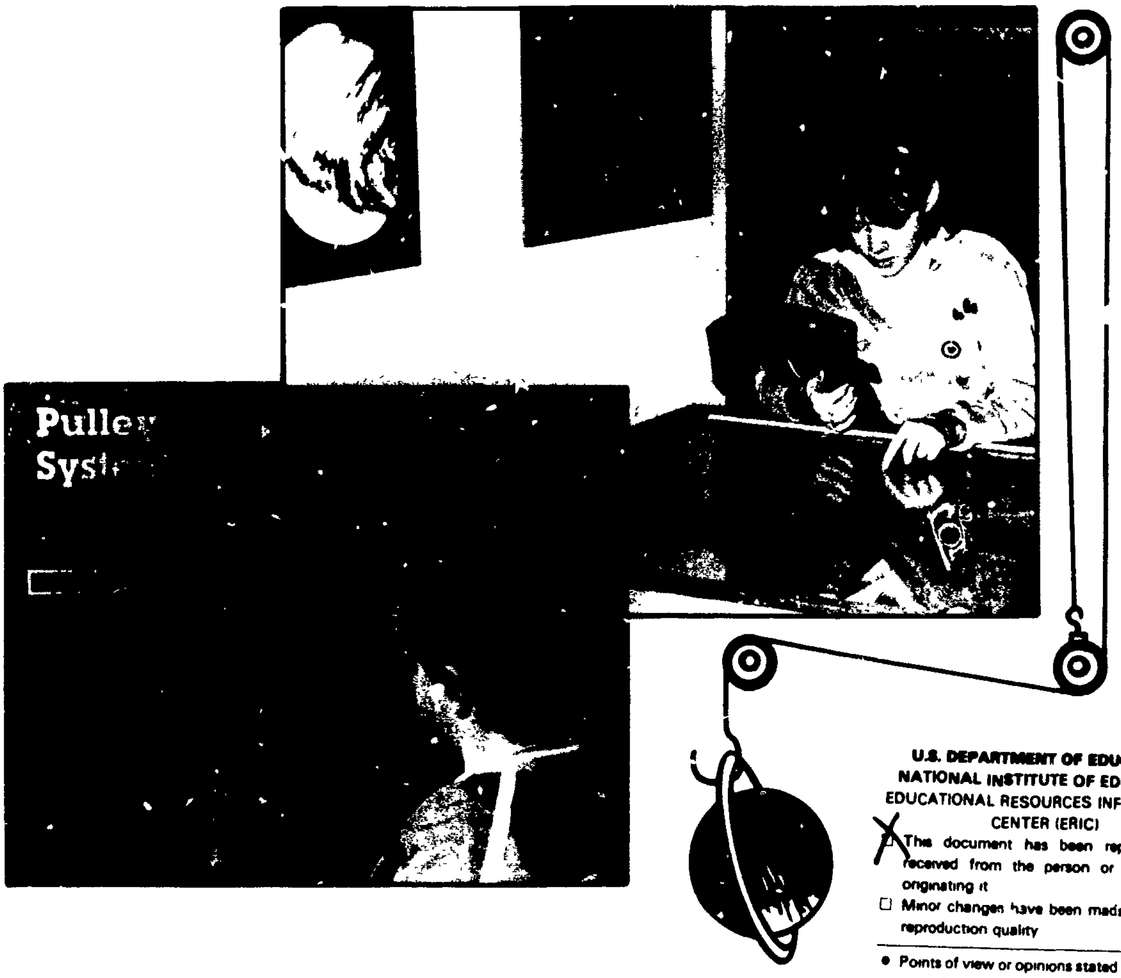
DOCUMENT RESUME

ED 267 965 SE 046 487

AUTHOR	Borun, Minda; And Others
TITLE	Planets and Pulleys: Studies of Class Visits to Science Museums.
INSTITUTION	Franklin Inst., Philadelphia, PA.; Museum of Science, Boston, MA.
SPONS AGENCY	EXXON Education Foundation, New York, N.Y.; National Endowment for the Arts, Washington, D.C.
PUB DATE	83
NOTE	149p.; Project also supported by a grant from the Guilliam H. Clamer Foundation. Some photographs may not reproduce clearly.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Elementary School Science; *Exhibits; Intermediate Grades; *Museums; Physical Sciences; Science Education; *Science Instruction; *Student Attitudes; Student Participation
IDENTIFIERS	Science Education Research.

ABSTRACT

Cognitive and affective outcomes of class visits to two participatory science museums were examined by comparing responses of fifth- and sixth-grade students randomly assigned to four conditions (control, exhibit only, lesson only, and exhibit followed by lesson) and two tests (verbal and visual). Students visiting a simple machines exhibit scored higher on a test of science than the control group, but lower than the group attending a classroom lesson in the museum. The study did not demonstrate conclusively a cognitive advantage of having the exhibit experience prior to the lesson. Scores on the visual test were consistently higher than scores on the verbal test. Study findings indicate that the particular strength of the science museum exhibit lies in the affective domain. Students found the exhibit much more enjoyable, interesting, and motivational than a classroom lesson. Implications of the findings for science education are discussed. (Supporting documentation, including verbal and visual tests, affective questionnaires, and performance tests, and lecture sheets are provided in appendices.) (Author/JN)



Planets and Pulleys

Studies of Class Visits to Science Museums

by

MINDA BORUN and BARBARA K. FLEXER
THE FRANKLIN INSTITUTE SCIENCE MUSEUM, PHILADELPHIA

and

ALICE F CASEY and LYNN R. BAUM
THE MUSEUM of SCIENCE, BOSTON

This project was supported by grants from the Exxon Education Foundation, the National Endowment for the Arts, and the Guilliam H. Clamer Foundation.

Copies may be purchased through:
The Association of Science-Technology Centers
1016 Sixteenth Street, N.W., Washington, D.C. 20036

CONTENTS

ILLUSTRATIONS

Philadelphia

Boston

TABLES

Philadelphia

Boston

FOREWORD

Museums and other informal learning centers support school science through the provision of unique resources not available in schools. These may be the collections of natural history museums, the living specimens of zoos, or the complex and imaginative exhibits of science centers.

Science museums, in particular, provide access to educational technologies that schools cannot afford or have not yet installed. From planetariums to microcomputers, science education equipment is often found in museums long before it reaches schools. Further, a trained educational staff can bring students into contact with scientists and introduce them to the techniques and processes of industry. Museums present objects, experiments, and concepts in tangible form. For students learning science, they represent the ideas of science in a physical form that allows active experimentation and investigation parallel to that of science itself.

What is needed is a new and far-reaching partnership between schools and museums. Only through careful study and research will we have the information necessary to develop joint programs which fully utilize the unique strengths of each institution.

Joel N. Bloom
Vice President and Director
The Franklin Institute Science Museum

PREFACE

The investigations discussed in this report represent a foray into a little known realm of museum research. Conducting experiments in museums, as in other field settings, is challenging. Experimental control is difficult to achieve without changing visitors' behavior or overly detracting from their museum experience. Moreoever, generalizability of research findings is restricted by the diversity of exhibits, even in the same museum. Nevertheless, if museums are to be effective centers of informal education, answers to basic questions about communication by means of exhibits need to be found. Hence, it is hoped that the research reported here will inspire further investigations.

Minda Borun
Director of Museum Research
The Franklin Institute Science Museum

ACKNOWLEDGMENTS

Philadelphia

We want to thank Dr. Rochel Gelman, Department of Psychology, University of Pennsylvania, who served as a consultant for this project and upon whose expertise we relied heavily throughout the long course of this investigation. Additional project consultants include Dr. Matthew Bruce, Department of Secondary Education, Temple University, and June Hyatt and Elizabeth Van Patten, of Hyatt Esserman Research Associates, who assisted in research design and instrumentation.

Conducting an experimental research study in a museum setting with visiting school children as subjects was a complex and ambitious endeavor which would not have been possible without the assistance of a great many members of the staff of The Franklin Institute Science Museum. We owe a debt of appreciation to J. Newlin, Director of Exhibits, whose "Mechanics" exhibit was the focus of this study and who served as a science content expert for the project.

Special thanks are owed to Patty DiTomaso, Ann Linton, Sherry Peck, and Mickey Sok of the Education Department for their consistent support and valuable help with preliminary planning, piloting of test instruments, and implementing the experimental procedure. We particularly appreciate the outstanding efforts of Mary Hardwick, who presented the classroom science lesson on simple machines with unflagging verve and enthusiasm to literally dozens of experimental groups. Mary also provided illustrations for the visual test and assisted with all phases of this project.

In addition, we are grateful to Gaye Slagle for word processing support and generous help whenever asked, and to Bill Rooney for his keen eye and excellent editing.

Various stages of this study involved the cooperation of many other members of the museum staff who so willingly gave us the assistance we needed. They include Kim Casper, Donna Claiborne, Kim Coleman, Etta Goodwin, Jack Hoffacker, Mildred Jones, John Milner, Sheryl Sankey, Lisa Scaven, Jim Singer, and Charles Van Horn.

Boston

We acknowledge with gratitude the invaluable assistance of the following persons:

Staff of the Education Department who helped generously and competently with all phases of the project.

William Kochanczyk, Director of Education, who provided administrative support essential to research efforts being conducted within an educational program.

Dr. Richard J. King, former Program Director of the Museum of Science, who assisted with all aspects of this study of museum learning, an area of inquiry of special interest to him and to which he has made substantial contributions.

Volunteer visitor guides who directed students on testing days efficiently and hospitably and with great patience.

Project consultants -- Dr. Sheldon White, Harvard University; Dr. George Hein, Lesley College; Stephen Baum, Systems Programmer, Kurzweil Computer Company; and Terry Kwon, former Coordinator of Science, Brookline Public Schools. They shared with the project staff their wisdom and experience as well as their understanding of, and concern for, learning in a museum.

SUMMARY

Participatory science museums provide unique learning opportunities not available in most schools. They display memorable visions of real objects and provide experiences whereby students can actively participate in the discovery of scientific principles. Thus, teachers often turn to museums to supplement their classroom science instruction. However, to date there is a paucity of experimental studies which investigate the outcomes of a school visit to the science museum.

The current study, conducted jointly by The Franklin Institute Science Museum and the Boston Museum of Science, examined the learning that takes place on a school trip to a science museum and investigated the facilitative effect of the visit on classroom learning. At both sites, cognitive and affective responses of middle school students randomly assigned to three treatment groups and a control group were compared. One treatment group visited a museum exhibit, one group attended a classroom lesson covering the same content, and one group attended the lesson following a visit to the exhibit. The control group was tested before having had either museum experience. Furthermore, each museum used verbal, visual, and performance tests, so that the relative effectiveness of the three types of tests for measuring museum-based learning could be assessed. Although the basic design was the same at The Franklin Institute and the Museum of Science, there were important methodological differences at the two sites.

The Philadelphia Study

Five participatory devices in The Franklin Institute's Mechanics Hall were selected as the focal exhibit. These devices, the most dramatic of which is the "Giant Lever", were designed to teach the concepts of three simple machines -- lever, inclined plane, and pulley -- by allowing visitors the opportunity to manipulate the devices, observe the effects of their actions under varied conditions, and form conclusions about the principles involved. Thus, kinesthetic learning was the emphasis of the Philadelphia component of the study.

The investigation at The Franklin Institute consisted of two separate experiments. The sample for the first experiment was comprised of 416 fifth- and sixth-grade students randomly assigned to four experimental groups (Control, Exhibit, Lesson, and Exhibit/Lesson) and two cognitive tests (Verbal and Visual). Both the Verbal and Visual Tests consisted of 10 multiple-choice items designed to measure the recognition, comprehension, and application of the three simple machines presented in the focal exhibit and the lecture written for the classroom lesson. A five-item affective questionnaire was also administered to students in Experiment 1 so that the enjoyment, interest, and motivation provided by the two treatments could be compared.

For the second experiment, 104 fifth graders were randomly assigned to a control group and the same three treatments described above. This time, an individually administered performance test was used to assess students' ability to apply simple machine concepts to three-dimensional materials, demonstrate and explain the basis of the mechanical advantage of each machine, and name the simple machines.

The analyses of data from both experiments indicate that, as hypothesized, students learn from the focal exhibit. Regardless of which type of instrument was used (i.e., verbal, visual, or performance), fifth and sixth graders visiting the simple machines exhibit scored significantly higher on a test of science content than students in the control group. More specifically, it was found that, although the Simple Machines Exhibit is not as effective a brief learning experience as a classroom lesson in conveying science vocabulary or suggesting applications, it teaches science concepts as well as a didactic lesson does.

Although scores on the Verbal and, particularly, the Performance Test reflected a tendency for students having both museum experiences to score higher than the other treatment groups, neither of the experiments conducted at The Franklin Institute provided statistically significant evidence of the motivational impact of the exhibit on subsequent classroom learning. Perhaps this can be explained by the fact that the lesson on simple machines was a substantive counterpart of the exhibit. That is, since a portion of the science content was effectively taught by the exhibit, the exhibit's effect on learning from a subsequent lesson conveying the same content would likely be obscured.

The experiments provided some interesting information regarding the effects of testing mode on test performance. The results of Experiment 1 indicated that the Visual Test was significantly easier across treatment groups than the Verbal Test. In fact, scores of the Lesson Group were so high on the Visual Test that there was little room for improvement by a group having both treatments. In addition, the analysis of Performance Test data (Experiment 2) showed that test scores of students visiting the exhibit compared more favorably with those of students attending the lesson than they did on either the Verbal or Visual Tests. Thus, the Performance Test, which provided the closest match to the exhibit experience, proved to be the most sensitive measure of exhibit-based learning.

Responses to the Affective Questionnaire indicated that the exhibit was perceived as more enjoyable and interesting than students' usual school classes or a museum lesson, even when presented by a dynamic educator in a novel environment. More importantly, it was found that the students viewed the exhibit as a learning experience and were interested in learning more about the science content it presented.

Finally, a survey of fifth- and sixth-grade teachers who brought their classes to The Franklin Institute revealed that their primary motivation for the visit was science education. Questionnaire data further indicated that teachers' educational expectations are generally being fulfilled.

The Boston Study

A temporary exhibit entitled "Planets and Moons" was selected for the focus of investigation at the Museum of Science. This highly visual exhibit presented planetary and lunar features and particularly stressed four major themes -- rings, craters, volcanoes, and turbulence.

The sample for the Boston study consisted of 535 sixth- and seventh-grade students who were randomly assigned to the four experimental groups and three tests (Verbal, Visual, and Performance). The tests used by the Museum of Science were open-ended and divergent. Moreover, they were designed to measure the influence of the treatments on perceptions and images and the ability to make connections and predictions, rather than the acquisition of science facts or concepts.

The Verbal Test was an essay test with five questions, each of which was phrased in such a way as to suggest no single correct answer. Students were put in the role of space travelers and asked to speculate about the features on or around planets. Analysis of scores on the Verbal Test yielded no significant differences among means for the four experimental groups.

The Visual Test was a picture-sorting activity whereby students were asked to sort a set of 13 pictures into two piles according to their own criteria. The procedure was repeated to produce a second set of responses. The set of pictures included whole planets and moons, parts of planets and moons, and specific features of planets and moons. The influence of the exhibit and/or classroom lesson on Visual Test performance was assessed by tallying the frequency with which categories used for picture sorting corresponded to the four main themes (i.e., rings, craters, volcanoes, and turbulence) of the exhibit and lesson. Although this test was tied in most closely with the design of the visually-oriented exhibit, there were no significant differences among experimental groups in the mean number of sorting categories which represented main themes.

For the Performance Test, students were provided with masonite boards covered with a thin layer of clay, some aluminum foil, additional clay, and coffee stirrers. They were asked to make four features that might be seen on or around planets. Again, answers that related to the four main themes were identified. The analysis of Performance Test data revealed that only students in the Lesson Group created significantly more main-theme features than students in the Control Group. Though not significant, a trend toward a positive effect of the exhibit on students' performance on this test was reflected by the number of main-theme responses for the Exhibit and Exhibit/Talk Groups.

Thus, the cognitive findings of the Boston study were sparse. No doubt the nature of the testing instruments themselves provides an explanation. The open-ended Verbal Test and broadly defined Visual Test apparently lacked the structure needed to tap the images and mental categories conveyed by a museum exhibit. The more focused Performance Test, however, provided at least some support for a notion that perceptions and images acquired by student visitors to a museum exhibit can be manifested through testing.

Affective data collected at the Museum of Science did, however, provide clear-cut evidence of the value of a museum exhibit in teaching science. A majority of students who visited the focal exhibit reported that they enjoyed the exhibit, found it interesting, and wanted to learn more about planets and moons. Furthermore, just as in the Philadelphia study, students perceived a discovery type of visit to a science exhibit as a learning experience meriting comparison with a school lesson.

