title]?"

D. Relate Contributions of Others to Problem [Note: This section should only be used if section Bb is used. If section Ba is used, ignore this section.]

1. "Can you tell me what the other groups are doing?"

2. "Why do you think their work is important?"

E. Acquire Skills and Pertinent Information

1. "Did you have to learn anything new to do this work?"

2. "Did you know how to do this kind of work beforehand?"

3. [Note: only ask this question if applicable.]
"How did you learn how to do this (saw, measure, graph, etc.)?"

F. Plan Action [Note: As for category B, section Fa should be used if the whole class is working on the USMES challenge as one unit. Section Fb should be used if the class has split up into several groups which are working independently of one another.]

Fa. Class Plan of Action

1. "How did the class decide what to do today?"

2. "Who decided what to do today?"

3. "Why did you decide to do this work?"

Fb. Group Plan of Action

1. "How did your group decide what to do today?"

2. "Who decided what you were going to do today?"

3. "Why did you decide to do this work?"

INTERVIEW QUESTIONS - LEVEL II

[Note: When asking questions from this section, we have not included a wide variety of questions pertaining to actual decisions and the work done by the individual groups (or the class) as a whole, as the variation between different USMES units is too large to write such general questions. Please try to ask the children how they made decisions within their group, and why they made these decisions to do whatever they are doing, ONLY if applicable. Please do not prompt them too much.]

G. Encounters Errors and Hang-ups

1. "Did you have any problems?"

2. "What did you do about the problem(s)?"

H. Plan Next Steps

1. "What are you going to do next?"

2. "How are you going to do this?"

3. "How did you decide to do this next?"

I. Organize, Analyze and Interpret Data

1. "What did you find out so far?"

2. "Did you keep records of what you found out?"

3. "Did you make reports of what you found out?"

J. Relate Results of Sub-task to Total Problem

1. "Okay, You found out that (repeat child's answer to 9a)." "How does this help you with [mention title of USMES unit]?"

2. "What are you going to do now?"

K. Communicates Findings

1. "Did you tell anyone else in the class what you found out?" [If the child answers 'no,' ask: "Will you?"]

2. [Note: only ask this question if the answer to part a is "yes."] "How did you explain it to them?"

3. "Did any of the other groups tell you what they found out? [If the child answers 'yes,' ask: "What?"]

APPENDIX O

PROFILES Observation Form

Name of Child: _____ Date: _____

Observer: _____ Visit Number: ____

Teacher: _____ Tape Number: ____

School: _____ Grade: _____

City, State: _____

PROFILES Observation

Child: _____ Visit #: _____ Date: _____

A. Define problem (challenge)

1. Has a problem (challenge) been defined? _____

2. What is the problem (challenge)? _____

_____

3. Who defined the problem (challenge)? _____

_____

B. Select and define sub-task

1. How is the class organized for work on the problem? _____

_____

_____

_____

_____

_____

2. What sub-task is <u>this</u> child involved in currently? _____

_____

3. Who decided what sub-task would be done by <u>this</u> child? _____

_____

4. Is <u>this</u> child doing the same thing he was during your last visit? (Answer only on second and subsequent visits.) _____

_____

Child: _____ Visit #: _____ Date: _____

C. Relate current sub-task to problem

1. How does this child's current work relate to the total problem? _____

_____

_____

_____

_____

_____

2. To your knowledge, was this relationship clarified/explained by the teacher?

_____

_____

D. Relate contributions of others to problem

1. What are others doing (give names or functions of groups, if possible)? _____

_____

_____

_____

_____

_____

2. How is the work of others (or other groups) related to the total problem? _____

_____

_____

_____

_____

_____

3. Was this relationship clarified/explained by the teacher, to your knowledge?

_____

Child: _____ Visit #: _____ Date: _____

E. Acquire skills and pertinent information

1. To your knowledge, has it been necessary for this child to acquire new skills or specific information to do his current work? _____

2. Name the most important new skills or information acquired, if applicable: _____

F. Plan action

1. What is this child doing today? _____

2. Who decided what this child would do today? _____

G. Encounter errors and hang-ups

1. Has this child encountered any significant problems with his work? _____

2. Briefly, what were the problems? _____

Who recognized these problems, to your knowledge? _____

3. Who decided what to do about these problems, to your knowledge? _____

Child: _____ Visit #: _____ Date: _____

H. Plan next steps

1. Does this child have plans for the next steps in the process, to your knowledge? _____

2. What are his plans? _____

3. Who made these plans? _____

I. Organize, analyze and interpret data

1. What records, reports, graphs or charts have been made by this child or his group, if any? _____

2. Did this child organize, analyze or interpret data himself, to your knowledge?

J. Relate results of sub-task to total problem

1. How do the results of this sub-task relate to the total problem? _____

2. Did the teacher clarify/explain this relationship, to your knowledge? _____

Child: _____ Visit #: _____ Date: _____

K. Communicates findings

1. Have any reports to the class been made by this child (or his group)? _____

2. Will any authorities outside the class be informed? _____

3. What were the final findings? _____

L. Reliability issues (Fill out after taping)

1. Describe the class environment today: _____

2. Describe the child's mood today: _____

3. From your observations, do you feel that the child understands more about the problem solving process discussed today on tape, than he was able to express?

Why? _____

4. Were there any unusual circumstances which may have affected your observation or interview today? _____

What? _____

APPENDIX P

Observer's Manual for the Administration of the PROFILES Technique

To the Observer:

This manual and the accompanying materials consist of the following:

1. Instructions to guide you in the administration of the Profiles of USMES Problem Solvers.

2. A form on which to record your observations of the children's problem solving ability.

3. A cassette tape for recording interviews with individual children.

PROFILES

General Instructions

This instrument has been developed as part of the evaluation of the USMES program. The purpose of the observations and interviews is t enable the observer to collect data on how well children grasp the process of problem solving during USMES experiences.