Conclusions

The results of the joint study conducted by The Franklin Institute and the Museum of Science indicate that although science museum exhibits can teach science, the greater strength of the museum experience is in the affective domain. Both investigations provide conclusive evidence that students visiting a science museum on a class trip enjoy the visit, find the exhibits interesting, and are motivated to learn more about the science content that is demonstrated or displayed. Even more importantly, it is clear that middle school students perceive their museum visit as a learning experience.

The findings of the present study have important implications for science education. Teachers who turn to a participatory science museum as a resource to supplement classroom learning will no doubt find that some learning does take place on a school trip. But more importantly, they can be assured that the science museum experience will stimulate an interest in learning science by presenting science content in a manner that students find exciting. Thus, whether used as an introduction to a classroom unit or as a culminating activity, a visit to a science museum can effectively complement classroom instruction. It is appropriate to conclude, then, that a science museum, by providing exciting exhibits which generate enthusiasm for and interest in learning science, can serve as a valuable adjunct to formal science education.

INTRODUCTION

Each year millions of school children visit museums on class field trips. Their teachers use such trips to complement classroom instruction, assuming that museums will provide learning opportunities not available in schools. Museums have the potential for serving as a valuable adjunct to formal education. However, it is necessary to develop an understanding of the educational impact of a museum visit before teachers are in a position to make optimal use of the museum as a supplemental resource.

Participatory science museums, in particular, provide visual and kinesthetic learning experiences that are qualitatively different from the experiences associated with classroom lessons or printed text. Science museums offer memorable three-dimensional visions of real objects and opportunities for students to actively explore demonstrations of scientific principles. The importance of such multisensory learning experiences has long been stressed by advocates of "discovery learning." For example, Inhelder and Piaget (1958) pointed to the child's action upon objects in the environment as an integral, even necessary, part of the process of cognitive development. Later, Jerome Bruner (1961), a seminal educator, extolled the virtues of a discovery approach to learning. He maintained that a child who finds things out for him/herself not only gains a deeper understanding of the principles involved but is more satisfied and motivated than a child who is taught didactically. Notions such as these led to the development of materials-based science programs (e.g., Science Curriculum Improvement Study, 1974) and an increase in hands-on activities in many science classrooms.

Science museums with their interactive devices and numerous opportunities for discovery may be seen as the epitome of the materials-centered program. While it might be expected that school trips to science museums have positive outcomes in both the cognitive and affective domains, little experimental evidence for museum-based learning is available in the educational research literature. Although Borun (1977) showed that casual visitors to The Franklin Institute Science Museum made measurable cognitive gains, it remained to be seen whether similar outcomes are derived from school class visits.

Studies conducted at the Lawrence Hall of Science in Berkeley, California, provide some evidence of learning through school trips. Eason and Linn (1976) compared the cognitive test responses of fifth to eighth graders who visited two types of interactive optics exhibits to responses of control subjects who had not been to the exhibits. They found that school students' knowledge of optics was improved by a visit to a participatory exhibit, regardless of exhibit format. Interestingly, the size of the treatment effect differed as a function of the type of question which was used. That is, a marked improvement was observed when a question was presented in diagram rather than written form.

Also at Lawrence Hall, Sneider, Eason, and Friedman (1979) compared high school students randomly assigned to visit the Star

Games exhibit to a control group on measures of cognitive learning, psychomotor skills in using telescopes, and attitudes toward astronomy materials. Their study revealed that the exhibit was an effective learning tool, as assessed by both an astronomy quiz and a psychomotor measure, but they did not find the hypothesized increase in subjects' selection of astronomy books for later reading, which was intended to demonstrate the motivational impact of the exhibit.

While there is at least some support for the notion that children learn on school trips to a science museum, it is necessary to turn to other than participatory science centers to find experimental evidence that museum-based learning has an impact on regular classroom learning. One such study conducted at the Cincinnati Museum of Natural History (Lambert, 1978) assessed the impact on fourth-grade students of supplementing a classroom unit on prehistoric Indians with a program conducted at the museum and/or a traveling exhibit lesson. Each type of museum lesson (on-site and traveling) was shown to enhance performance on a multiple-choice test. Further, the group which had both treatments had the highest scores.

More recently, Wright (1980) examined the effect of a multisensory experience at the Kansas Health Museum on the comprehension and application of human biology concepts taught during a six-week classroom unit. While sixth-grade classes in the experimental group spent three hours at the museum visiting the Transparent Lady, two films, and an exhibit on human body systems, the control group had three hours of traditional classroom review. Museum activities were found to contribute to significantly better achievement on the unit test.

The Lambert and Wright studies suggest that museum-based learning can reinforce classroom learning. Neither study exposed a treatment group to a museum exhibit prior to a classroom lesson to look at the possible cognitive and motivational impact of the exhibit on subsequent classroom learning. Further, with the exception of the Eason and Linn (1976) study, none of the research cited above considered the relative effectiveness of various types of instruments for assessing museum-based learning. Yet it has been suggested that typical classroom tests which are highly verbal in nature may not be appropriate for measuring the visual or kinesthetic learning that occurs during a visit to a science museum (see Borun, 1977; Kimche, 1978).

A review of relevant studies clearly reflects a need for further investigation of museum-based learning -- a need which is frequently discussed in museum literature (e.g., Kimche, 1978; King, 1981; Wright, 1980). The present research, conducted as a joint project of The Franklin Institute and the Boston Museum of Science, was an attempt to answer the following important research questions:

1. What are the cognitive outcomes of a school visit to a museum exhibit?
2. What are the affective outcomes of a school visit to a museum exhibit?
3. Is classroom learning facilitated by a visit to a museum exhibit?

4. Does measurement of museum-based learning depend on a match between the nature of the learning experience and the test mode?

The study was designed to compare the cognitive and affective responses of middle school students who visit a museum exhibit, students who attend a classroom lesson covering the same content, and students who attend the lesson following a visit to the exhibit. A control group was included in the design to provide an indication of baseline knowledge. In order to assess the relative effectiveness of various instruments for measuring museum-based learning, three types of tests -- verbal, visual, and performance -- were employed.

Although the same questions were addressed at both The Franklin Institute Science Museum and the Boston Museum of Science, there were methodological differences at the two sites due to differences in focal exhibit, museum audience, and educational philosophy. These differences are as follows:

1. The sample was comprised of fifth and sixth graders at The Franklin Institute and sixth and seventh graders at the Museum of Science.
2. The focal exhibit was essentially manipulative at The Franklin Institute and visual at the Museum of Science.
3. Testing instruments at The Franklin Institute were highly structured and objective, while those at the Museum of Science were open-ended.
4. At The Franklin Institute, two separate experiments are described. The experimental procedures were implemented

in the fall for the verbal and visual tests and in the spring for the performance test. At the Museum of Science, all data were collected in the fall.

It was felt that this variation would permit a broader base from which to draw conclusions. Further, since there has recently been much discussion concerning the relative merits of objective testing versus testing with an open-ended response approach (see King, 1981), it was hoped that the differences in testing instruments cited in Number 3 above would shed some light on questions of instrumentation.

Because of the methodological differences listed above, reports of the investigations conducted at each museum are presented separately in this monograph.

THE PHILADELPHIA STUDY

The study conducted at The Franklin Institute was one attempt to measure and document the educational outcomes of a school trip to a science museum. The basis for the Philadelphia study is participatory or kinesthetic learning. Accordingly, five participatory devices in Mechanics Hall served as a focal exhibit for testing purposes. These devices were designed to teach basic facts and concepts in Physics by allowing visitors to participate actively in the learning process. That is, visitors can manipulate the devices, observe the effects of their actions under varied conditions, and form conclusions about the principles involved.

Study samples at The Franklin Institute consisted of fifth- and sixth-grade students for several reasons. First, preliminary investigation of science curricula in the Philadelphia area indicated that by selecting fifth and sixth graders, a substantial number of subjects would not have previously learned in school the physics concepts conveyed by the exhibit. Second, given the science concepts and the level of their presentation in the focal exhibit, early fifth grade was found to be an appropriate lower limit. Finally, museum reservation records indicated that there were sufficient registrations of fifth- and sixth-grade classes among the regularly scheduled school groups to satisfy the sample requirements.

Since the intent of the Mechanics Exhibit is to teach basic science facts and concepts, it was considered appropriate to use tests of science content to assess cognitive outcomes of a visit to this exhibit. Furthermore, since preliminary interviews of fifth and sixth graders in Mechanics Hall revealed their inability to verbalize the principles which were demonstrated, it became apparent that, in order to tap the subjects' competence, tests would have to stimulate recognition or recall of these science concepts. Accordingly, the verbal and visual tests designed for the Philadelphia Study are multiple-choice in format and responses are recorded with paper and pencil. The performance test, on the other hand, employs three-dimensional objects and questions are open-ended. Nonetheless, all three tests are highly structured and yield quantitative scores for museum learning.

For practical reasons pertaining to the size of the sample, the number of experimental groups, and the need to minimize testing time, the study at The Franklin Institute consisted of two separate experiments. The first experiment, conducted in the fall, was designed to assess the museum-based learning of a large sample of fifth and sixth graders and to examine the effect of test mode (verbal vs. visual) on their test performance. The same experiment provided the context for the collection of affective data. The second experiment, conducted the following spring, focused on cognitive outcomes of a museum visit, assessed by a performance test individually administered to a smaller group of fifth-grade students.

The following research hypotheses were tested:

1. Students visiting an exhibit will score significantly higher on a test of science content than students in a control group.

2. Students will perceive an exhibit as significantly more enjoyable, interesting, and motivating than a classroom learning experience.

3. Students attending a classroom lesson following a visit to a museum exhibit will score significantly higher on a test of science content than students attending only the lesson.

4. Students visiting a museum exhibit will score higher on a test which closely resembles the learning experience than on a traditional paper-and-pencil test.

METHOD

Experiment 1

Sample Selection

In order to solicit their students' participation in the study, a randomly-selected group of fifth- and sixth-grade teachers who had made reservations for class trips to The Franklin Institute Science Museum for the month of November 1981 was contacted by telephone. A free readmission pass for each student was offered as compensation. Of the teachers who were approached, all of those whose museum schedules allowed time for the experimental procedure agreed to participate.

Data were collected from 432 fifth- and sixth-grade students from nine schools. Four were suburban public schools, three were suburban parochial (Catholic and other) schools, and two were urban parochial schools.[1]

---

[1]No Philadelphia public school groups could be included because of a long-term teachers' strike.

Experimental Procedure

Students in each participating school group were randomly assigned to one of four experimental groups and one of two tests. Color-coded stick-on labels were used to designate group membership and numbers on the labels indicated which test the student would take (i.e., verbal or visual).

Following a brief orientation, subjects were directed to join their group leaders (museum educators), who escorted them through the appropriate sequences of activities. Although all experimental groups had an equivalent set of experiences in the Museum, the order was varied. Events which were experimentally irrelevant for a particular group occurred after the tests had been administered. The sequence for each group was as follows:

	Time 1	Time 2	Time 3
Control Group:	Test	Exhibit	Lesson
Exhibit Group:	Exhibit	Test	Lesson
Lesson Group:	Lesson	Test	Exhibit
Exhibit/Lesson Group:	Exhibit	Lesson	Test

Following the third activity, groups were escorted to a prearranged meeting place in order to rejoin their teachers.

Exhibit. A group of five manipulative displays in the Simple Machines section of the Mechanics Exhibit was chosen as the focus of the study. The displays, the most dramatic of which is the "Giant Lever", afford visitors an opportunity to experience the mechanical advantage provided by a lever, inclined plane, and pulley (see Figure A-1).

Since the selected displays are located in Mechanics Hall, a very large room which houses many other manipulative displays of appeal to middle school children, a portion of the room was roped off. In addition, gold stars were affixed to the five displays used in the study.

At the entrance to the exhibit, subjects were asked to be sure to see each of the five displays with a gold star. No science instruction of any kind was provided by the group leader. Subjects were permitted to spend up to 15 minutes in the exhibit area.

Lesson. The Simple Machines Lecture was written to convey the same scientific concepts demonstrated by the exhibit. There was particular concern for keeping the content and vocabulary at a fifth-grade level[1] with the exception of necessary technical terms (i.e., lever, fulcrum, inclined plane, pulley, and mechanical advantage). The lecture (as presented in Appendix A-1) was edited and approved by a panel of museum educators.

In order to maintain procedural consistency, the lesson was conducted in a museum classroom by the same museum educator each time. The lecturer made use of a few small, common demonstration materials which would be available to any classroom science teacher (see Figure A-2). No question-and-answer period was permitted.

Instruments

The data-collection instruments for each subject were assembled into a booklet which was color-coded to indicate his/her

[1]Botel's "Readability Levels" (1962) was used as a vocabulary guide.

Figure A-1. Simple Machines Exhibit

Giant Lever

Lever

Pulley

experimental group. One researcher administered all tests. The testing procedure took approximately 15 minutes per group. Subjects were instructed to answer every item in their booklets and were encouraged to ask for help in reading any words which they found difficult (see Figure A-3).

Demographic Data Sheet. Subjects were asked to provide the following information: sex, grade, age, and school. In addition, they were to indicate whether or not they had studied simple machines in school and to report the number of times they had previously visited the focal exhibit. (See Appendix A-2 for a copy of the Demographic Data Sheet.)

Affective Questionnaire. This short questionnaire was designed to assess affective outcomes of the museum experiences. Specifically, subjects were required to evaluate the exhibit and/or lesson in terms of their overall reaction, enjoyment, interest, extent of learning, and motivation for future learning.

The form of the questions differed somewhat for experimental groups. Whereas the Exhibit and Lesson Groups were asked to compare the museum experience to regular school classes, the subjects who had received both treatments were asked to compare the exhibit to the lesson. Subjects in the Control Group did not complete an affective questionnaire. (See Appendix A-3 for copies of the Affective Questionnaires.)

Cognitive Tests. Two forms of a simple machines test (i.e., verbal and visual) were developed to assess the cognitive outcomes of a visit to the focal exhibit and/or attendance at the lesson on simple machines. Since visiting school groups had a limited amount of time to devote to the study, an effort was made to minimize testing time and yet construct measures which are sufficiently valid and reliable for research purposes.

The content validity of potential items was established by a panel of four science educators. Readability level of each item was kept at or below fifth-grade level[1] with the exception of the names of the simple machines and their main components. Furthermore, each version of the test was piloted with fifth- and sixth-grade students. On the basis of the analyses of pilot data, items were revised if they were shown to be unclear or ambiguous. Other items were eliminated if they did not discriminate well between high and low scorers on the total test.

1. Verbal Test. The Verbal Test consists of 10 four-option multiple-choice items. Items were designed to measure the students' recognition of vocabulary and comprehension and application of the operational concepts of three simple machines (i.e., lever, inclined plane, and pulley). For example, the following items were used to measure the recognition, comprehension, and application, respectively, of the inclined plane:

Which one is an inclined plane?

A. a swing
B. a seesaw
C. a jungle gym
D. a sliding board

[1]Since the formulas traditionally used for assessing the readability level of text are not appropriate for multiple-choice items, the vocabulary level of the tests was controlled by using Botel's (1962) guidelines.

Figure A-2. Presenting the Lesson

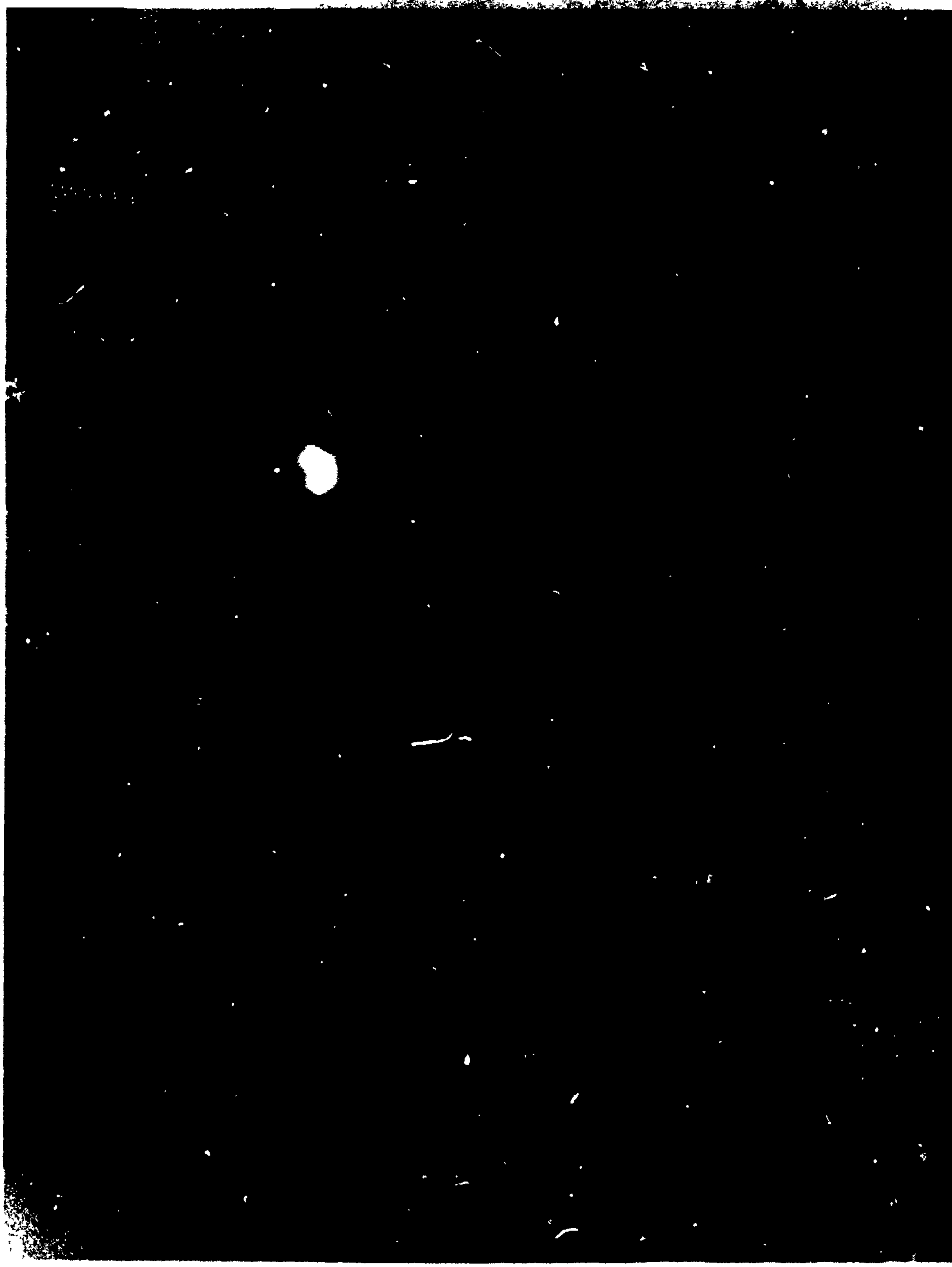
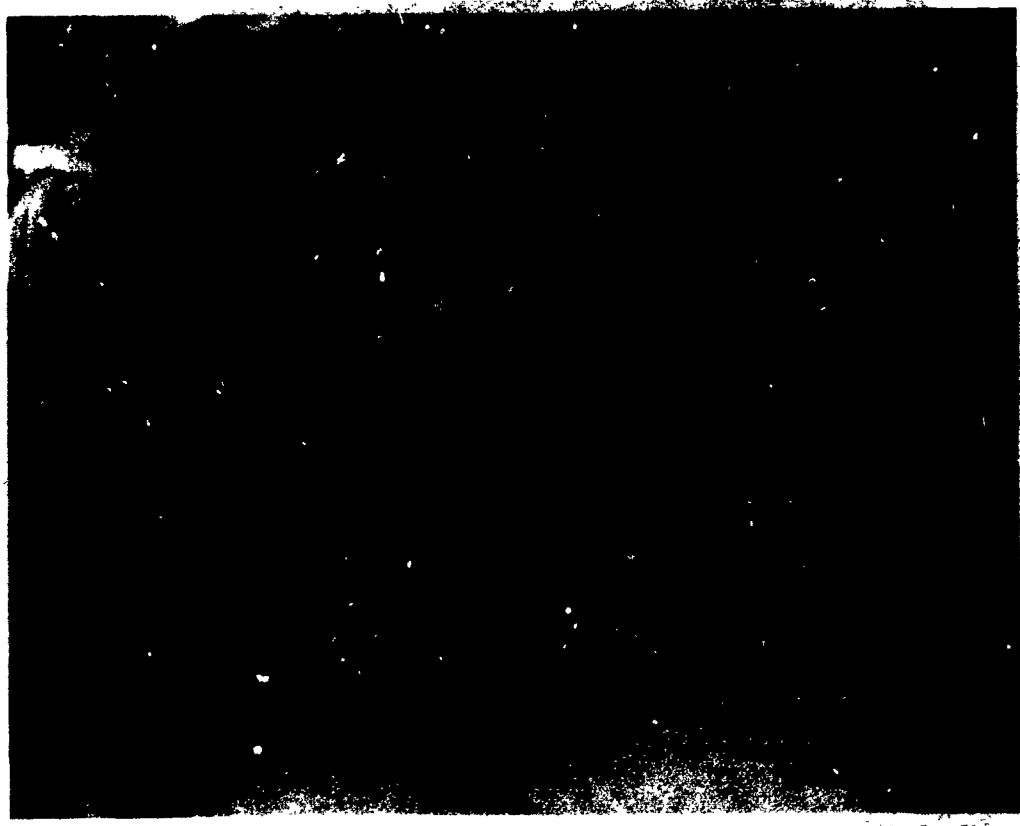


Figure A-3. Testing -- Verbal and Visual



It is easiest to pull a weight up a ramp if the ramp is

A. very steep.
B. sort of steep.
C. hardly steep at all.
D. The steepness has nothing to do with it.

A heavy box must be moved from the floor to a platform. Which way of moving the box would take the least force?

A. picking it up and placing it on the platform
B. tying a rope around it and pulling it up over the side of the platform
C. picking it up and carrying it up steps leading to the platform
D. tying a rope around it and pulling it up a long board leading to the platform

The instrument includes parallel sets of questions for the lever and pulley and an additional item requiring the recognition of a simple machine from a group that includes complex machines. Subjects respond to each item by circling the letter in front of the selected response. The score for each subject is the number of items answered correctly. (See Appendix A-4 for a copy of the Verbal Test.)[1]

2. Visual Test. The 10 items on the Visual Test are parallel in content to those on the Verbal Test. However, the stem of each multiple-choice item contains a minimal number of words. For example, the last item cited above appears as follows in the Visual Test: "Which picture shows the way of raising the box that takes the least force?" The optional responses are then presented pictorially. Thus, the test provides a learning

---

[1]The internal consistency reliability coefficient (K-R Formula 20) computed from test data for the present sample was .50. While not high, such a reliability coefficient is considered satisfactory for assessing group differences and is, in fact, typical of reliabilities for tests with only 10 items.

assessment which is far less dependent on linguistic ability. (See Appendix A-5 for a copy of the Visual Test.)[1]

Experimental Design

A 4 (Treatment) X 2 (Test) factorial design was employed to analyze the data from Experiment 1.

Sample Description

The final sample consisted of 416 fifth- and sixth-grade students.[2] Specifically, each of the eight cells in the experimental design included 26 subjects (13 boys and 13 girls) from each grade.

Analyses of demographic data indicated that the groups were not significantly different in terms of chronological age[3] (see Table A-1) and the number of subjects reporting previous visits to the exhibit (see Table A-2). On the other hand, it can be seen that groups taking the Verbal Test differed significantly in the proportion of subjects reporting having studied simple machines in school, with the greatest discrepancy between percentages for the Exhibit and Lesson Groups (see Table A-3).

---

[1]The internal consistency reliability coefficient (K-R Formula 20) computed from test data for the present sample was .67.

[2]Sixteen students from whom data were collected were randomly eliminated from the sample in order to achieve an equal distribution of sexes and grades in each cell of the design.

[3]$F < 1$, $df = 3/408$, ns.

However, this difference would not have had a significant effect on group scores, since statistical analyses showe that test scores are not related to having studied the topic in school. Further, analysis of individual responses revealed that the dis-crepancy in reporting having studied the subject in school may be due to the fact that, unlike students in the Lesson Group, students who visited the exhibit tended not to recognize "simple machines" as the same topic they studied in school.

TABLE A-1

Means and Standard Deviations of Age in Years for Experiment 1

	Group							
	Control		Exhibit		Lesson		Exhibit/Lesson	
Test	M	SD	M	SD	M	SD	M	SD
Verbal	11.18	(.79)	11.02	(.63)	11.03	(.75)	11.10	(.74)
Visual	11.05	(.75)	11.06	(.82)	11.23	(.82)	11.05	(.66)

Note: N = 52 in each experimental group.

TABLE A-2

Frequencies and Percentages of Reports of Previous Visits to the Exhibit

Test	Control f	Control %	Exhibit f	Exhibit %	Lesson f	Lesson %	Exhibit/Lesson f	Exhibit/Lesson %	χ^2
Verbal	29	56%	31	60%	25	48%	25	48%	ns
Visual	30	58%	28	54%	27	52%	29	56%	ns

Note: N = 52 in each experimental group.

TABLE A-3

Frequencies and Percentages of Reports of Previous Study of Simple Machines in School

Test	Control f	Control %	Exhibit f	Exhibit %	Lesson f	Lesson %	Exhibit/Lesson f	Exhibit/Lesson %	χ^2
Verbal	24	46%	16	31%	33	64%	26	50%	11.31*
Visual	18	35%	25	48%	26	50%	29	56%	ns

Note: N = 52 in each experimental group.

*$p < .05$.

Experiment 2

Sample Selection

Data for Experiment 2 were collected from 111 students in four classes randomly selected from fifth-grade groups with reservations for school trips to The Franklin Institute Science Museum during the month of April 1982. Since it was the second half of the school year, this sample was at approximately the midpoint between the two age groups tested in Experiment 1.

Experimental Procedure

The procedure for randomly assigning subjects to experimental groups and the sequence of activities for each group were the same as described for Experiment 1. However, since there was no need to divide subjects into two test groups for this experiment, individual subject ID numbers, rather than test numbers, were written on the color-coded labels.

Exhibit. The five displays in the Simple Machines Exhibit described earlier were again the focus of investigation. The exhibit experience was the same as that described for Experiment 1.

Lesson. The Simple Machines Lecture described earlier was delivered by the same museum educator who participated in Experiment 1. However, for Experiment 2, drawings of simple machines, rather than small three-dimensional objects, were used to complement the lecture. This change was dictated by a need to use different modalities in the lesson and the test in order to avoid bias in favor of the Lesson Group. Thus, for Experiment 1, which involved a picture test, three-dimensional demonstration objects were used; whereas for Experiment 2, which was an object-manipulation test, pictures were used. Minor revisions were made in the text of the lecture in order to accommodate this change.

Instruments

Performance Test Materials. In Experiment 2, three-dimensional manipulative materials were used to assess the cognitive outcomes of a visit to the focal exhibit and/or attendance at the lesson on simple machines. The testing materials were presented at three stations (i.e., one for each simple machine). Test materials at each station included two miniature men ("Smurfs"), a 3 1/2 inch wooden cube which served as a platform, a weight with a hook on it, and three examples of one of the three simple machines. The machines were constructed as follows:

- Lever -- One metal fulcrum stand and three wooden lever arms which varied in length, but each of which had a hook for the weight located the same distance from the fulcrum attachment
- Inclined Plane -- Three solid wooden ramps, all 3 1/2 inches high, but varied in slope and, hence, in length
- Pulley -- Three pulley systems, each attached to a separate ring stand, as follows: one fixed pulley (one rope), one fixed pulley and one movable pulley (two ropes), and two fixed pulleys and one movable pulley (three ropes)

All materials were housed in large open-sided cardboard cartons which served as testing carrels (see Figure A-4).

Figure A-4. Testing -- Performance

Inclined Plane

45

Lever

Pulley

Testing Procedure. At each testing session, all subjects proceeded from station to station in the same order. However, the locations of the three simple machines were randomized for the four test sessions.

At each station the following problem situation was posed: "This little man (a 'Smurf') has to raise this heavy weight to the top of the platform. He is too little to do it without some help. Show me how the Smurf could use one of these to make the job easier." A subsequent open-ended question -- "Which one would make it easiest to lift the weight?" -- assessed the subjects' ability to apply simple machines concepts to three-dimensional materials. A further question -- "Why did you choose that one?" -- required subjects to explain the basis of the mechanical advantage of each of the machines. Finally, students were asked to name the simple machines. Responses were recorded by the examiner assigned to each testing station. (See Appendix A-6 for copies of the data sheets.)

Scoring. All data sheets were scored by the same researcher. One point was awarded for each correct response. Thus, there was a potential total test score of 12 points (i.e., four points for questions pertaining to each of the three simple machines).[1]

Experimental Design

Since all subjects participating in Experiment 2 took the same test, there was only one independent variable (Treatment)

---

[1]The internal consistency reliability coefficient (K-R Formula 20) computed from test data for the present sample was .70.

comprised of four levels (Control, Exhibit, Lesson, and Exhibit/Lesson).

Sample Description

The final sample for Experiment 2 consisted of 104 fifth-grade students. Each of the experimental groups included 14 boys and 12 girls.[1]

Mean chronological ages for the four groups were not significantly different (see Table A-4).[2] Furthermore, the overall mean age (11.09 years) was the same as that for the entire sample in Experiment 1.

TABLE A-4

Means and Standard Deviations of Age in Years for Experiment 2

Group							
Control		Exhibit		Lesson		Exhibit/Lesson	
M	SD	M	SD	M	SD	M	SD
11.00	(.40)	11.09	(.47)	11.10	(.50)	11.16	(.46)

[1]Seven students from whom data were collected were randomly eliminated from the sample in order to achieve the same proportion of boys and girls in each treatment group.

[2]$F < 1$, $df = 3/100$, ns.

Teacher Survey

A Visiting Teacher Questionnaire was used to gain insights into museum-based learning beyond those yielded by Experiments 1 and 2.

Sample

The questionnaire was mailed to 64 fifth- and sixth-grade teachers who had visited The Franklin Institute Science Museum with their classes in Fall 1981. A reminder letter was mailed to non-respondents two weeks later. A total of 32 teachers returned completed questionnaires.

Instrument

The Visiting Teacher Questionnaire was designed to collect information about teachers' motivations for scheduling class trips to The Franklin Institute and to elicit their observations of the educational outcomes of their students' visits. An open-ended format was used in order to encourage comments outside the domain of obvious responses. (See Appendix A-7 for a copy of the questionnaire.)

Data Analysis

Data from the Visiting Teacher Questionnaire were analyzed qualitatively. That is, teachers' responses to the questions were summarized and used to supplement experimental findings.

RESULTS

Cognitive Outcomes

To evaluate how effectively a manipulative museum exhibit can convey scientific ideas, several types of analysis were used. First, total scores on the Simple Machines Tests were analyzed by means of analysis of variance. Second, proportions of subjects exhibiting a high level of competence were compared across experimental groups to provide a qualitative assessment of museum learning. Third, subjects' subscores for groups of related items on the Verbal and Visual Tests were examined to provide information about the specific nature of the learning that takes place in the museum exhibit. Finally, teachers' survey responses describing cognitive outcomes of their visits were summarized.

Experiment 1: Verbal and Visual Tests

Analysis of Total Test Scores. Means and standard deviations of total scores on the Verbal and Visual Tests, separately and for both tests combined, for subjects in each of the four experimental groups are presented in Table A-5.

Total test scores were analyzed by means of a 4 (Treatment) X 2 (Test) analysis of variance. The results of the analysis can be summarized as follows:

1. Experimental groups differed significantly in their performance on the Simple Machines Tests.[1] Subsequent comparisons[2] indicated that the mean score of the Exhibit Group,

---

[1] $F = 26.40$, $df = 3/408$, $p < .001$.

[2] The Newman-Keuls Test was used to make pairwise comparisons.

while significantly higher than that of the Control Group, is significantly lower than that of the Lesson Group. Furthermore, the mean score of the Exhibit/Lesson Group does not differ significantly from that of the Lesson Group.

2. Subjects taking the Visual Test scored significantly higher than subjects taking the Verbal Test.[1]

3. Mean scores of the treatment groups were not differentially affected by the type of test (i.e., verbal vs. visual).[2]

TABLE A-5

Means and Standard Deviations of Total Scores on the Verbal and Visual Tests

	Group							
	Control		Exhibit		Lesson		Exhibit/Lesson	
Test	M	SD	M	SD	M	SD	M	SD
Verbal	4.87	(1.72)	5.77	(1.89)	6.25	(2.06)	6.70	(2.01)
Visual	5.79	(2.29)	6.94	(1.92)	8.29	(1.38)	7.98	(1.57)
Both Tests	5.33	(2.06)	6.36	(1.99)	7.27	(2.02)	7.34	(1.90)

Note: Maximum score on each test = 10.