Profiles of USMES Problem Solvers should be administered to designed USMES classes. Two children are to be selected randomly from each USMES class in the evaluation sample. During each visit, the observer will first observe these two children working on their USMES projects, and then interview them individually afterwards, on tape.

If possible, each child should be taken to a quiet room, with no distractions, for the interview. The actual interview schedule and the role you are to play as an observer will be explained in detail shortly.

Our analysis of the Profiles will be based on two kinds of records: (a) a tape recording of the child's verbal presentation during the interview and (b) your observations of the child during the USMES project and your perceptions of the part the child plays in the problem solving process.

In general, your role as an observer will be to orgainze the interviews, to observe and record the role of each child in your sample in problem solving, and to interview each child. Specific instructions for carrying out the interviews are given in the following sections of this manual.

ORGANIZATION

1. Selection of Children

A random sample of two children should be picked from each USMES class in your sample. It is not acceptable if children are not selected randomly. When children are picked on the basis of good academic performance on the one hand, or on the basis of "getting rid of the troublemaker" on the other, the entire interview will have to be disregarded. The only two exceptions to this rule are: (1) in the case where a very shy child has been selected (this would probably only apply to the 1st and 2nd graders). If such a child either (a) seems extremely upset at the thought of being interviewed or (b) refuses to cooperate in the interview, then it would be permissible to replace this child with another, following the above procedure. (2) the child selected has a very high rate of absenteeism or must attend special classes frequently. In these cases also, it would be permissible to replace this child with another.

It would be best for you to pick the children yourself, but the teacher can also make the selections if correct procedures are used. The easiest appropriate method is to write the names of each child on a piece of paper, put each piece in a hat, and then select two.

2. When to Interview Each Child

Each time you visit the classes in your sample you should observe the two children during USMES class time for approximately 30 minutes. The length of the observation will depend on what activities the class is participating in on that particular day, how familiar the observer is with the class (or group) activities, and the amount of time it takes to fill out the observation sheets. After the observation period, each child should be interviewed individually, (one at a time) for approximately 5 to 10 minutes.

Each subject should be observed and interviewed 6 to 8 times over the course of an entire USMES unit. The length of the particular USMES unit will determine the interval between visits. Since most USMES units span from 3 to 4 months, visits once every two weeks would most likely be appropriate. Try to keep equal time intervals (approximately) between visits.

Try to conduct the interviews as close to the beginning of the day as possible. Since USMES classes take place several times a week, try to interview the children on a day when USMES is done in the morning. Also, try not to conduct interviews on the day before or after vacation periods, or on days when special school events are to take place. If children are preoccupied with other things, or are tired from a long day at school, less information will be gathered than if more optimal conditions exist. Use your own good judgement.

3. Where to Administer the Profile Interview

Each child should be taken out of the regular classroom to a quiet room with no distractions. (An empty library or cafeteria would do, if there are no empty rooms). There are 2 reasons for doing this: (1) in a noisy, busy environment, the child's attention will not be on the interview; it will be elsewhere and (2) it is very difficult to decipher what is said on the tapes when there is noise in the background. In the past, interviews have been conducted in rooms, for example, right next to a music class. If the interview is not audible, it is useless. Again, use your own good judgement in making this decision.

OBSERVATIONS

When you arrive at the classroom, you should spend approximately 30 minutes observing activities going on in the classroom. During this time period, you should determine which of the following aspects of the problem solving process each of the 2 children you are observing is involved in at the time:

1. Define problem (challenge)
2. Select and define subtask
3. Relate current subtask to problem
4. Relate contributions of others to problem
5. Acquire skills and pertinent information
6. Plan action
7. Encounter errors and hangups and reformulate plans

8. Plan next steps
9. Organize, analyze and interpret data
10. Relate results of subtask to total problem
11. Communicate findings

(This time will also serve as an opportunity for you to become familiar with the activities of the total class.) After you have determined which categories are applicable to the class you are observing, you should answer the questions only in the categories you have checked off on the Observation Form numbered A thru K. It is important for you to fill out this form before you interview the children, so their responses will not influence your answers to the questions. If it is impossible for you to completely fill out the form for both children before interviewing them (e.g., the USMES session was short on a particular day) then try to jot down as many notes as possible, and then fill it out completely as soon as the child interviews are over (but try not to let the child s responses influence yours). If you are not certain how to answer a particular question, write that down, elaborating on the problem.

You will have received intensive training in the use of this Observation Form as the Observers' Training Workshop.

Preparation for Taping

After you are done observing the class, ask one of the randomly selected children to come with you to a quiet room. Explain to the child that you are interested in what is going on in the classroom, and that you would just like to ask him a few questions about what everybody in the class is doing. Do not explain to him the nature of this evaluation.

Let the child say his name and a sentence into the microphone. Ask the child to speak distinctly. Then play the tape back to the child. This will give him a chance to get used to recording his voice, and it will give you a chance to see how well his voice is being picked up. It should put the child more at ease. After this "voice test" is done, the tape should be rewound and recorded over, so that this part of the session will be erased. When the entire session is over, we would like to have only the interview returned to us.

INTERVIEW SESSION

The interview with the child should be tape recorded in its entirety. The interview should be no longer than 15 minutes. Before starting the interview, make sure that you identify yourself, the child's name, the child's teacher, and the interview Profile number on the tape recording.

Please follow the script that will be provided to you for the interview. There are two levels of the Profile Interviews. Level I is to be administered first through 3rd graders; Level II is to be administered to 4th through 6th graders. The vocabulary on Level I is a little simpler than that for Level II.