[1] $F = 54.57$, $df = 1/408$, $p < .001$.

[2] $F = 1.71$, $df = 3/408$, ns.

In order to assure that the experimental results were not affected by the discrepancy reported earlier in the number of subjects per cell who indicated that they studied simple machines in school, data were analyzed separately for subjects who did and did not report having a unit on simple machines in school. Both analyses yielded results which are consistent with those for the entire sample.

It can be seen, then, that a visit to the Simple Machines Exhibit <u>is</u> a learning experience for fifth- and sixth-grade students. Moreover, the finding that the lesson was, over all, a more effective brief learning experience than the exhibit is not particularly surprising. The lecture is concise, well written, and was delivered by an exceptionally effective museum educator. Furthermore, discovery learning such as that taking place in the exhibit is apt to be a much slower and more subtle experience (cf. Guthrie, 1967) than the transmission of information through an organized lecture.

On the other hand, it is surprising that subjects attending the lesson after a visit to the exhibit did not score significantly higher on the cognitive tests than subjects having only the lesson. However, an examination of means for each test separately (see Table A-5) yields some additional insights not revealed by the analysis of variance. Scores on the Verbal Test <u>are</u> somewhat higher for the Exhibit/Lesson Group than for the Lesson Group. However, a slight trend in the <u>opposite</u> direction can be seen in the means for the Visual Test. This may be due to the fact that there is evidence of a ceiling effect on the Visual

Test; the mean score for the Lesson Group is so high that there is little room for improvement. These trends are more clearly seen in the mastery-level analysis to follow.

A lesser dependence on reading ability may explain the finding that the Visual Test yielded a higher overall mean score than the Verbal Test. Finklestein and Hammill (1969) similarly found that fifth graders performed significantly better on a pictorial version of a science inventory than on a reading-based version of the same test.

Finally, since scores on the Visual Test are consistently higher across treatment groups than scores on the Verbal Test, Experiment 1 did not provide evidence of a link between learning mode and testing mode. Nevertheless, it is important to keep in mind that the Simple Machines Exhibit can more appropriately be classified as manipulative than visual. Thus, the Visual Test does not really match the learning mode.

Analysis of Mastery-Level Achievement. This analysis focused on the quality of individual achievement on the Verbal and Visual Tests rather than on mean behavior of groups. For the analysis, a subject was classified as having mastered the science ideas conveyed by the exhibit and/or lesson if he/she responded correctly to at least seven of the ten items on the test. Seven was chosen as the criterion in order to exceed the level at which a score would be likely to result from guessing.[1]

[1]For a 10-item, 4-option multiple-choice test, the probability of attaining a score of seven or more by guessing alone is less than .C03.

Frequencies and percentages of subjects in each experimental group scoring at a mastery level on the Verbal and Visual Tests are presented in Table A-6.

TABLE A-6

Frequencies and Percentages of Mastery on the Verbal and Visual Tests

	Group							
	Control		Exhibit		Lesson		Exhibit/Lesson	
Test	f	%	f	%	f	%	f	%
Verbal	8	15%	18	35%	20	38%	27	52%
Visual	19	37%	29	56%	47	90%	45	86%
Both Tests	27	26%	47	45%	67	64%	72	69%

Note: N = 52 in each group.

Although the general pattern of results is the same as that shown by the analysis of variance, the percentages in Table A-6 reflect clearer evidence of two trends reported earlier. Specifically, Verbal Test data provide some evidence of the impact of the exhibit experience on subsequent classroom learning. While 38% of subjects in the Lesson Group had mastery-level scores on the Verbal Test, 52% of subjects in the Exhibit/Lesson Group achieved the same level. With regard to the Visual

Test, the ceiling effect mentioned before is even more pronounced. Specifically, 90% of the subjects in the Lesson Group achieved test scores of seven or more.

Analysis of Subscores. In an attempt to delve further into the nature of the learning that takes place in the exhibit, test items were grouped into the following categories:

Vocabulary -- items requiring recognition of a lever, a pulley, and an inclined plane (i.e., Items 3, 4, and 6)

Concept -- items assessing knowledge of the principle of each simple machine (i.e., Items 7, 9, and 10)

Application -- items describing a practical application of each of the simple machines (i.e., Items 2, 5, and 8)

Each category, then, is comprised of three items, one for each of the simple machines which are the focus of the study.[1]

Each of the three sets of subscores was analyzed by means of a 4 (Treatment) X 2 (Test) analysis of variance. As was the case with total test scores, significant treatment and test effects were revealed by each analysis, but mean scores of the treatment groups were not differentially affected by type of test.

---

[1]To maintain symmetry of categories, Item 1, which deals with the meaning of the term "simple machine", was not included in the present analysis.

Further analyses f significant treatment effects for Vocabulary[1], Concept[2], and Application[3] subscores revealed that the pattern of means differs somewhat according to the type of test item. (See Table A-7 ı . group subscore means.)

TABLE A-7

Means and Standard Deviations of Subscores on the Verbal and Visual Tests

	Group							
	Control		Exhibit		Lesson		Exhibit/Lesson	
Category	M	SD	M	SD	M	SD	M	SD
Vocabulary	1.85	(.96)	1.91	(.96)	2.29	(.73)	2.32	(.78)
Concept	1.31	(.86)	1.86	(.84)	1.94	(.95)	1.87	(.89)
Application	1.60	(.92)	1.82	(.96)	2.16	(.93)	2.30	(.86)

Note: Each category includes three items.

In the case of Concept items, all three treatment groups scored significantly higher than the Control Group and their means are approximately the same. Thus, the analysis of Concept subscores provides empirical evidence that the exhibit is as effective as the lecture in teaching the principles of the lever, inclined plane, and pulley.

---

[1] $F = 8.78$, $df = 3/408$, $p < .001$.

[2] $F = 13.98$, $df = 3/408$, $p < .001$.

[3] $F = 13.09$, $df = 3/408$, $p < .001$.

Additional scrutiny of individual test items revealed that the lever principle is the concept most effectively conveyed by the exhibit. To understand why, one needs only observe activity in Mechanics Hall for a few minutes. The Giant Lever is clearly the focal point of the large exhibit area. It is particularly appealing to visiting school groups who obviously enjoy experiencing the mechanical advantage provided by a lever in lifting a 500-pound weight with their own bodies. More importantly, children learn the principle conveyed by the Giant Lever without needing to rely either on printed operating instructions or explanations (neither of which is available in this display).

With respect to Vocabulary and Application items (see Table A-7), on the other hand, only the lesson, alone or in combination with the exhibit, was found to be effective in producing significantly higher scores than those for the Control Group. This is not surprising in view of earlier research (Borun & Miller, 1980) which has shown that very few young visitors (under age 19) read museum labels. Therefore, it is readily understood that, in the absence of any other form of linguistically-based information (such as oral instruction), fifth and sixth graders in the Exhibit Group did not learn to identify the simple machines by name.

With respect to Application items, the focal exhibit (unlike the lesson) does not present information on everyday applications. The absence of such suggestion is especially evident for the devices constructed to teach the concepts of the inclined plane and the lever. Moreover, the levers and inclined plane found in the exhibit do not even resemble commonly-used forms of these simple machines; thus, it is understandable that the students had difficulty relating the exhibit devices to real effort-saving situations.

Experiment 2: Performance Test

Analysis of Total Test Scores. Means and standard deviations of total scores on the Performance Test for each of the experimental groups are presented in Table A-8.

TABLE A-8

Means and Standard Deviations of Total Scores on the Performance Test

Group							
Control		Exhibit		Lesson		Exhibit/Lesson	
M	SD	M	SD	M	SD	M	SD
5.27	(1.95)	7.[illegible]9	(2.53)	7.38	(2.19)	8.42	(2.60)

Note: Maximum score = 12.
N = 26 in each group.

A one-way analysis of variance performed on test scores showed a significant treatment effect.[1] Pairwise comparisons of the means[2] indicated that subjects in each of the three treatment groups scored significantly higher on the Performance Test than subjects in the Control Group. Unlike in Experiment 1, means for

[1] $F = 8.27$, $df = 3/100$, $p < .001$.

[2] Newman-Keuls Test

the Exhibit, Lesson, and Exhibit/Lesson Groups are not significantly different from one another.

Thus, the exhibit was shown, again, to be effective in conveying science content. In fact, test scores of subjects visiting the exhibit compare more favorably with scores of subjects attending the lesson than they do on either the Verbal or Visual Tests. This finding can be interpreted as evidence that a match between learning and testing modes is important, since the Performance Test, which provides the closest match to the exhibit experience, proves to be the most sensitive measure of exhibit-based learning.

In addition, the Performance Test data reflect a tendency (though not statistically significant) for subjects experiencing both the exhibit and lesson to attain higher test scores than subjects attending only the lesson.

Analysis of Mastery-Level Performance. For a qualitative comparison of the Performance Test results across experimental groups, a score of nine out of 12 possible points (75%) was selected as the criterion for mastery.

Frequencies and percentages of subjects in each treatment group achieving a mastery level are shown in Table A-9. An examination of the table reveals that the overall pattern of treatment effects on test performance is essentially the same as that for the analysis of total test scores. Again, there is evidence that students learn from the exhibit and that there is a small, but positive, impact of a prior visit to the exhibit on subsequent learning from the lesson.

TABLE A-9

Frequencies and Percentages of Mastery

on the Performance Test

Group							
Control		Exhibit		Lesson		Exhibit/Lesson	
f	%	f	%	f	%	f	%
1	8%	8	31%	10	38%	13	50%

Note: N = 26 in each group.

A second analysis of mastery was suggested by the Performance Test data. Unlike the Verbal and Visual Tests, which consist entirely of multiple-choice items requiring recognition of correct responses, the Performance Test includes open-ended questions which demand explanations of the principles of the simple machines. These items are not only difficult to answer by guessing, but actually test the central message of the exhibit and/or lesson. Therefore, a comparison was made across experimental groups of the proportion of subjects who displayed competence by correctly describing the principles of at least two of the three simple machines.

The number and percentage of subjects in each experimental group who met the criterion described above are shown in Table A-10. It can be seen that the proportion for the Exhibit Group compares favorably with that for the Lesson Group and that the proportion for the Exhibit/Lesson Group is highest of all. This analysis provides the most convincing evidence that prior experience with a hands-on museum exhibit stimulates later classroom learning.

TABLE A-10

Frequencies and Percentages of Correct Explanations for at Least Two of Three Science Principles

Group							
Control		Exhibit		Lesson		Exhibit/Lesson	
f	%	f	%	f	%	f	%
4	15%	10	38%	9	35%	15	58%

Note: N = 26 in each group.

Comparison of Results of the Two Experiments

Although Experiment 1 (Verbal and Visual Tests) and Experiment 2 (Performance Test) data were not collected from the same sample, or even samples of the same size, it is interesting to compare the results of the two investigations. For this comparison, scores on the three tests were converted to percentages. Mean scores on the three tests for the four experimental groups are presented in Table A-11.

A close examination of the table yields a number of observations which are worthy of comment. First, although one might expect Control Group scores to be lowest on the Verbal Test because of its reading demands, it can be seen that they are lowest on the Performance Test. The explanation probably lies in the fact that the tests differ not only in response mode (verbal vs. performance), but also in the level of difficulty of the required response. While the Verbal Test consists of multiple-choice items requiring only recognition of the correct answer and allowing for successful guessing, the Performance Test questions are open-ended, requiring that students recall and state the correct answer.

TABLE A-11

Means of Percentage Scores on All Three Tests

	Group			
Test	Control	Exhibit	Lesson	Exhibit/Lesson
Verbal	49%	58%	63%	67%
Visual	58%	69%	83%	80%
Performance	44%	60%	61%	70%

Note: N = 52 for Verbal and Visual Test groups.
N = 26 for Performance Test groups

Second, Table A-11 provides evidence that the Performance Test is more sensitive than the other two tests in tapping learning from the exhibit. Not only is the difference between scores for the Control and Exhibit Groups greatest on the Performance Test, but subjects in the Exhibit Group did very nearly as well as subjects in the Lesson Group on this test.

Third, the combined effect of the two treatments on Performance Test scores is clearly apparent here and would most likely have shown a significant increase over groups having just one treatment had the sample size been larger.

Teacher Survey

Teachers' responses to the questionnaire clearly indicate that their primary motivation for bringing classes to the Museum is science education. In fact, when asked to list and rank order reasons for class visits, 66% of the teachers who responded cited science learning as most important. Those teachers frequently reported that the visit was a culminating activity for a particular science unit. One teacher, on the other hand, wrote that the Museum "provided an apperceptive base for future lessons." Furthermore, while 38% of the responding teachers commented on the variety of learning experiences provided by the Museum, 34% specifically mentioned the "hands-on" nature of science museum learning as a reason for the visit.

A large proportion of teachers felt that their goals for the visit were, indeed, fulfilled. Educational outcomes cited by the teachers support this perception. The reinforcement of science concepts was most often listed (50% of respondents) as an outcome of the museum visit. Exposure to new science concepts was another frequently mentioned outcome. The only unfulfilled expectations reported were two teachers' disappointment in not being able to use the Observatory because of inclement weather.

When asked to list specific student behaviors which demonstrated educational outcomes, teachers most often mentioned that museum learning was seen in contributions to class discussion (55%) or questions students asked in class (21%). Some teachers also cited their students' written reports, science projects, and even art projects as indications of museum-based learning.

Teachers were asked about their classroom activities in preparation for the museum visit. Many respondents described pre-trip activities involving an orientation to the museum and suggestions about what to see. This is consistent with the findings of a recent study at the Chesapeake Bay Nature Center of the effects of three types of advance preparation for a school visit. In this study, Falk (1982) found that orientation to the site is most effective.

Almost the entire sample (94% of responding teachers) followed up their museum visit with classroom activities. Most frequently the follow-up activity was an oral discussion about the students' experiences in the Museum; however, some teachers assigned written reports about the visit. A few mentioned that a science unit was initiated as an outgrowth of the museum experience. One teacher specifically listed a unit on the circulatory system which grew out of the children's walk through the Heart exhibit. A particularly powerful endorsement for museum learning came from the teacher who reported that the class he took to the Museum for a visit to the Weather Station "performed better on a meteorology test than other groups who did not have the museum lesson."

In summary, then, it can be seen that the teacher questionnaire data clearly complement the cognitive findings of the experiments conducted at The Franklin Institute. Teachers do bring their classes to the Museum primarily to learn science and, in fact, to complement school science lessons. Both teacher reports and experimental data indicate that their educational expectations are being met.

Affective Outcomes

The Affective Questionnaire, which was completed by all subjects in Experiment 1 except those in the Control Group, provides valuable information regarding students' attitudes toward the exhibit and/or lesson, their perception of the activities as learning experiences, and their motivation to learn more about simple machines.

Overall Reaction

Subjects expressed their reaction to the exhibit and/or lesson by selecting the cartoon face and descriptive wording which best indicate how much they liked the activities. Percentages of subjects choosing each response are presented in Table A-12.

The data were analyzed by assigning numerical values from 5 for "I love it" to 1 for "I hate it." When mean ratings for the Exhibit and Lesson Groups were compared, it was found that, although the reaction to both activities was very positive, the exhibit ($M = 4.30$) was significantly better liked than the lesson ($M = 3.70$).[1] A similar analysis was used to compare reactions to the exhibit and lesson, respectively, among subjects who received both treatments. Again, the mean rating is significantly higher for the exhibit ($M = 4.22$) than the lesson ($M = 3.48$).[2] It is interesting to note that, although the exhibit was almost equally rated by subjects in the Exhibit and Exhibit/Lesson Groups, the reaction to the lesson was somewhat less positive among subjects

---

[1] $t = 5.13$, $df = 206$, $p < .001$ (independent-samples t test).

[2] $t = 9.17$, $df = 103$, $p < .001$ (correlated-samples t test).

who experienced both activities and thus had a basis for comparison.

TABLE A-12

Overall Reaction

	Group			
	Exhibit	Lesson	Exhibit/Lesson[1]	
Response			Exhibit	Lesson
I love it	41%	15%	40%	8%
I like it	48%	56%	45%	48%
I don't like or dislike it	10%	18%	12%	35%
I dislike it	1%	5%	2%	4%
I hate it	0%	6%	1%	6%

[1]Subjects in this group reacted to both exhibit and lesson.

Note: N = 104 in each group. Therefore, frequencies are approximately equivalent to the percentages.

Finally, reference to Table A-12 reveals that the difference in overall reaction to the treatments is most dramatically seen in the selection of the cartoon face with the broadest smile. While 41% of subjects in the Exhibit Group reported that they loved the exhibit, only 15% of those in the Lesson Group felt the same way about their activity. Furthermore, the difference is even more striking in the responses of the subjects who had both treatments. Although 40% reported that they loved the exhibit, only 8% reported loving the lesson.

Enjoyment

The second item of the Affective Questionnaire assesses the enjoyment provided by the exhibit and lesson. The nature of the question differs according to whether subjects received one or both treatment(s). Subjects in the Exhibit and Lesson Groups compared their enjoyment of the museum activity to regular school classes, whereas subjects receiving both treatments compared the exhibit to the lesson.

Percentages of subjects in the Exhibit and Lesson Groups choosing each response are presented in Table A-13. An analysis of item responses indicates that when compared to school classes, the exhibit is perceived as significantly more enjoyable than the lesson. Mean ratings (on a 5-point scale) for the Exhibit and Lesson Groups are 4.47 and 3.46, respectively.[1] Further reference to the table reveals that a total of 87% of subjects in the Exhibit Group, as compared to 51% of subjects in the Lesson Group, found the activity more enjoyable than school classes.

---

[1] $t = 2.97$, $df = 206$, $p < .01$ (independent-samples t test).

TABLE A-13

Enjoyment

Comparison of Museum Activity to School Classes	Group	
	Exhibit	Lesson
Much more enjoyable	68%	24%
A little more enjoyable	19%	27%
About as enjoyable	6%	29%
A little less enjoyable	5%	10%
Much less enjoyable	2%	10%

Note: N = 104 in each group.

The responses of subjects in the Exhibit/Lesson Group, who compared the exhibit to the lesson, indicate that 75% enjoyed the exhibit more, 5% enjoyed the lesson more, and 20% considered them equally enjoyable.[1] Thus, the data provided by the Exhibit/Lesson Group reflect even more clearly the greater enjoyment afforded by the exhibit.

Interest

On the third item, subjects in the groups receiving one experimental treatment compared the exhibit or lesson to their school classes in terms of interest. Percentages of subjects selecting each response are shown in Table A-14.

A comparison of mean ratings (on a 5-point scale) indicates that the exhibit ($M = 4.29$) compares significantly more favorably

---

[1]chi square = 81.51, df = 2, $p < .001$.

with school classes than the lesson (M = 3.31).[1] Moreover, as seen in Table A-14, a total of 80% of subjects in the Exhibit Group, as compared to 47% of subjects in the Lesson Group, found the activity more interesting than school classes.

TABLE A-14

Interest

Comparison of Museum Activity to School Classes	Group	
	Exhibit	Lesson
Much more interesting	57%	14%
A little more interesting	23%	33%
About as interesting	13%	31%
A little less interesting	6%	13%
Much less interesting	1%	9%

Note: N = 104 in each group.

The Exhibit/Lesson Group data provide even stronger evidence of the effectiveness of the exhibit in stimulating interest. While the exhibit as considered more interesting by 81% of the subjects who had both treatments, the lesson was selected by only 10%, and the activities were considered of equal interest by 9% of the sample.[2]

---

[1] $t = 6.53$, $df = 206$, $p < .001$ (independent-samples t test).

[2] chi square - 102.27, $df = 2$, $p < .001$.

Extent of Learning

In order to determine the extent to which the exhibit and lesson are perceived as learning experiences, subjects in the Exhibit and Lesson Groups were asked to compare their museum activity to their school classes. Percentages of subjects selecting each response are shown in Table A-15.

TABLE A-15

Extent of Learning

Comparison of Museum Activity to School Classes	Group	
	Exhibit	Lesson
Learned much more	25%	18%
Learned a little more	41%	31%
Learned about the same	24%	39%
Learned a little less	8%	7%
Learned much less	2%	5%

Note: N = 104 in each group.

The analysis of item responses (on a 5-point scale) revealed that the exhibit (M = 3.80) compares significantly better than the lesson (M = 3.51) with school classes in terms of the amount that subjects felt they had learned.[1] Furthermore, reference to Table A-15 shows that a total of 66% of subjects in the Exhibit Group, as compared to 49% of subjects in the Lesson Group,

[1] $t = 2.02$, $df = 206$, $p < .05$ (independent-samples t test).

responded that they learned more from the museum activity than they learn in school classes.

Subjects in the Exhibit/Lesson Group, on the other hand, did not see a similar difference in the instructional effectiveness of the two treatments. Specifically, 37% reported learning more from the exhibit, 36% from the lesson, and 27% considered them equivalent.[1]

It is important to note here that, although cognitive test scores tend to show more learning from the lesson, many subjects perceived themselves as learning as much, or even more, from the exhibit. Therefore, a class visit to a participatory science exhibit is clearly considered by the students to be a learning experience and not merely a recreational activity such as visiting a playground.

Motivation for Additional Learning

As a measure of motivation, subjects in each of the three treatment groups were asked to report how interested they would be in learning more about simple machines. Percentages of subjects giving each response are presented in Table A-16.

Item responses (on a 4-point scale) were analyzed by means of a one-way analysis of variance. The results revealed that the groups differed significantly in their motivation for future learning.[2] Specifically, the mean for the Exhibit Group ($M = 3.27$) is significantly higher than that for the Lesson ($M = 2.94$) or the Exhibit/Lesson ($M = 3.00$) Groups. Moreover, reference to

---

[1] chi square = 1.81, df = 2, ns.

[2] $F = 4.88$, $df = 2/309$, $p < .01$.

Table A-16 indicates that the percentages of subjects who were very interested in learning more about simple machines are 43% for the Exhibit Group, 27% for the Lesson Group, and 28% for the Exhibit/Lesson Group.

TABLE A-16

Motivation for Additional Learning

Extent of Interest in Additional Learning	Group		
	Exhibit	Lesson	Exhibit/Lesson
Very interested	43%	27%	28%
Somewhat interested	44%	50%	46%
Not very interested	9%	12%	24%
Not at all interested	4%	11%	2%

Note: N = 104 in each group.

Interestingly, the motivational impact on future learning of the combined treatments was no more than that of the lesson alone and less than that of the exhibit alone. Thus, somewhat of a satiation effect seems to be in evidence. That is, there may be a tendency for students who have encountered the information in two ways not to feel the need for additional exposure.

Additional Observations

The data from the Affective Questionnaire yield some additional insights which are important to consider when interpreting the results of the study. It was revealed that the lesson is perceived as being like a school class. That is, a majority of subjects in the Lesson Group considered the lesson to be about the same as, or just a little better than, school classes with respect to enjoyment (56%), interest (64%), and learning (70%). (See Tables A-13, A-14, and A-15, respectively.) Thus, there is support for the contention that study findings concerning the museum lesson can be related to school lessons.

Finally, it is interesting to note that the kinds of affective outcomes indicated by the experimental data were also mentioned in the teacher survey. Teachers frequently used the words "enjoyment", "interest", and "enthusiasm" when describing outcomes of their class visits. One teacher described the value of the museum experience to her students as follows: "They were motivated to learn and discover in a fun way."

DISCUSSION

Results of the investigations of visiting school groups conducted at The Franklin Institute clearly indicate that participatory museum displays can and do teach science. Regardless of the type of cognitive measure employed (i.e., verbal, visual, or performance), the data unequivocally support the hypothesis that fifth and sixth graders visiting a museum exhibit score significantly higher on a test of science content than students in a control group. In addition, the cognitive data provide valuable information regarding the nature of the learning that takes place in a participatory exhibit. That is, the simple machine displays chosen for the study were found to be considerably more effective in conveying science principles than in teaching vocabulary or suggesting practical applications.

The more pronounced and important findings of the study, however, are in the affective domain. As hypothesized, the affective data indicate that the museum exhibit is perceived as far more enjoyable and interesting than a classroom lesson. This is true whether the basis of comparison is a museum lesson or the students' own school classes.

In addition, responses to the affective questionnaire provide evidence for the hypothesized motivational effect of school visits to a science museum. More specifically, a large proportion of students who visited the exhibit indicated a desire to learn more about simple machines. The longevity of the motivational power of a science museum visit and the extent to which it carries over to other settings remains to be investigated.

However, the teacher survey does provide indications that this motivation is evidenced in behavior back in the classroom.

Most important is the finding that students visiting the exhibit clearly perceive the visit as a learning experience. In fact, a large majority of students reported that they learned more from the exhibit than they learn in school classes. Thus, the study provides assurance to museum staff and classroom teachers alike that middle-school students do <u>not</u> view their school trips to the science museum as merely a recreational activity.

It was not a primary intention of this research to contrast discovery learning and didactic instruction. Rather, the intent in looking at both kinds of experience was to study the motivational impact of museum-based discovery learning on subsequent classroom learning. Nonetheless, the inclusion in the research design of a treatment group that had only a classroom science lesson permits a comparison of the relative educational effectiveness of a participatory museum exhibit and a classroom lesson.

That comparison yielded the following interesting results: When paper-and-pencil tests were used (Experiment 1), students attending the classroom lesson answered more science questions correctly than students visiting the exhibit. However, when the test involved the manipulation of three-dimensional materials (Experiment 2), the two treatment groups performed equally well. Thus, although the exhibit was not found to be <u>more</u> effective than a classroom lesson in conveying science content, it is seen to be equally good. Further, students visiting the exhibit compared most favorably with students who had a traditional lesson when the test most closely resembled the museum learning experience. Support for the hypothesis regarding the importance of a match between learning experience and test mode can be derived from this finding. That is, participatory museum-based learning is best measured with a performance test.

The hypothesis that the motivational impact of a visit to a museum exhibit would be seen in higher test scores of students having both treatments is not conclusively supported by these data. Nevertheless, a pronounced trend in this direction is seen in mean scores and mastery level percentages on the Verbal and Performance Tests. The trend is particularly clear on the Performance Test which requires "recall," a skill which goes beyond simple recognition.

The lack of statistically significant evidence of the impact of a visit to a museum exhibit on subsequent classroom learning may be explained by the design of the study itself. First, the classroom lesson was written as a substantive counterpart of the focal exhibit. Thus, it becomes logically clear that if a portion of the science content was successfully taught by the exhibit, as it was shown to have been, the exhibit's effect on a test of a subsequent lesson conveying the same content would tend to be obscured. Second, for purposes of experimental control, the museum instructor made no mention of the students' earlier experiences in the exhibit area. Most classroom teachers, on the other hand, would build on those experiences by supplying the cognitive link between discoveries made in the exhibit and the science concepts being presented in class. Third, conducting the lesson in a museum classroom probably introduced novelty effects which gave the classroom lesson motivational power in its own right. In fact, the study might have yielded a better measure of the impact of a museum visit if the lesson had been taught back in the school classroom by the students' regular teachers. However, this change in design would have introduced an extraneous variable which is very difficult to control (i.e., the individual style and effectiveness of the teacher).

Some speculation regarding the order of treatments is appropriate here. In light of Ausubel's (1968) concept of advance organizers for discovery learning, it is possible that the educational forte of participatory museum exhibits is in their enhancement and reinforcement of science concepts which have already been outlined for the students. It would be interesting, in a future investigation, to compare the relative effectiveness of participatory exhibit experiences occurring before and after classroom instruction.

The conclusive evidence presented here that children are able to learn science concepts through a nonverbal, kinesthetic experience provides support for theorists who advocate multisensory learning experiences (e.g., Bruner, 1961; Piaget, 1958). Hence, the present investigation has implications for both education and psychology.

The study has further implications for exhibit design. In addition to identifying science principles, rather than vocabulary or applications, as the strongest learning outcome of participatory science museum experiences, the study isolates the

Giant Lever as the most effective of the group of five simple machine displays. These findings confirm earlier evidence that children do not tend to read science museum labels (Borun & Miller, 1980). They also suggest that children find it difficult to transfer science principles from unique mechanical devices to real-world applications. Further study of the potential of an exhibit for conveying an understanding of applications seems warranted. Finally, in view of the particularly strong impact of the Giant Lever, it might be speculated that children learn more from large landmark exhibits. Interestingly, it has been observed elsewhere (Weiss & Boutourline, 1963) that they do, in fact, pay more attention to such exhibits.

In summary, the Franklin Institute study shows that science museums do teach science, but more importantly, they motivate and stimulate an interest in learning science by presenting science content in a manner that is enjoyable, interesting, and perceived to be instructional. In this way, they complement other science education media.

THE BOSTON STUDY

Learning in a museum environment has long been a topic of major concern at the Museum of Science (King, 1979, 1981). Through the course of numerous discussions, the museum staff developed a philosophy of museum-based learning which gave direction to this research study.

The initial planning phases of the project addressed two critical issues. The first was the development of a test format that would reflect the philosophy of learning in the Museum of Science. The second issue was the choice of an exhibit for study.

In considering the relationship of museum learning and test format, museum staff affirmed the belief that learning in or from an exhibit is not limited to the ability to answer discrete questions about the exhibit's subject matter. Rather, learning also takes the form of an increasing interest in speculating about exhibit elements and in perceiving connections among them. It is thought that museum learning is the acquisition of perceptions and images, and that it enhances the ability to make predictions. Furthermore, it is believed that testing should be in the same mode as learning.

It was hypothesized that museum learning would be evidenced in the study in the following ways:

1. increased enthusiasm for subject matter relating to the exhibit used in the study

2. ability to perceive connections between various components of the exhibit, as well as with experiences outside the Museum

3. formation of perceptions and images gained from the museum experience

4. ability to make predictions based on information and experiences acquired in the exhibit

After a review of a number of exhibit possibilities, the exhibit entitled Planets and Moons was tentatively chosen for the study.[1] This exhibit focused on features that can be found on or around the planets and moons in the solar system. The exhibit, housed in a gallery which enclosed 2,000 square feet, consisted of sixty photographs, two interactive devices designed to demonstrate turbulence, a meteorite which could be touched, and a video presentation of the milkdrop photographs of Dr. Harold Edgerton. (See Figure B-1 and Appendix B-1.) Much of the exhibit labeling consisted of questions, the answers to which could be found by examining photographs and then lifting covers which revealed explanations.

Museum staff involved in the selection decision felt that Planets and Moons was a logical exhibit to use for this study for a number of reasons:

1. Planets and Moons was a new exhibit. It was, therefore, unlikely that the students participating in the study would have seen it previously.

---

[1] Planets and Moons was a temporary exhibit and is no longer on display in the Museum.

Figure B-1. Planets and Moons Exhibit

Looking at Kalliroscope

Kalliroscope

Turbulence Machine

2. The subject matter was appropriate for middle school students.

3. The focus of the exhibit was a subtle aspect of astronomy. It was thought that students would not have formally studied this subject matter in school and that the influence of the treatments (exhibit viewing and/or talk) would be more readily discernible.

4. Planets and Moons was a very visually oriented exhibit and, therefore, appropriate to that area of inquiry of the project which relates to visual learning in a science museum.

5. Logisticall Planets and Moons offered considerable advantage. Its location in the Peabody Gallery, an enclosed area, permitted the observation of students while they were viewing the exhibit.

As a final assurance that this exhibit would be appropriate for the study, project staff and consultants observed middle school students in Planets and Moons. After watching more than four hundred students and talking with them about their impressions, it was decided that the exhibit was of interest to sixth- and seventh-grade students.

Once the decision about exhibit choice was made, the design of the talk component began. In presenting planetary and lunar features the exhibit focused on four major themes: rings, craters, volcanoes, and turbulence. The text of the talk portion of the experiment also deals with these themes and is based on information found in the exhibit. (See Appendix B-2.) The only visual aid was a chart of the solar system. ame instructor delivered all of the talks so that there would be consistency in the presentation.

METHOD

Sample Selection

The sample for the Boston study consisted of 535 sixth- and seventh-grade students. Of these, 317 attended a school in a suburban community, 92 attended an inner-city (Boston) school, and 126 attended two schools in a community bordering the core city. The suburban and border schools volunteered to participate in the study. The inner-city school had established a relationship with the Museum of Science and, when asked, was most interested in becoming involved with the study. While the suburban and inner-city students were in the sixth grade at the time of the study, students from the border community were in the seventh grade.

Experimental Procedure

After the school groups had been selected for the study, a schedule of project activities was set up for each school. Since free admission is routinely given to all Massachusetts school groups, free transportation to the Museum was offered to all schools agreeing to participate in the experiment.