Please follow the questions on the script very closely. Ask all the questions from each category that you have checked off during the observation period as being applicable to the class you are observing. If possible, it would be preferable to ask all the questions verbatim as they are written on the sheet, as to maintain uniformity between observers. Although you may have to ask additional questions to encourage the child to respond, or to explain what he means in greater detail, please do not stray far from the prescribed script. Above all, fo not prompt the children to say anything, since any responses from children to prompting questions will have to be ignored for the analysis. You may ask additional questions but you must ask all questions on the script.

You will have received intensive training in the using of the interview form at the Observer's Training Workshop.

After you have finished interviewing the first child, go back to the classroom and take the second child to a quiet room and follow the same procedure outlined for the first child. If one child happens to be absent on a particular day, try to get back to that classroom as soon as possible in the next few days to interview the child. It is not acceptable to interview another child instead.

After the entire interview is over, review the tape on your own. If you think any part of the conversation will be difficult for us to understand, please make a note of what was said and attach it to the observation form.

At this time, fill out category "L" on the Observer sheet.

Please be sure to return to us all tapings and observation sheets. Each interview should be sent back to us as soon as it has been completed--it would probably be easiest to send all the interviews you have done each week, back to us on Fridays. The reason for this is so that we can monitor the tapes in order to make certain that everyone is following standard procedures.

APPENDIX Q

Transcription and Analysis of Interview Tapes to Illustrate Application of the PROFILES Techniques For Assessing the Problem Solving Behaviors of Individual Children Using USMES Units

The PROFILES technique can be used for careful systematic study of how individual children progress through the process of solving real, complex problems. Observers who apply the method must be well trained in the use of these observational/interview techniques for reliable, valid results.

Many interviews were conducted in the Spring of 1975 during the pilot study of the PROFILES Interviews, in order to find out which questions elicited the best responses from children as to their understanding of the sub-tasks of the problem solving process. Three of these interviews follow in the next pages to illustrate the interviewing technique that should be used by the observer. Included after each interview is a sample of how the observation form should be filled out by the observer during the 30 minute observation and a sheet illustrating how the scoring protocol would be applied for each particular interview.

Two things should be noted during the examination of these interview samples. First, these interviews were conducted during the pilot study, and only after all of the interviews were completed, was a final script for the interviews devised, by pooling together all the best questions from various interview forms. Therefore, the 3 interviews presented here do not correspond exactly to the interview scripts included in either Appendix M or N. Several questions in these interviews are poor, but this could not be discerned until they were actually tested out. The interviews are presented here to give the reader an illustration of what the interviews would be like.

Second, one should note slaient characteristics about the sample subjects. Interview #1 was conducted with a very outgoing, verbal child in the 6th grade. His class was working on a developmental USMES unit called "Planning a Vacation." Interviews #2 and #3 were pursued with one very shy, non-verbal child in the 3rd grade. This child had been interviewed two times previous to these 2 interviews. His class was working on the USMES unit "Consumer Research." These 2 very different children were included in this section to show how well this interview script works with different types of children and different complex problems.

On the Observation Forms and Scoring Protocol Sheets included after each interview, only sections relating to components of the problem solving process applicable to the particular USMES class observed, have been filled out. By looking at the appropriate appendices, (Appendix O and R), the reader can follow the response pattern for these forms.

Interview #I - Transcription of Audio Tape

Obs. - Okay, this is Bob Farius's class, and this is Patrick who I'm speaking to now and I just wanted to know what was going on in the class. Can you tell me what everybody's doing now?

S. - Well, uh, the two classes, mine and Mrs. Serni's, are being split up into groups for different activities that we are going to do at this camp. And the reason for this is that we don't want there to be any mix-ups and we'd like you know, to keep everything in order and see who's going to do this, and who's going to do that.

Obs. - How come you're going on this camping trip?

S. - Well, mainly the 2 classes did save up and it's really a fun thing and we're going to do alot of projects like prong studies, and stuff like that.

Obs. - Who decided what you're going to do at the camp, or did you decide yet?

S. - To do at the camp?

Obs. - Um, Um, you just said that you're going to do a prong study?

S. - Well, yeah we're going to do a -- they really haven't decided, but some of the things the activity groups is saying, like canoeing, swimming, nature studies, lunch, they even had the amount of crayons, maybe, to make stuff out of it.

Obs. - And how did you decide to go on this camping trip?

S. - Well, I don't know, but a couple of years ago it all started, well, I don't know alot of it. The kids have saved up all the money by themselves.

Obs. - How did you save up all of this money?

S. - Well, help from the parents and the teachers. We had a bake sale. We had one last year.

Obs. - What kind of sale?

S. - Sale. We sold stuff. like chairs, books, anything we could our hands on. Last year we made $800.00 off it. This year I'm sure we made something like $600.00.

Obs. - Yeah, that's alot of money.

S. - From bake sales, and book sales, and everything. We just made $800.00 this year.

Obs. - So you can go on a really nice camping trip, huh? Um, what group are you working in right now?

INTERVIEW #I

S. - I'm working in the food group right now.

Obs. - Um, Um, okay, and what are the different groups that there are?

S. - There's maintenance, activities, food, first aid, and which is you know, to see who can swim, who can't. There is adults to check how many adults can go on this thing. And, uh, there is alot of other things, I can't remember all of them.

Obs. - And how did the class pick all of these different groups?

S. - Well, uh, each boy would say, or somebody, you know, or if they had an idea what we could do down here and we needed to organize ourselves and see what foods and the prices so we had to split up into groups.

Obs. - And who came up with the suggestions for all these different groups?