When the students arrived, they were randomly assigned to four groups. The sequence of experimental activities for each group was the same as that used in the Philadelphia study. The Boston schedule of events was as follows:

Activity Schedule

1. Welcome, orientation, group formation, general directions (all students, teachers, chaperones, volunteers)
2. Visit to exhibit, lecture, test for control and treatment groups
3. Closing remarks, directions (all students, teachers, chaperones, volunteers)

Group Rotation Schedule

	Activity 1	Activity 2	Activity 3
Control Group	Test	Exhibit	Talk
Exhibit Group	Exhibit	Test	Talk
Talk Group	Talk	Test	Exhibit
Exhibit/Talk Group	Exhibit	Talk	Test

Each of the three project activities was allotted a fifteen-minute time block. The visit to the exhibit and the talk each took ten minutes, with the additional five minutes in the time block set aside for traveling from one part of the museum to another. The testing procedure took fifteen minutes, including time for completion of the Demographic and Affective Questionnaires.

Following a brief orientation, volunteer guides escorted students to the various activity areas. As in the Philadelphia study, all students were exposed to each of the treatments, including treatments which took place after the tests but which were provided to keep the experiences parallel.

Instruments

Demographic Data Sheet. All students were asked to complete a demographic sheet similar to that used at The Franklin Institute. They were requested to indicate their sex, grade, school, age, and date of birth. In addition, students were asked to reply Yes or No to the question, "Have you studied planets and moons in school?" They were assured, by the absence of their names on the sheets, that the museum was not interested in individual achievement, but rather in the total effort of each group.

Affective Questionnaire. Members of the Exhibit, Talk, and Exhibit/Talk Groups were asked to complete Affective Questionnaires so that a measure of their reaction to the museum experience could be obtained. Again, the questionnaires are parallel in content and format to those used by The Franklin Institute.

Cognitive Tests. The forms of assessment developed by the Museum of Science staff are reflective of the Museum's philosophy of learning. It was considered important to measure museum learning through tests that are as open ended and evocative as possible of divergent thinking, thus providing an opportunity for flexibility in the way students expressed themselves. Using tests with student-generated responses in a museum research study is innovative and presents challenges in the development of effective scoring and analysis procedures. However, Museum staff decided that this strategy was worthy of exploration. The tests were designed to measure the influence of the exhibit and the talk. It was a concern that tests be structured so that all students, including the Control Group, could make a serious attempt at the tasks. Furthermore, an effort was made to assure that, as much as possible, students would enjoy the experience of taking the tests.

Although all of the tests in the Boston study were developed with the above criteria in mind, the result was three very different formats, which allowed students to respond to questions and tasks in a variety of modes. All tests were piloted with middle school students. Revisions to the experimental procedure and to the tests themselves were made prior to the actual experiment.

1. Verbal Test. The Verbal Test is an essay test with five questions. Each question is phrased in such a way that no single correct answer is suggested. Students are put in the role of space travelers and asked to speculate about the features on or around planets. (See Appendix B-3.)

2. Visual Test. The Visual Test is a picture-sorting activity. Each student is given a set of thirteen pictures and asked to sort them into two piles. The set of pictures includes whole planets and moons, parts of planets and moons, and specific features of planets and moons. The pictures are in color and are visually appealing. Students determine the categories by which they sort. This procedure is repeated to produce a second set of responses. This type of test allows for a different medium of expression from the Verbal Test and is tied in most closely with the design of the visually oriented exhibit. (See Figure B-2 and Appendix B-4.)

3. Performance Test. The Performance Test requires the manipulation of materials. Students were provided with masonite boards about one foot square covered with a thin layer of clay. They were also given some aluminum foil, an additional ball of clay, and coffee stirrers. Students were asked to use the foil, clay, and cotton to make four features that might be seen on or around planets. (See Figures B-3 and B-4 and Appendix B-5.)

Scoring Procedures

Verbal Test. All questions on the Verbal Test attempt to elicit responses from students that pertain to features on or around planets and moons. The point value assigned to each question was based on the difficulty of the question as assessed by a consultant on middle schools and the number of responses required by each question. The maximum total score was 52. Two scorers evaluated all Verbal Test responses. The score assigned for each item was an average of the two raters' scores.

Visual and Performance Tests. Since the purpose of the Visual and Performance Tests was to measure the influence of the treatments, all answers that pertain to the four main themes -- rings, craters, volcanoes, and turbulence -- were identified and tallied for each of the four experimental groups. This procedure does not imply that other answers are less correct. Rather, determining the frequency of answers designated as main themes is a means of assessing the impact of the experimental treatments. It was expected that the treatment groups would have a greater number of responses in the main-theme categories than the Control Group. In order to check the reliability of the

Figure B-2. Testing -- Visual

Figure B-3. Testing -- Performance

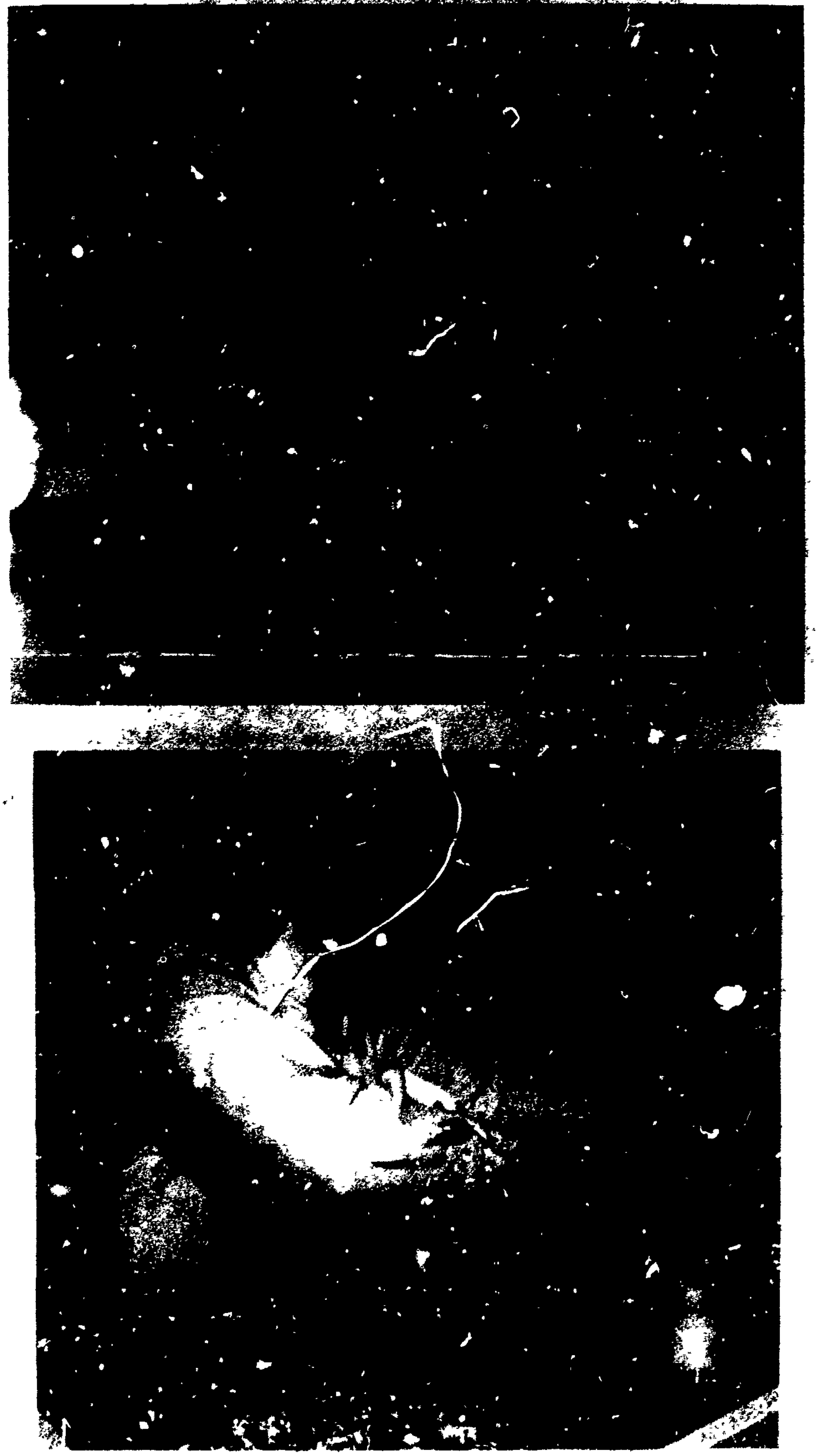
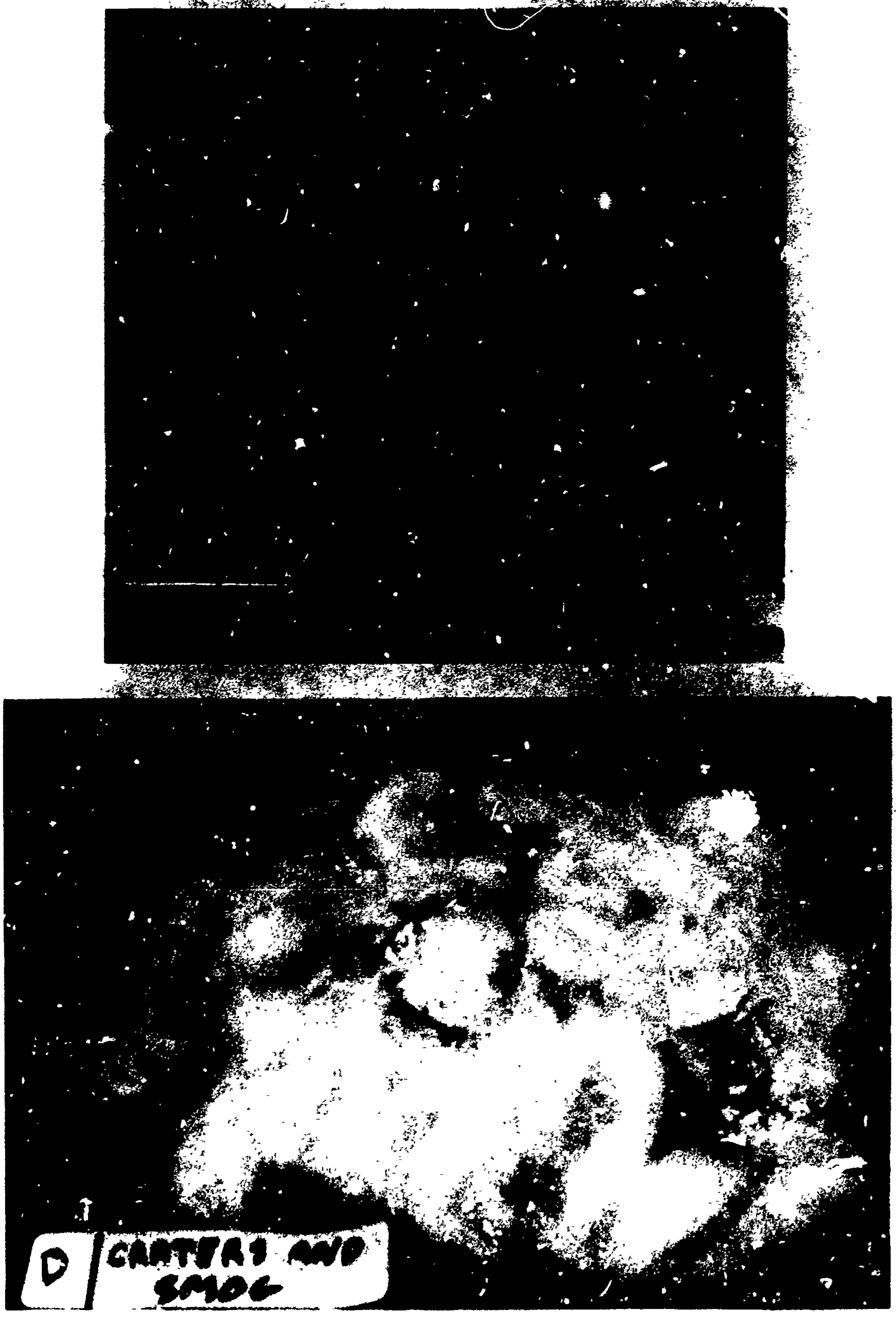


Figure B-4. Performance Test Products



scoring procedure, all tests were scored by two persons. On the Visual Test there were only 15 disagreements out of 852 responses. There were no disagreements between scorers on the Performance Test.

Sample Description

The distribution of subjects in each of the experimental cells for each test is presented in Table B-1.

TABLE B-1

Distribution of Subjects by Test

	Group				
Test	Control	Exhibit	Talk	Exhibit/Talk	Total
Verbal	55	54	56	53	218
Visual	54	51	50	58	213
Performance	28	26	26	24	104

Mean chronological ages for the experimental groups were not significantly different.[1] Means and standard deviations of subjects' ages in years are presented in Table B-2.

[1] F = .31, df = 3/531, ns.

TABLE B-2

Means and Standard Deviations of Age in Years

	Group							
	Control		Exhibit		Talk		Exhibit/Talk	
Test	M	SD	M	SD	M	SD	M	SD
Verbal	11.96	(.70)	11.84	(.79)	11.94	(.88)	11.82	(.66)
Visual	11.73	(.61)	11.80	(.62)	11.79	(.65)	11.71	(.55)
Performance	11.47	(.33)	11.40	(.37)	11.41	(.43)	11.46	(.52)
All Tests	11.77	(.63)	11.73	(.68)	11.78	(.74)	11.71	(.60)

RESULTS

Cognitive Outcomes

The procedures used to assess treatment effects on cognitive test performance differed according to the nature of the test. That is, total scores were the basis of comparison on the 5-item Verbal Test, while frequencies of main-theme responses were compared for both the Visual and Performance Tests.

Verbal Test

The influence of the exhibit and/or talk on Verbal Test responses was evaluated by means of an analysis of variance of total test scores (maximum score = 52 points). Means and standard deviations of scores for students in each of the four experimental groups are presented in Table B-3.

TABLE B-3

Means and Standard Deviations of Scores on the Verbal Test

Group							
Control (N - 55)		Exhibit (N = 56)		Talk (N = 54)		Exhibit/Talk (N = 53)	
M	SD	M	SD	M	SD	M	SD
28.71	(9.00)	24.57	(11.65)	28.02	(12.81)	30.36	(10.11)

Results of the analysis of variance of Verbal Test scores can be summarized as follows:

1. There are no significant differences among the scores of the Control Group and the treatment groups.[1]
2. There is no meaningful trend in the means of the experimental groups.

The above findings suggest that an open-ended verbal test of the type used in this study does not serve as an effective means of determining the level of learning achieved by students who have visited a museum exhibit and/or heard a short talk at the museum. It appears that the evaluation of such brief learning experiences requires a more sensitive test to elicit responses from participants in the treatment groups that are significantly different from responses of the Control Group.

Visual Test

The influence of the exhibit and/or talk on Visual Test performance was assessed by tallying the frequency with which categories used for picture sorting correspond to the four main themes of the exhibit and talk. A summary of frequencies and percentages for the four experimental groups is presented in Table B-4.

A one-way analysis of variance was used to measure group differences in the numbers of main theme categories used by individual students. The results of the analysis indicate that there are no significant differences among the means for the Control (M = . , Exhibit (M = .65), Talk (M = .54), and

---

[1] $F = .98$, $df = 3/214$, ns.

Exhibit/Talk Groups (M = .81). However, some tendency towards a positive treatment effect can be observed in the Exhibit and Exhibit/Talk Groups.[1]

TABLE B-4

Frequencies and Percentages of Main Theme Responses on the Visual Test

Group							
Control (N = 216)		Exhibit (N = 204)		Talk (N = 200)		Exhibit/Talk (N = 232)	
f	%	f	%	f	%	f	%
31	14%	33	16%	27	14%	47	20%

N = number of items (i.e., four potential responses per subject).

The Visual Test may lack adequate structure and, therefore, is not sufficiently sensitive. Students were not asked to sort their pictures according to specific or suggested topics. On the other hand, the mode of the Visual Test does parallel the visual emphasis of the exhibit; thus, further efforts to develop a visual test that is more structured and yet capable of generating responses that reflect perception acquisition, connection making, and student creativity would be worthwhile.

---

[1] F = 2.64, df = 3/209, ns.

Performance Test

The influence of the exhibit and/or talk on Performance Test behavior was assessed by tallying the frequency with which surface features made by the subjects correspond to the four main themes of the exhibit and talk. The frequencies and percentages of all such features for the four experimental groups are presented in Table B-5.