S. - Well, it's been going for a couple of years now and before we go, uh, Mr. Farius, I'm sure was the one who came up with it.

Obs. - I see, okay, and how did you get involved in this particular group, the food group?

S. - Well, I, we had to pick and this is the second time on the food group, and I enjoyed the first one, you see you know what foods everybody's going to eat and I just joined up with this one.

Obs. - So you just kind of volunteered yourself?

S. - Yeah.

Obs. - And why did you get involved in that group?

S. - Why? Because I was interested to see what kind of food you know, we would pick, what kind of food everybody else likes, you know.

Obs. - I see, okay. And why do you think it is important to have the food group for the camping trip?

S. - Oh, well, without it, you wouldn't know what to eat. Um, we have to buy all the stuff by ourselves, and you know, if we didn't have a food group it wouldn't be organized, and you wouldn't know what to buy, how much it would cost, what the people's choices were.

Obs. - Okay, and just for one second, could you tell me alittle bit about what the other groups are doing? And why it's important for them to be doing those things?

INTERVIEW #I

S. - Well, they're all making surveys and I'm sure the activities group came up with alot of nice things we could do on this campout, and also a school came before and we had to talk for them and there was a T.V. and they were watching us doing one of these activities and they suggested some things, and stuff like that.

Obs. - Okay, so what is, like, the equipment group doing?

S. - To see who is going to bring what tents, who's going to bring what stuff, and we have to limit, to see that nobody brings some things like, that you're not allowed.

Obs. - Okay. And which group do you think is the most important group?

S. - Well I guess they're all important.

Obs. - You don't think one's more important than all the other ones?

S. - Well they're all of the equal difference, because if you know, if you were without a food group, you wouldn't have the food, without the activities group this place would be so boring.

Obs. - Yes.

S. - To get organized.

Obs. - Yes, I guess they're all just about the same in importance. Did you have to learn any new things in order to know which foods to pick?

S. - Well, we, in math, and stuff, we learn how to do surveys and charts and graphs and that helped us out to see what people wanted what stuff, because maybe some people didn't like pancakes or were allergic to eggs and we had to get that straightened out.

Obs. - That's really good. Is there anything else you had to know about how to pick out which foods?

S. - Well the problem we might be running into with the survey is we got all the stuff and I'm sure everybody likes one of them. Like there's pancakes, eggs, bread on stick and cereal, for let's say breakfast and in the survey, people would have to pick what they wanted so we have to make certain amount so we know how much money we have saved.

Obs. - Yeah, so do you have a budget, or anything like that?

S. - Right now we haven't priced any of the food yet.

Obs. - Do you know how much money you have to spend?

INTERVIEW #I

S. - Well it's going to cost us, I don't know, maybe $500.00 just for the camp and I guess it's going to cost money for the food, and some of the activities and....

Obs. - Do you know, say, if you only have so much money....

S. - I'm sure, that's what Mr. Farias said today, there's a certain limit, you know, you can't overexaggerate the foods.

Obs. - Oh, so you don't really have so far -- you don't know how much you are going to be spending.

S. - No.

Obs. - Okay. Did you ever do this kind of work, like surveys or anything like that before you did this survey?

S. - Oh, yeah, in math, we, and I'm sure they did it last year, but I couldn't go, I would have gone, but had to leave.

Obs. - Oh, where did you go?

S. - Overseas.

Obs. - Oh, I see. You just said one problem that you had. Did you have any other problems?

S. - Well, before we ran into a couple of problems. We might be, well I don't know if we have some problems or not, we had some problems with the total surveys, if nobody wants any of the stuff we have or if they want something else to eat, and you know that's going way out of our reach, and it's a problem if somebody can't eat that stuff, what do we do then?

Obs. - So what are you going to do? Do you know?

S. - We haven't decided.

Obs. - And are you having any other problems?

S. - This really isn't a problem, but we have to get somebody, it will probably be the food group's concern to find somebody, like, one teacher and 3 adults can't cook a for something like 56 kids, so you have to get some kids, and to cook with them, and you have to tell them how to cook the stuff, you know you won't want to have scrambled eggs and they turned out, you know....

Obs. - Looking like mush.

S. - Yeah.

INTERVIEW #I

Obs. - Okay, it sounds like things aren't that bad. I just wanted to ask you a couple more questions about the actual food. How are you actually deciding which foods to pick?

S. - Mainly the survey. We surveyed before, and we came out with the people with the highest votes of food, you know like you would have something like 50 votes for one food, and 10 votes for the other, so the majority wins.

Obs. - And, so that's the only way that you pick the foods, by the survey?

S. - Well, I don't know, we haven't--I would guess so, yeah.

Obs. - Well, and are most people picking the same things so you don't have that problem picking....

S. - No, we don't have that much problem.

Obs. - Yeah, it's not like everybody says a different thing and it's kind of hard to pick out....

S. - There would be another problem though.

Obs. - What?

S. - If somebody wanted pancakes right, and we got right enough amount of pancakes for 10 people, and only 5 wanted them, and 5 wanted something else, like hamburgers, if we have hamburgers and nobody wnats them they could spoil overnight.

Obs. - So do you know what you're going to do about that?

S. - Well, my teacher suggested that we eat, if we had extra hamburgers, somebody better eat them because, you know if it spoiled, it would be no good. And a worse problem is the racoons. They could get into our food shed and devour the stuff.

Obs. - Oh. Did that ever happen before?

S. - I don't know but there were alot of racoons somebody told me last year.

Obs. - Oh, that's a big problem. So, in other words, the only thing you use to pick out which food is the answers people gave in the surveys. So how do you pick out, like you come up with some answers to the surveys, how do you pick which foods from the surveys.

S. - As I said, the majority wins.

INTERVIEW #I

Obs. - Oh, so it's just fair and square then.

S. - Yeah.

Obs. - Okay. What are you going to do next with your food group?

S. - Well, as soon as we get everything organized, we're going to buy the stuff, and if we have to make something before we go there, I'm sure we're going to fix it. And the food group, as soon as we get done we just going to have, well I'm sure they're going to pick it before, the people who are going to cook, and maybe one of the parents wants to cook if there's nothing to do and like one of the kids doesn't want to cook anymore down there, we have to find somebody else, so the food group will go to the end of the campout.

Obs. - It sounds like you still have alot of things to do.

S. - Well, we only have 3 weeks to do it in.

Obs. - That's true. So do you think that your class will really be able to plan a good camping trip?

S. - Yeah, I'm sure it's going to be a real nice campout.

Obs. - Yeah, it sounds like a really good time. Thanks alot. You were a big help.

Observation Form for Interview #1

Patrick
5/21/75
Farius
6
Adams School

Linda Hench
#1
1
Lexington, Massachusetts

A. 1. Yes.

2. Planning a vacation - they are planning a 3-day camping trip.

3. The teacher.

B. 1. The class has divided up into 6 groups.

2. The food group.

3. The child - each child in the class decided what group they wanted to be in, and if too many children wanted to be in any particular committees, then names were picked out of a hat, in order to democratically pick group members.

C. 1. His work on the food group is directly related to helping to plan a vacation, since a menu is needed to plan a vacation.

2. Yes, but the teacher has not presented this as an USMES "challenge"--it is a developmental unit.

D. 1. The 5 other groups are:
a. activities
b. transportation
c. first aid
d. maintenance
e. equipment
Functions of the groups are pretty much self-explanatory.

2. All the groups are integral in helping to plan a vacation - since each of these factors needs to be considered.

3. No specifically.

E. 1. Yes.

2. Doing surveys, graphs and tables, and how to interpret results.

F. 1. He voted for the group he wanted to be in and is now helping to plan a menu with a survey.

2. Partly the teacher, partly the group.

OBSERVATION FORM FOR INTERVIEW #I

G. 1. Yes.

2. One problem is that the group must be certain as to how much food to actually buy - otherwise they will waste alot of money. Another problem is that they must check to see if they are ordering food that everyone will enjoy eating - otherwise they might order food that some people will not eat - and it will be wasted.

3. He did, and other children in the class did.

4. It has not been decided yet.

H. 1. Yes.

2. To complete the menu by surveys, and then but the food.

3. His group.

L. 1. Normal.

2. Normal.

3. No.

4. No.

Scoring Protocol Results for Interview #I

A.	1.	1
	2.	2
	3.	4
	4.	0
B.	1.	2
	2.	2
	3.	1
	4.	1
	5.	0
C.	1.	2
	2.	1
D.	1.	1
	2.	2
E.	1.	2
	2.	1
F.	1.	3
	2.	3
G.	1.	2
	2.	2
	3.	3
	4 I.a.	2
	4 b.	2
	4 c.	6
	4 d.	6
	4II.a.	2
	4 b.	2
	4 c.	6
	4 d.	6
H.	1.	1
	2.	2
	3.	2
	4.	2
L.	1.	2

Interview #II - Transcription of Audio Tape

Obs. - Okay, this is Mrs. Meade's class and now I'm talking to Joseph. Can you tell me what you've been doing today so far?

S. - Testing the paper towels to see how strong they are.

Obs. - What were you doing to test the paper towels?.... How were you testing the paper towels?

S. - By putting weight on them.

Obs. - What kind of weights were you putting on them?

S. - A couple of erasers, the blackboard erasers.

Obs. - How did you decide to put blackboard erasers on them? I remember last time I was here you were using glue containers. How come you changed your mind and now you're using erasers?

S. - I don't know because I came in late today and I don't know what the other kids decided, how they decided.

Obs. - I see. So how many erasers are you using?

S. - Maybe 6.

Obs. - Six erasers? And how is it working out?

S. - Good, so far.

Obs. - Have you had any problems?

S. - Not today, but before.

Obs. - What were the problems from before?

S. - We were disorganized.

Obs. - Um, Um. And why were you disorganized?

S. - Because everybody was doing different things.

Obs. - What were they doing?

S. - Some were doing different things, they were with the paper towels they were testing all the other ones except they're not testing the same ones at each time.

INTERVIEW #II

Obs. - And how come they were doing that?.... You don't know?

S. - No.

Obs. - Okay. How did you find out that you were disorganized?

S. - Because we weren't doing everything, everybody was doing different things.

Obs. - And what are you going to do to help correct that problem?

S. - Do the same thing everytime, like so instead of doing all these different things, you can just do all these people on one thing at a time.

Obs. - What do you mean? Can you go into that alittle more?

S. - So if you're not doing all these different kinds of things, you're just doing one thing at a time.

Obs. - What are you talking about by "things?"

S. - The paper towels. I guess so, because last time they were doing all these different paper towels. Everybody was going out and wetting every one.

Obs. - Oh, so what do you think they should be doing instead?.... Like, what did your group decide to do instead of everybody doing something else?.... What are you doing now, in other words?

S. - Testing paper towels.

Obs. - How are you testing them, because you said before that everybody was testing different towels, and it was very disorganized. Why is it organized now?

S. - Because now, most of the people are doing one thing at a time.

Obs. - What is the one thing that they are doing at a time?

S. - They're testing only one paper towel at a time.

Obs. - And how are they testing the paper towels?

S. - Putting weights on them.

Obs. - And is it, are you having any problems with that?

S. - No.

Obs. - Everything's working out so far?