TABLE B-5

Frequencies and Percentages of Main Theme Responses on the Performance Test

Group							
Control (N = 11[illegible])		Exhibit (N = 104)		Talk (N = 104)		Exhibit/Talk (N = 96)	
f	%	f	%	f	%	f	%
32	29%	43	41%	53	51%	42	44%

N = number of items (i.e., four potential responses per subject)

A one-way analysis of variance of numbers of main-theme surface features created by individual subjects revealed a significant treatment effect.[1] Subsequent pairwise comparisons indicated that only the means for the Talk ($M = 2.04$) and Control ($M = 1.14$) Groups are significantly different.[2] Although means

[1] $F = 4.44$, $df = 3/100$, $p < .01$.

[2] Newman-Keuls Test

for the Exhibit (M = 1.65) and Exhibit/Talk (M = 1.75) Groups are not significantly greater than that for the Control Group, they do reflect a definite tendency for subjects visiting the exhibit, with or without the talk, to make surface features representing the four main themes. Thus, of the three tests, the Performance Test proved to be the most sensitive to the experimental treatments.

Affective Outcomes

Responses of Boston students to the Affective Questionnaire provide valuable insights into student impressions of museum visits and museum learning.

Overall Reaction

The selection of cartoon faces and associated wording as indicators of how much students liked the exhibit or talk showed that Boston students reacted positively to both activities. The frequencies and percentages of students in the Exhibit and Talk Groups choosing each response are represented in Table B-6. An independent-samples t test revealed no significant difference between the means (on a 5-point rating scale) of the Exhibit (3.66) and Talk (3.60) Groups.[1]

Students in the Exhibit/Talk Group were asked to evaluate both of their museum experiences. The frequencies and percentages of students in the Exhibit/Talk Group choosing each response are shown in Table B-7. A correlated-samples t test showed that for the Exhibit/Talk Group reaction to the exhibit (M = 3.85) was

[1] $t = .64$, $df = 260$, ns.

TABLE B-6

Overall Reaction of Exhibit and Talk Groups

	Group			
	Exhibit		Talk	
Response	f	%	f	%
I love it	14	11%	16	12%
I like it	72	55%	63	48%
I don't like or dislike it	37	28%	43	33%
I dislike it	2	2%	2	1%
I hate it	6	4%	8	6%

Note: N for Exhibit Group = 131.
N for Talk Group = 132.

TABLE B-7

Overall Reaction of Exhibit/Talk Group

	Museum Experience			
	Exhibit		Talk	
Response	f	%	f	%
I love it	27	20%	11	8%
I like it	76	56%	69	51%
I don't like or dislike it	24	18%	41	30%
I dislike it	2	2%	6	5%
I hate it	6	4%	8	6%

Note: N for Exhibit/Talk Group = 135.

significantly more positive than reaction to the talk (M = 3.40).[1] The difference in the extent to which the two activities were liked is best seen in the percentage who loved the exhibit (20%) versus the percentage who loved the talk (8%).

Thus, while those students who experienced each of the treatments separately had an equivalently positive reaction, the stronger affective response to the exhibit is clearly seen in the responses of students who had both experiences.

Enjoyment

Students in the Exhibit and Talk groups compared their enjoyment of the museum activity to regular school classes. An analysis of item scores (on a 5-point scale) indicates that the exhibit was perceived as significantly more enjoyable than the talk in comparison to school classes.[2] Mean scores for the Exhibit and Talk Groups are 3.92 and 3.52, respectively.

Frequencies and percentages of subjects choosing each response are presented in Table B-8. Reference to the table reveals that 70% of the students in the Exhibit Group, as compared with 54% of the students in the Talk Group, found the activity more enjoyable than their school classes.

Exhibit/Talk Group students compared the exhibit to the talk. It was found that 59% enjoyed the exhibit more, 8% enjoyed the talk more, and 33% considered them equally enjoyable.[3]

---

[1] $t = 5.42$, $df = 134$, $p < .001$.

[2] $t = 3.00$, $df = 261$, $p < .01$.

[3] chi square = 52.30, $df = 3$, $p < .001$.

TABLE B-8

Enjoyment

Comparison of Museum Activity to School Classes	Group: Exhibit f	%	Talk f	%
Much more enjoyable	42	32%	28	21%
A little more enjoyable	50	38%	44	33%
About as enjoyable	18	18%	34	26%
A little less enjoyable	15	11%	21	16%
Much less enjoyable	2	1%	5	4%

Note: N for Exhibit Group = 131.
N for Talk Group = 132.

Interest

A comparison of mean scores (on a 5-point scale) indicates that students found both the exhibit and the talk more interesting than their school classes but did not find the exhibit (M = 3.83) significantly more interesting in comparison with the talk (M = 3.58).[1] Frequencies and percentages of students in the Exhibit and Talk Groups choosing each response with respect to interest are represented in Table B-9.

[1] $t = 1.79$, $df = 261$, ns.

TABLE B-9

Interest

Comparison of Museum Activity to School Classes	Group: Exhibit f	Exhibit %	Talk f	Talk %
Much more interesting	37	28%	34	26%
A little more interesting	55	42%	39	30%
About as interesting	24	18%	38	29%
A little less interesting	10	8%	12	9%

Note: N for Exhibit Group = 131.
N for Talk Group = 132.

The Exhibit/Talk Group data, which compare student responses to the exhibit and talk, show that 58% of the students found the exhibit more interesting, 21% considered the talk more interesting, and 21% felt that the exhibit and talk are of equal interest.[1] As was the case with overall reaction, students who experienced both the exhibit and talk and compared the two more frequently found the exhibit more interesting than did students who had been to the exhibit only and compared it to their school classes.

[1] chi square = 23.60, df = 3, $p < .001$.

Extent of Learning

Students in the Exhibit and Talk Groups saw both museum activities as offering more learning than school classes. Frequencies and percentages of subjects selecting each response are shown in Table B-10. Analysis of item scores (on a 5-point scale) showed that there is no significant difference between the exhibit (M = 3.57) and talk (M = 3.61) in terms of students' perceptions of the learning provided.[1]

TABLE B-10

Extent of Learning

Comparison of Museum Activity to School Classes	Group: Exhibit f	Exhibit %	Talk f	Talk %
Learned much more	22	17%	27	21%
Learned a little more	56	43%	46	35%
Learned about the same	36	27%	40	30%
Learned a little less	10	8%	16	12%
Learned much less	7	5%	3	2%

Note: N for Exhibit Group = 131.
N for Talk Group = 132.

[1] $t = .29$, $df = 261$, ns.

103

Similarly, students in the Exhibit/Talk Group did not see a difference in the instructional effectiveness of the two treatments. Specifically, 32% reported learning more from the exhibit, 36% from the talk, and 32% considered them both the same.

Motivation for Additional Learning

The degree of interest which students expressed in learning more about planets and moons is shown in Table B-11, which presents the frequencies and percentages of students giving each response. For the Exhibit Group, 78% indicated interest in learning more; for the Talk and Exhibit/Talk Groups, 82% wished to learn more. Obviously, all three treatments had a strong motivational effect.

TABLE B-11

Motivation for Additional Learning

	Group					
	Exhibit		Talk		Exhibit/Talk	
Extent of Interest in Additional Learning	f	%	f	%	f	%
Very interested	27	21%	44	33%	41	30%
Somewhat interested	75	57%	64	49%	70	52%
Not very interested	19	14%	19	14%	17	13%
Not at all interested	10	8%	5	4%	7	5%

Note: N for Exhibit Group = 131.
N for Talk Group = 132.
N for Exhibit/Talk Group = 135.

Summary

Two findings of the affective testing merit particular consideration. First, the overall choices of the Exhibit/Talk Group give a strong indication that students prefer viewing an exhibit to attending a talk. Second, it is interesting to note that sustained motivation for learning about the subject of an exhibit, rather than satiation, appears to result from student participation in both an exhibit and a talk. Such affective findings support the use of a science museum exhibit by teachers as an effective introduction to a topic of classroom study.

DISCUSSION

While the findings of the Boston study are not dramatic, they do provide indications of the positive effects of a museum experience, particularly in the affective domain. The first research hypothesis (see p. 58), that enthusiasm for the subject matter would be generated by a school visit to a museum exhibit, was clearly supported by the results of the study conducted at the Museum of Science. Responses to the Affective Questionnaire indicate that a substantial majority of students in treatment groups visiting the focal exhibit liked and enjoyed the exhibit, found it interesting, and wanted to learn more about the topic on display. Given that today's middle school students have access to such a wide range of multisensory experiences, these affective findings should be encouraging to museum educators and other museum professionals. Trips to Disney World, classroom computers, movies, and television programs provide competition to the museum for student endorsement.

Moreover, the fact that 60% of the students in the Exhibit Group felt that they learned more from their experience than from school lessons is worthy of particular consideration. While students might be expected to view the talk as a traditional learning situation, it is notable that they also perceive a non-guided or discovery visit to the exhibit as a learning experience meriting comparison with a school lesson.

The affective results of the Boston study provide assurance to both museum staff and teachers regarding the educational value of a museum exhibit. Enjoying an exhibit and feeling that the exhibit, like a classroom lesson, is a learning experience can be expected to increase student motivation.

The Boston study provides little support, however, for the hypotheses regarding cognitive outcomes of a museum visit. Group differences on all three tests were slight, with a significant treatment effect shown only by the Performance Test. Nevertheless, students found the testing instruments to be both interesting and enjoyable, as evidenced by their comments and observed reactions. They stayed with the tasks, and few failed to attempt all the test items. Thus, the goal of presenting students in the testing situation with a set of activities that were fun to do was achieved.

Perhaps the nature of the testing instruments themselves provides an explanation for the sparse cognitive findings in the Boston study. While the Verbal Test was designed to measure the hypothesized ability to make connections and predictions, the open-ended answer form may not provide sufficient structure for the effects of brief museum experiences to be clearly identified. Further, although the Visual Test requirements are potentially well suited to measuring the perceptions and images provided by a highly visual exhibit, the picture-sorting activity is no doubt too broadly defined. A more structured task might yield significant group differences. Both the Visual and the Performance Test provided students with opportunities to experiment. Activities of this type could be used effectively in science classes for both the introduction of a new topic or as culminating activities for a unit of study.

The more focused Performance Test provides at least some support for the hypothesis that perceptions and images are acquired by student visitors to a museum exhibit and can be manifested through testing. Thus, this test seems to point to the most promising path for future investigations of the nature of museum-based learning.

The study findings suggest, then, that brief open-ended museum experiences involving the acquisition of divergent perceptions and images and the ability to make connections and predictions are perhaps not best measured by a test that generates so wide a range of responses, but rather by a test with a more limited focus. Alternatively, an open-ended test might yield significant treatment effects if students spend a considerably longer time visiting exhibits in the museum. Information gained from this kind of testing could provide insights into how learning beyond the levels of recognition and recall takes place in nonstructured settings.

The need to gather more information about student learning in museum exhibits and to share this information with museum and school staff is clearly urgent. The Boston study provided the opportunity for trying out open-ended strategies for assessing discovery learning in the museum and provided an impetus for further work on the development of sensitive, innovative testing instruments. In this decade of lessening financial support for the mounting of new exhibits in museums and for science programs in the schools, it is especially important that we investigate ways to use existing exhibits more productively, particularly with student visitors. It is critical that a continuing effort be made to design tests that will investigate effectively aspects of museum learning including speculation and prediction, making connections, perception and image acquisition. This study has provided a direction for further work on the measurement of learning in a science museum.

CONCLUSIONS

The experiments conducted at The Franklin Institute Science Museum and the Museum of Science in Boston, on the cognitive and affective outcomes of a school trip to a science museum, indicate that a brief visit to a museum exhibit produces some "learning," whether traditionally defined, as at The Franklin Institute, or more broadly construed, as at the Boston Museum of Science. Both studies compare responses of middle school students randomly assigned to four treatments: control, exhibit only, lesson only, and exhibit followed by lesson. However, the two studies differ on a number of dimensions, the most salient of which are the nature of the focal exhibit and the type of instruments used for cognitive data collection.

While The Franklin Institute focused on participatory displays which provide experiences of a highly kinesthetic nature, the Museum of Science chose an exhibit which was predominantly visual. Further, although the tests at both museums incorporate the same three stimulus modes (i.e., verbal, visual, and performance), they differ significantly in form and content. These differences are a function of the inherent natures of the focal exhibits as well as a contrast in philosophies regarding instrumentation.

At The Franklin Institute, the emphasis was on the learning of science content from a group of displays designed to teach scientific principles; the tests used were highly structured. At the Museum of Science, on the other hand, the main concern was not the amount of science information conveyed by their primarily visual and less didactic exhibit, but the images ("percepts") which were acquired and the mental connections which were made. Since the Boston research team felt that traditional objective tests of science content were not adequate measures for this kind of exhibit influence, more innovative open-ended instruments were used to assess museum learning. Items having one right answer were thought to be inappropriate. Thus, the tests used at the Museum of Science, particularly the Visual Test, allowed for a wide range of correct responses.

The Philadelphia study shows that fifth and sixth graders who visit a museum exhibit score significantly higher on a cognitive test than students in a control group, regardless of test mode -- verbal, visual, or performance. While study findings indicate that discovery learning in a science museum is not as efficient in conveying vocabulary and applications information as a well-structured classroom lecture, they do show that an exhibit can effectively convey science concepts. At the Museum of Science, however, there is only suggestive evidence of exhibit effects on test responses, as seen in the frequency with which subjects taking the Performance Test made surface features that represented the four main themes of the Planets and Moons Exhibit.

It is interesting that in both studies a performance test was found to be the most sensitive measure of exhibit learning. In the case of the Franklin Institute study, this finding can be construed as support for the hypothesized relationship between learning experience and test mode. At the Museum of Science, however, where the Visual Test, rather than the Performance Test, incorporates a stimulus mode most like the focal exhibit, it is an unexpected finding. One explanation may be the relative amount of structure provided by the Performance Test, since it is the most specific of the three tests used in the Boston study.

The fact that the significant treatment effects yielded by the Philadelphia study were not replicated in Boston can probably be explained primarily by differences in the testing instruments used at the two sites. At The Franklin Institute, highly structured objective tests were used to assess the mastery of science content. At the Museum of Science, on the other hand, innovative testing methods were used in an attempt to tap the mental images and categories conveyed by their exhibit. It was thought that such alternative methods of cognitive testing would provide greater breadth to the joint study of museum-based learning.

Indeed, the use of nontraditional, indirect testing methods has often been proposed as the appropriate type of measurement for informal educational settings (Kimche, 1978; King, 1981). But the development of valid and reliable novel tests is no easy task. The Franklin Institute study, on the other hand, shows that the traditional learning that clearly does take place during a short visit to a museum exhibit can successfully be tapped by objective instruments.

Results of the affective testing at the two institutions are more closely in agreement than the cognitive findings. Both investigations provide conclusive evidence that the majority of students visiting a science museum on a class trip enjoy the visit, find the exhibits interesting, and are motivated to learn more about the science content that is demonstrated or displayed. Even more importantly, both sets of data indicate that middle school students perceive their museum visit as a learning experience, not just an enjoyable day away from school routine.

In summary, the investigations at The Franklin Institute and the Boston Museum of Science suggest that the real strength of the museum experience lies not in its efficacy in conveying extensive amounts of information, but in its capacity for generating enthusiasm for and interest in science learning. Discovery learning takes time. Given an equivalent amount of time and attention, more science facts can be learned from a lecture or book than from an interactive exhibit. But how often do people go to a lecture or pick up a science book? The attracting power of a science museum exhibit assures that some learning will, indeed, take place. But, more importantly, exhibits are exciting experiences which stimulate an interest in learning science. In this way, they complement other science education media.

APPENDIX A-1

SIMPLE MACHINES LECTURE

Today I am going to talk about doing work and ways we can make it easi_ by using simple machines.

First of all, I want to make sure we all agree about what work is. When I talk about work, I don't mean homework or the work your parents do to earn money. In science, work is a force (a push or a pull) that moves through a distance. If I push very hard on this wall, although I'm using a lot of energy, no work is being done. A lot of force is being applied, but nothing is moving. However, if I push the table, work <u>is</u> being done because a push is moving through a distance.

In the next few minutes I am going to show you how some simple machines help to make work easier. The work I am going to do is to lift this weight.

The first simple machine is called a lever. A lever usually is made of some kind of board and something called a fulcrum that the board moves on. When I want to use this lever to lift the weight, the best place to push down is as far from the weight as I can get. That's how the lever will help me the most. As I move closer to the fulcrum, it becomes harder to lift the weight. And when I go all the way down to where the fulcrum is, it is almost as if I lifted the weight without any help at all. The help that is given by the lever is called mechanical advantage. The amount of mechanical advantage you get depends on the distance of the force from the fulcrum.