INTERVIEW #II

S. - Yeah.

Obs. - Okay. Did you have to learn any new things to work on this problem?.... Or did you know how to do this kind of work before, how did you know how to test the paper towels?

S. - We voted on every kid's idea.

Obs. - And what were the different ideas?

S. - I forget.

Obs. - You forget?

S. - Yeah, because it was about a month ago.

Obs. - Okay. Can you tell me how you think the work you're doing today on the paper towels will help with consumer research?.... Like, how do you think it is going to help solve the research?.... Like, why do you think it's any help at all?

S. - Yeah, so we can tell which one is the best.

Obs. - Do you think you will be able to tell which one is the best?

S. - I don't know.

Obs. - Okay. What are you planning on doing next?.... You're testing the paper towels now, so what are you going to do after you finish testing the paper towels?

S. - I don't know.

Obs. - You don't know? Okay. What did you find out so far about the paper towels?.... Did you learn anything yet?

S. - No.

Obs. - Nothing at all?.... Do you know why you're testing them?

S. - To see which is best.

Obs. - And why do you want to find out which is the best?

S. - So my mother can buy them.

Obs. - Okay. Thanks very much.

Observation Form for Interview #II

Joseph
4/17/75
Meade
3
Hardy

Linda Hench
#3
1
Arlington, Mass.

B. 1. 4 groups have formed: (1) peanut butter; (2) scotch tape; (3) paper towels; and (4) soda groups. Each group is investigating an aspect of the product they were assigned to.

2. He is in the paper towel group.

3. The teacher - the children wrote down their group preferences on a piece of paper, and the teacher assigned children to groups from this information.

4. Yes, he is still testing paper towels, although this time he is using erasers as weights, rather than the heavy bottles he used last time.

C. 1. He is testing different brands of paper towels for strength. From this investigation, he will find out which is the strongest paper towel, which will help him decide which is the best paper towel to buy.

2. Yes, to a degree.

E. 1. Yes.

2. How to use a stop watch.

F. 1. He is testing different brands of paper towels to see how stron ; they are. This is done by wetting the middle of the towel, placing an eraser on the towel, and then one child in the group uses a stop watch to time how long it takes for the paper towel to break.

2. The teacher went over and helped the group, but it seemed that the group made the decision as to what they would do today.

G. 1. Yes - two.

2. P #1 - The group is disorganized.
P #2 - The group was using very heavy weights for the strength test, and consequently, all the towels were breaking immediately.

3. P #1 - The teacher.
P #2 - The group.

4. P #1 - The teacher, and the group discussed the problem and came up with a solution.
P #2 - The group.

OBSERVATION FORM FOR INTERVIEW #II

H. 1. No.

I. 1. The group has just written down the time it takes each paper towel to break, on a piece of paper.

2. No.

L. 1. Normal.

2. This child does not take the interviews seriously, and is fairly uncooperative.

3. Yes, he just seems unwilling to be interviewed.

4. No.

Scoring Protocol Results for Interview #II

B.	1.	2
	2.	1
	3.	4
	4.	4
	5.	2
C.	1.	1
	2.	1
E.	1.	0
	2.	1
F.	1.	3
	2.	3
G.	1.	2
	2.	2
	3.	1
	4a.	3
	4b.	2
	4c.	3
	4d.	2
H.	1.	2
	2.	3
	3.	6
	4.	6
I.	1.	1
	2.	0
L.	1.	2

Interview #III - Transcription of Audio Tape

Obs. - Okay, now we're talking to Joseph from Mrs. Meade's class, and I was wondering if you could tell me what you were doing today in the class. When I came in the class you were doing math but you had done something before I came in, right?

S. - Yeah.

Obs. - What were you doing?

S. - We were just finishing up.

Obs. - What were you finishing up?

S. - We finished what was the best paper towel.

Obs. - How did you finish it up?.... Were you writing anything down?

S. - Yeah.

Obs. - What were you writing down?

S. - How many seconds it takes.

Obs. - And did you make--I saw you make a graph up, right?

S. - Yes.

Obs. - Okay, did you have to learn any new things to work on this problem?

S. - No.

Obs. - No? Did you know how to make a graph beforehand?

S. - No.

Obs. - So, how did you learn how to make a graph then?

S. - The teacher showed us.

Obs. - And, so do you know how to make one now?

S. - Yes.

Obs. - Okay, you do. Did you have any problems with any part of the problem about the paper towels.

S. - Yeah.

INTERVIEW #III

Obs. - What were the problems?

S. - In the beginning, we were disorganized.

Obs. - You were disorganized--I talked to you last time about that. Is it all better, did you correct the problem?

S. - Yeah.

Obs. - So you don't have any problems anymore?

S. - No.

Obs. - Okay, What are you planning on doing next?

S. - We finished today.

Obs. - Today is the last day of consumer research?

S. - Yeah.

Obs. - Okay, and what did you find out?

S. - That Bounty was the best.

Obs. - That Bounty was the best? And what were the other paper towels that you tested?

S. - Viva, Scott, and A&P.

Obs. - And was Bounty alot better?

S. - Yeah, alot better.

Obs. - Really? Do you remember how much?

S. - It was about, from the second one I think it was around, it was around 22 seconds.

Obs. - And why is it better? What were you testing to make it better? Like, what were you doing to see which one was better?

S. - The strength.

Obs. - Oh, you were seeing how strong it is. And why do you think it's a good thing to see how strong a paper towel is?

S. - So it won't break easy.

INTERVIEW #III

Obs. - And what, you were putting weights on, but do paper towels usually break?

S. - Yeah.

Obs. - Do you think it's a good thing to see if they break or not?

S. - Yeah.

Obs. - Like, when you're at home, do paper towels break alot?

S. - Yeah.

Obs. - How do they break?

S. - When you rip them off, they rip in half.

Obs. - Oh, when you're ripping them off the roll?

S. - Yeah.

Obs. - Oh, so you think if it's stronger when it's wet, it won't rip off the roll so fast. Is that what you're saying?

S. - Yes.

Obs. - Okay. And what do you think it means that you found out that Bounty was the best?.... Why was it important to find out which was the best paper towel?

S. - So you can buy it.

Obs. - So that would be the one that you would want to buy?

S. - Yeah.

Obs. - Okay, and how does what you learned today about Bounty being the best paper towel help with the problem of Consumer Research?.... Do you think it helps? Why do you think you were testing paper towels?

S. - To see which is the best.

Obs. - And why do you want to see which is best? Why do you think finding out which is the best paper towel is important because you were looking at Consumer Research?

S. - I don't know.

INTERVIEW #III

Obs. - You don't know?.... Okay, what are you going to do now? It sounds like you're all done now.

S. - Yeah.

Obs. - But did you tell or show any of the other kids whay you found out that Bounty is the best?

S. - No.

Obs. - Do you plan on telling anybody else?

S. - They already know.

Obs. - How do they know?

S. - They heard us.

Obs. - Oh, they overheard you. Do you know what the results of all the other groups are?.... You don't know what anybody else found out? Do you think that everybody is going to tell everybody else?

S. - Yeah.

Obs. - Okay, you were a big help today. Thanks alot.

Observation Form for Interview #III

Joseph
5/8/75
Meade
3
Hardy

Linda Hench
#4
2
Arlington, Mass.

B. 4. Yes.

E. 1. Yes.

2. He has learned how to construct a bar graph.

F. 1. He is constructing a bar graph to show how ling it took before each paper towel that was tested broke, under weight.

2. The teacher.

G. 1. No.

H. 1. No.

I. 1. A bar graph has been made to see how strong each paper towel is.

2. No, each child in the group made the identical graph - (copying a model graph) which the teacher showed the group how to construct.

J. 1. The group has discovered that Bounty towels are the strongest, which is an important factor to take into consideration when buying paper towels.

2. Yes.

K. 1. No.

2. The class, and parents.

3. That Bounty towels are the strongest, and it is important to remember this when buying paper towels.

L. 1. Normal.

2. A little more talkative than normal - but still reserved.

3. Yes - see previous profile.

4. No.

Scoring Protocol Results for Interview #III

B.	4.	4
	5.	1
E.	1.	0
	2.	1
F.	1.	4
	2.	4
G.	1.	1
	2.	0
	3.	0
H.	1.	2
	2.	3
	3.	6
	4.	6
I.	1.	1
	2.	2
J.	1.	0
	2.	1
K.	1.	2
	2.	1
	3.	1
	4.	2
L.	1.	2

APPENDIX R

General Instructions for Using the PROFILES Scoring Protocol

The PROFILES interviews should be analyzed in terms of several rating scales constructed for various categories of behaviors in the problem solving process. In the application of this scoring protocol, one should observe the following guidelines:

1. Score only those components that the observer has checked on the observation sheet as being relevant to a particular interview. All other components should be coded with a "9".

2. The Observer Form and the tape of the child interview should be used together as the bases to answer all questions. For example, when coding how well a child explains or defines a situation, refer back to the Observation Form for criteria defining a good explanation. In other words, make a judgement based on both sources of information.

3. The best way to score each interview would be to read the Observation Form, listen to the tape once in its entirety, and <u>then</u> score the interview by referring back to the Observation Form and replaying pertinent parts of the tape. This practice should be followed to score each group of component behaviors and each specific item as needed.

4. All of the tapes for each child must be scored sequentially since pertinent information overlaps on different interviews. The situation will be clearer to the coder if tapes are scored in this manner. Therefore, no tapes should be scored for a particular child until the child has completed the USMES unit. (A researcher may wish to monitor the conduct of observers in the interviews more frequently, however. Information lost through inadequate interview technique could not be reclaimed at the end of the child's problem activities.)

Scoring on Groups of Component Behaviors

A. Define Problem (Challenge) Problem Definition:

1. According to the observer, has a problem challenge been defined? (See observation A.1.)

   0. Observer not able to be certain.
   1. Yes.
   2. No.
   9. N/A = (category not used on this visit).

2. Child defines problem (challenge): (see observation A.2 for criteria).

   0. No response or response not relevant.
   1. Gives partial definition or mentions limited or selected aspects of the problem.
   2. Adequately defines problem in own words.
   9. N/A.

3. Decision making: According to the observer who defined the problem? (See observation A.3.)

   0. Observer not able to be certain.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   9. N/A.

4. Decision making: According to the child, who defined the problem?

   0. No response or response not relevant.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   9. N/A.

B. Select and Define Sub-Task

1. Class organization: According to the observer, how is the class organized? (See observation B.1.)

   0. No response, or response not relevant.
   1. One unit/total class.
   2. Groups.
   3. Individuals working separately.
   9. N/A.

2. Child defines or explains sub-task: (see observation B.2 for criteria).

   0. No response, or response not relevant.
   1. Mentions limited aspects of sub-task.
   2. Defines or explains sub-task clearly in own words.
   9. N/A.

3. Decision making: According to the observer, who selected the sub-task for this child? (See observation B.3.)

   0. Observer not able to be certain.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   9. N/A.

4. Decision making: According to the child who selected the sub-task?

   0. No response, or response not relevant.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   9. N/A.

5. Duration of sub-task: Has the child changed sub-tasks since the observer's last visit? (Code 0 for first visit.)

   0. Observer not able to be certain.
   1. Yes.
   2. No.
   9. N/A.

C. Relate Current Sub-Task to Problem

1. Child relates current work to problem: (see observation C.1 for criteria).

   0. No response or response not relevant.
   1. Mentions limited or selected aspects of current work or problem.
   2. Relates current work to problem in own words.
   9. N/A.

2. Teacher clarification: According to the observer, did the teacher clarify or explain the relationship between this child's current work and the problem?

   0. Observer not able to be certain.
   1. Yes.
   2. No.
   9. N/A.

D. Relate Contributions of Others to Problem

1. Child names other groups and/or functions. (See observation D.1 for criteria.)

   0. No response or response not relevant.
   1. Yes.
   2. No.
   9. N/A.

2. Child relates work of other groups to total problem (see observation D.2 for criteria).

   0. No response or response not relevant.
   1. Mentions limited aspect but does not relate work of others to total problem.
   2. Adequately relates work of others to problem in own words.
   9. N/A.

E. Acquires Skills and Pertinent Information

1. Child names acquired skills or information.

   0. No response or response not relevant.
   1. Child mentions aspect of work or problem, but not specifically new skills or information.
   2. Child names acquired skills or information
   9. N/A.

2. Child acquires skills or information: According to the observer did the child acquire new skills or information. (See observation E.1.)

   0. Observer not able to be certain.
   1. Yes.
   2. No.
   9. N/A.

F. Plan Action

1. Planning: According to the observer, who decided what this child would do on this date?

   0. Observer not able to be certain.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   9. N/A.

2. Planning: According to the child, who decided what he would do?

   0. No response or response not relevant.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   9. N/A.

G. Encounters Errors and Hang-Ups

1. Child describes problems encountered.

   0. No response or response not relevant.
   1. Child says there were no problems.
   2. Child describes problem/or problems.
   9. N/A.

2. Number of problems: According to the observer, how many significant problems did the child encounter? _____

3. Number of problems: According to the child, how many problems were encountered? _____

4. For each problem code the following four items:

   a. Decision making: According to the observer, who recognized the problem? (See observation G.2.)

   0. Observer not able to be certain.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   9. N/A.

   b. Decision making: According to the child, who recognized the problem?

   0. No response or response not relevant.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   9. N/A.

   c. Planning: According to the observer, who decided what to do about the problem? (See observation G.3.)

   0. Observer not able to be certain.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   6. Decision not yet made.
   9. N/A.

d. Planning: According to the child, who decided what to do about the problem?

0. No response or response not relevant.
1. Child.
2. Children/group.
3. Children and teacher.
4. Teacher.
5. Other adult/authority.
6. Decision not yet made.
9. N/A.

H. Plans Next Steps

1. Planning: According to the observer, does this child have plans for the next steps in the process? (See observation H.1.)

   0. Observer not able to be certain.
   1. Yes.
   2. No.
   9. N/A.

2. Child describes plans:

   0. No response or response not relevant.
   1. Mentions aspects of current work or aspects of unrelated future activities.
   2. Describes plans for next steps.
   3. No plans for the future or work is finished.
   9. N/A.

3. Decision making: According to the observer, who made plans for future steps? (See observation H.?.)

   0. Observer not able to be certain.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   6. No plans for future have been made.
   9. N/A.

4. Decision making: According to the child, who made plans for future steps?

   0. No response or response not relevant.
   1. Child.
   2. Children/group.
   3. Children and teacher.
   4. Teacher.
   5. Other adult/authority.
   6. No plans for future have been made.
   9. N/A.

I. Organize, Analyze and Interpret Data

1. Child organizes, analyzes or interprets data. Has the child been involved in these activities? (See observation I.1 and I.2.)

   0. Observer unable to be certain.
   1. Yes, this child has.
   2. Yes, his group has although he has not.
   3. Yes, his class has although he has not.
   4. No.
   9. N/A.

2. Child describes data in interview.

   0. No response or response not relevant.
   1. Gives inaccurate description.
   2. Adequately describes data in own words.
   9. N/A.

J. Relate Results of Sub-Task to Total Problem

1. Child relates results of sub-task to total problem: (see observation J.1 for criteria).

   0. No response or response not relevant.
   1. Child talks about sub-task but does not relate it to problem.
   2. Child relates results of sub-task to total problem.
   9. N/A.

2. Teacher clarification: According to the observer, did the teacher clairfy or explain the relationship between the sub-task and the total problem? (See observation J.2.)

   0. Observer not able to be certain.
   1. Yes.
   2. No.
   9. N/A.

K. Communicates Findings

1. Child communicates findings: According to the observer, has this child communicated findings to others? (See observation K.1.)

   0. Observer not able to be certain..
   1. Yes.
   2. No.
   9. N/A.

2. Child communicates findings: According to the child, has he participated in communicating findings to others?

   0. Observer unable to be certain.
   1. Yes, this child has.
   2. Yes, his group has although he has not.
   3. Yes, his class has although he has not.
   4. No.
   9. N/A.

3. Communication to outsiders: According to the child or the observer, will others outside the class be told of findings? (See observation K.2.)

   0. Observer not able to be certain.
   1. Yes.
   2. No.
   9. N/A.

4. Child communicates findings to observer: of sub-task or total problem. (See observation K.3 for criteria.)

   0. No response or response not relevant.
   1. Child talked about limited aspects of work or problem.
   2. Child adequately communicated findings to observer in own words.
   9. N/A.

L. Reliability

1. Were there any factors which made the observer feel the data for this date might be unreliable? (See observations L.1 to L.4.)

   0. Observer not able to be certain.
   1. Yes.
   2. No.
   9. N/A.