The next simple machine is called an inclined plane or a ramp. If I want to move the weight up to here, I could lift it straight up, but I would have to work harder than if I use the ramp. When I push the weight up the ramp, I am moving it a longer distance, but it is easier to do because the ramp is giving me mechanical advantage. The amount of mechanical advantage given by an inclined plane has to do with its steepness. More force is needed with a steep inclined plane and less force is needed when an inclined plane is not very steep.

The third simple machine is called a pulley. There are two kinds of pulleys. This one is called a fixed pulley. It does not provide mechanical advantage, but it changes the direction of the force. That is useful, too. If I attach the weight to a cord and wrap it around this fixed pulley, I can lift the weight up by pulling down on the other end. However, I am using as much effort as if I just lifted it up by myself.

By adding another pulley, I can make a pulley system that does provide mechanical advantage. This pulley is called a movable pulley because it is attached to the weight itself and moves up and down with the weight. When I use both a fixed pulley and a movable pulley, I am using a lot of rope. By applying my force over that whole distance of rope, it is easier to pull the weight up. The mechanical advantage of a pulley system depends on the number of cords supporting the weight. So it takes less force to lift a weight which is supported by two movable pulleys than a weight which is supported by only one movable pulley.

You have seen three simple machines today -- the lever, the inclined plane, and the pulley. They all gave mechanical advantage because the force moved through a distance. The bigger the distance, the less work you have to do and the more mechanical advantage the machine gives you.

If you look carefully, I think you will see examples of these simple machines in the world around you.

APPENDIX A-2

DEMOGRAPHIC DATA SHEET

Please tell us about yourself by completing the following:

1. I am a _____ girl.

   _____ boy.

2. I am in _____ Grade 5.

   _____ Grade 6.

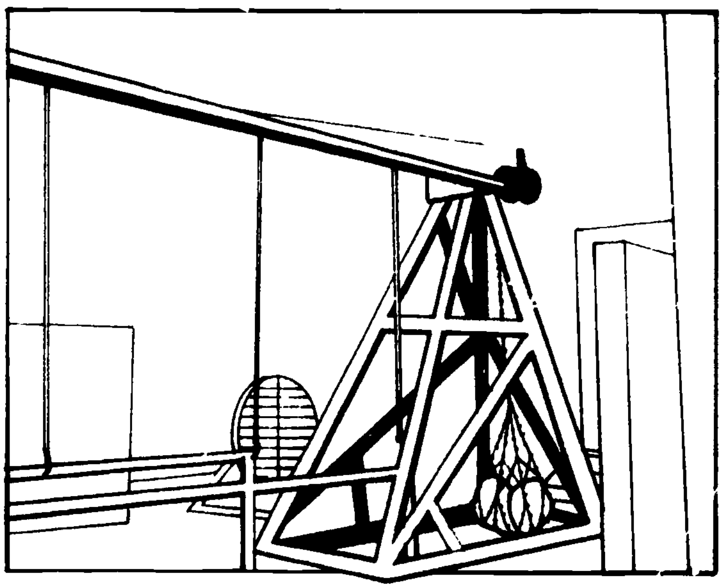
3. Right now I am _____ years old.

4. My birthday is on ______________________________.
   month day

5. Have you studied simple machines in school?

   _____ Yes _____ No

6. Not counting today, how many times have you seen this exhibit in The Franklin Institute?



_____ never

_____ once

_____ twice

_____ three or more times

117

APPENDIX A-3

AFFECTIVE QUESTIONNAIRES

<u>Exhibit/Lesson Group</u>

You have just seen an exhibit and heard a talk about simple machines.

1. a. Which face below shows how you feel about the <u>exhibit</u>? (Write the letter.) _____

   b. Which face below shows how you feel about the <u>talk</u>? (Write the letter.) _____

A	B	C	D	E

I love it	I like it	I don't like it or dislike it	I dislike it	I hate it

2. Now let's compare the exhibit with the talk. Which did you enjoy more? (Check one.)

   the exhibit _____
   the talk _____
   both the same _____

3. Which was more interesting?

   the exhibit _____
   the talk _____
   both the same _____

4. From which do you think you learned more?

   the exhibit _____
   the talk _____
   both the same _____

5. How interested would you be in learning more about simple machines?

   very interested _____
   somewhat interested _____
   not very interested _____
   not at all interested _____

Exhibit Group

You have just seen an exhibit about simple machines.

1. Which face below shows how you feel about the exhibit? (Circle one letter.)

A	B	C	D	E

I love it	I like it	I don't like it or dislike it	I dislike it	I hate it

2. Let's compare the exhibit with your school classes. The exhibit was: (Check one.)

much more enjoyable	____
a little more enjoyable	____
about as enjoyable	____
a little less enjoyable	____
much less enjoyable	____

3. Compared with your school classes, how interesting was the exhibit?

much more interesting	____
a little more interesting	____
about as interesting	____
a little less interesting	____
much less interesting	____

4. Compared with your school classes, how much do you think you have learned from the exhibit?

much more	____
a little more	____
about the same	____
a little less	____
much less	____

5. How interested would you be in learning more about simple machines?

very interested	____
somewhat interested	____
not very interested	____
not at all interested	____

<u>Lesson Group</u>

You have just heard a talk about simple machines.

1. Which face below shows how you feel about the talk? (Circle one letter.)

A	B	C	D	E

I love it	I like it	I don't like it or dislike it	I dislike it	I hate it

2. Let's compare the talk with your school classes. The talk was: (Check one.)

   much more enjoyable _____
   a little more enjoyable _____
   about as enjoyable _____
   a little less enjoyable _____
   much less enjoyable _____

3. Compared with your school classes, how interesting was the talk?

   much more interesting _____
   a little more interesting _____
   about as interesting _____
   a little less interesting _____
   much less interesting _____

4. Compared with your school classes, how much do you think you have learned from the talk?

   much more _____
   a little more _____
   about the same _____
   a little less _____
   much less _____

5. How interested would you be in learning more about simple machines?

   very interested _____
   somewhat interested _____
   not very interested _____
   not at all interested _____

APPENDIX A-4

VERBAL TEST

Directions: Please read each of the following questions slowly and carefully. Then circle the letter in front of the answer you think is correct.

1. Which one is a simple machine?

   A. an electric motor

   B. a pulley

   C. a bicycle

   D. a sewing machine

2. A heavy box must be moved from the floor to a platform. Which way of moving the box would take the least force?

   A. picking it up and placing it on the platform

   B. tying a rope around it and pulling it up over the side of the platform

   C. picking it up and carrying it up steps leading to the platform

   D. tying a rope around it and pulling it up a long board leading to the platform

3. Which one is a lever?

   A. a swing

   B. a seesaw

   C. a jungle gym

   D. a sliding board

4. What is a pulley?

   A. a wheel with a rope around it

   B. a wheel with a rope through the middle

   C. a wheel hanging from a rope

   D. a wheel tied to the end of a rope

5. A heavy rock is stuck in your garden. Which way of getting it loose takes the least force?

   A. bending down and lifting the rock with your hands

   B. putting one end of a pole under the rock and pushing down on the other end

   C. tying a rope around the rock and pulling it

   D. pushing against the middle of the rock with a pole

6. Which one is an inclined plane?

   A. a swing

   B. a seesaw

   C. a jungle gym

   D. a sliding board

7. It is easiest to lift a weight with a lever if you push or pull

   A. as far from the fulcrum as possible.

   B. halfway to the fulcrum.

   C. as close to the fulcrum as possible.

   D. It doesn't make any difference.

8. A heavy trunk must be moved through a window to an upstairs room. Which way of lifting it would take the least force?

   A. with a rope tied around the trunk

   B. with two crates to put on top of each other and climb on

   C. with a pulley attached to the building

   D. with a pulley attached to the building and a pulley attached to the trunk

9. It is easiest to pull a weight up a ramp if the ramp is

   A. very steep.

   B. sort of steep.

   C. hardly steep at all.

   D. The steepness has nothing to do with it.

10. It is easiest to lift a weight with a pulley system that has

   A. one cord supporting the weight.

   B. two cords supporting the weight.

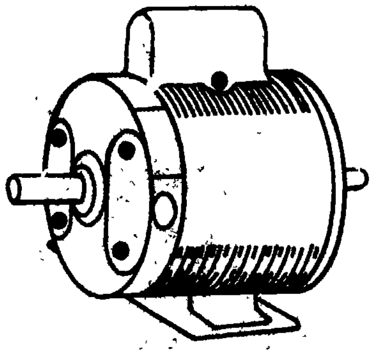
   C. three cords supporting the weight.

   D. The number of cords has nothing to do with it.

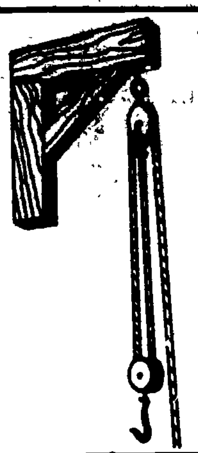
VISUAL TEST

1. Which picture shows a simple machine? (Circle one letter.)

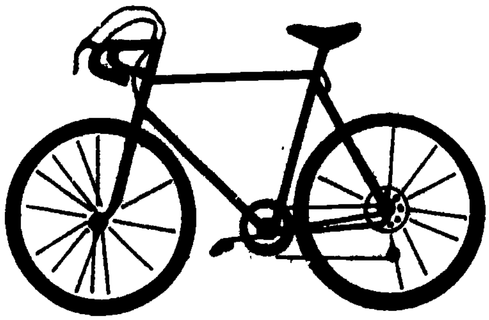
A



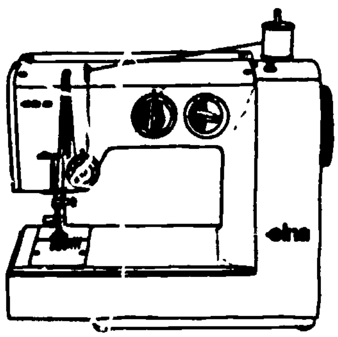
B



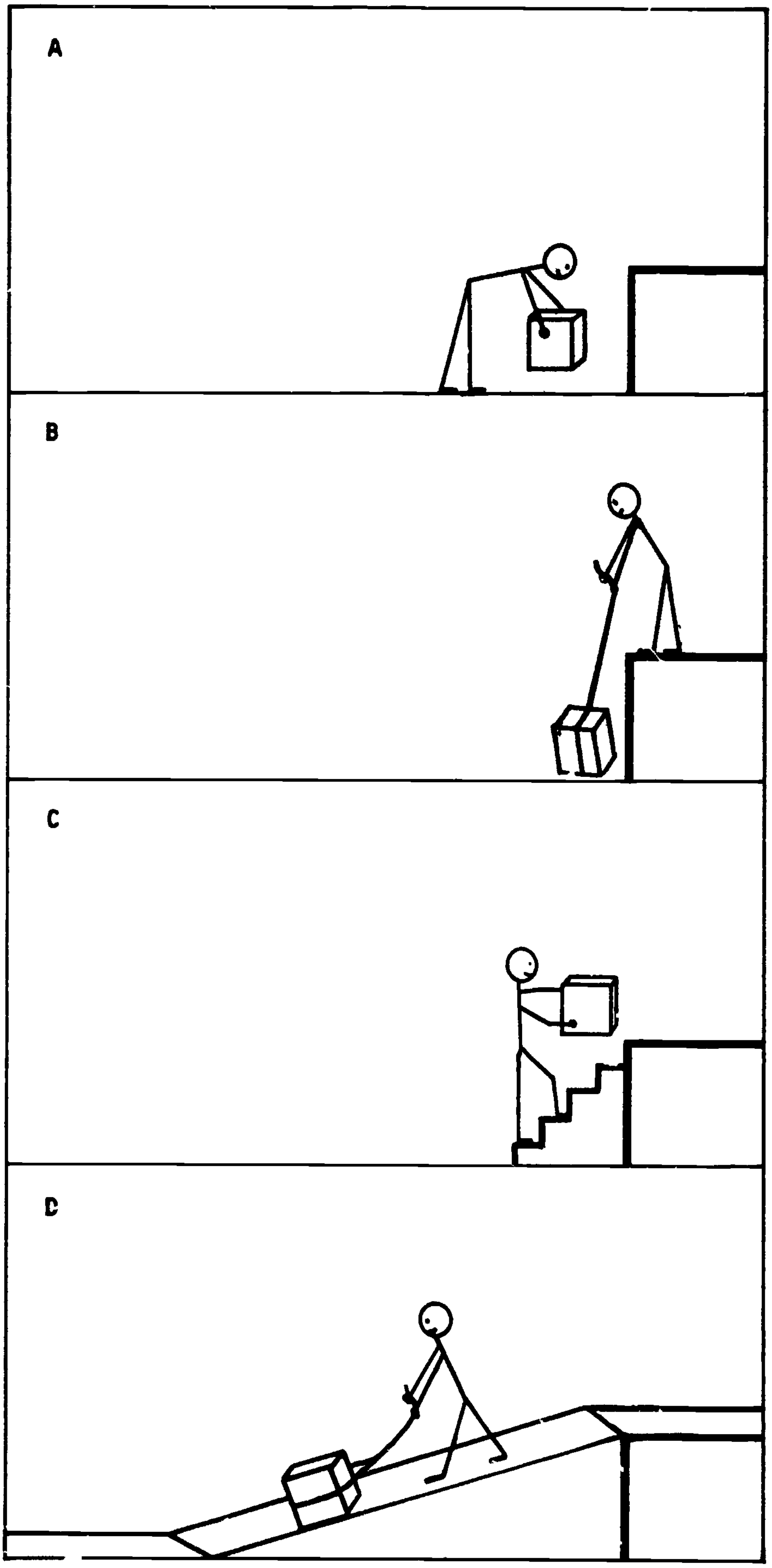
C



D

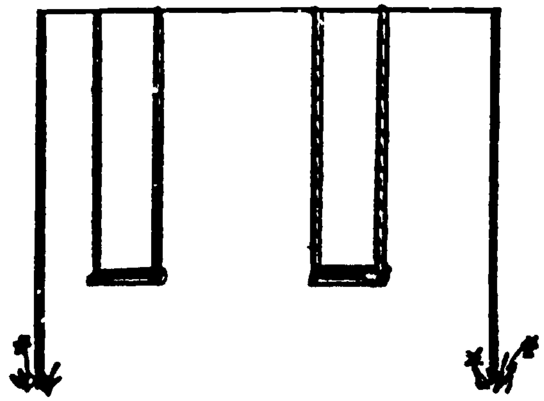


2. Which picture shows the way of raising the box that takes the least force?

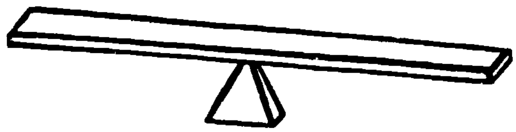


3. Which picture shows a lever?

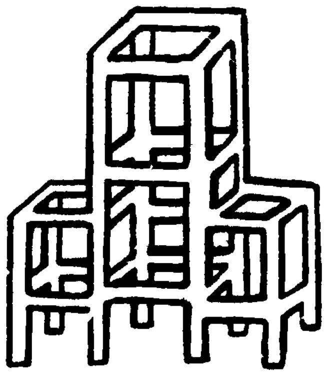
A



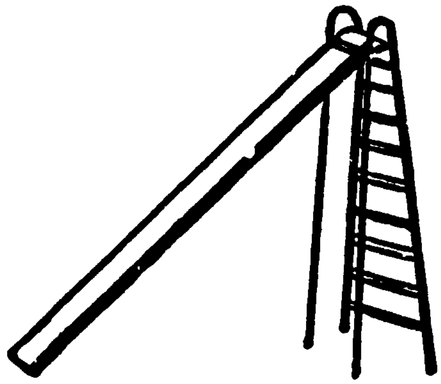
B



C

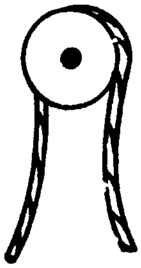


D

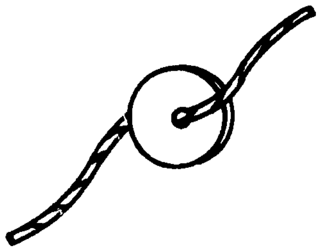


4. Which picture shows a pulley?

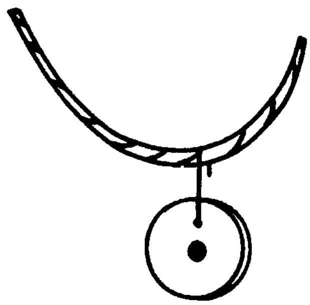
A



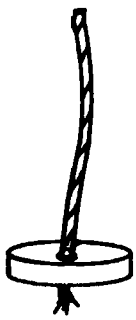
B



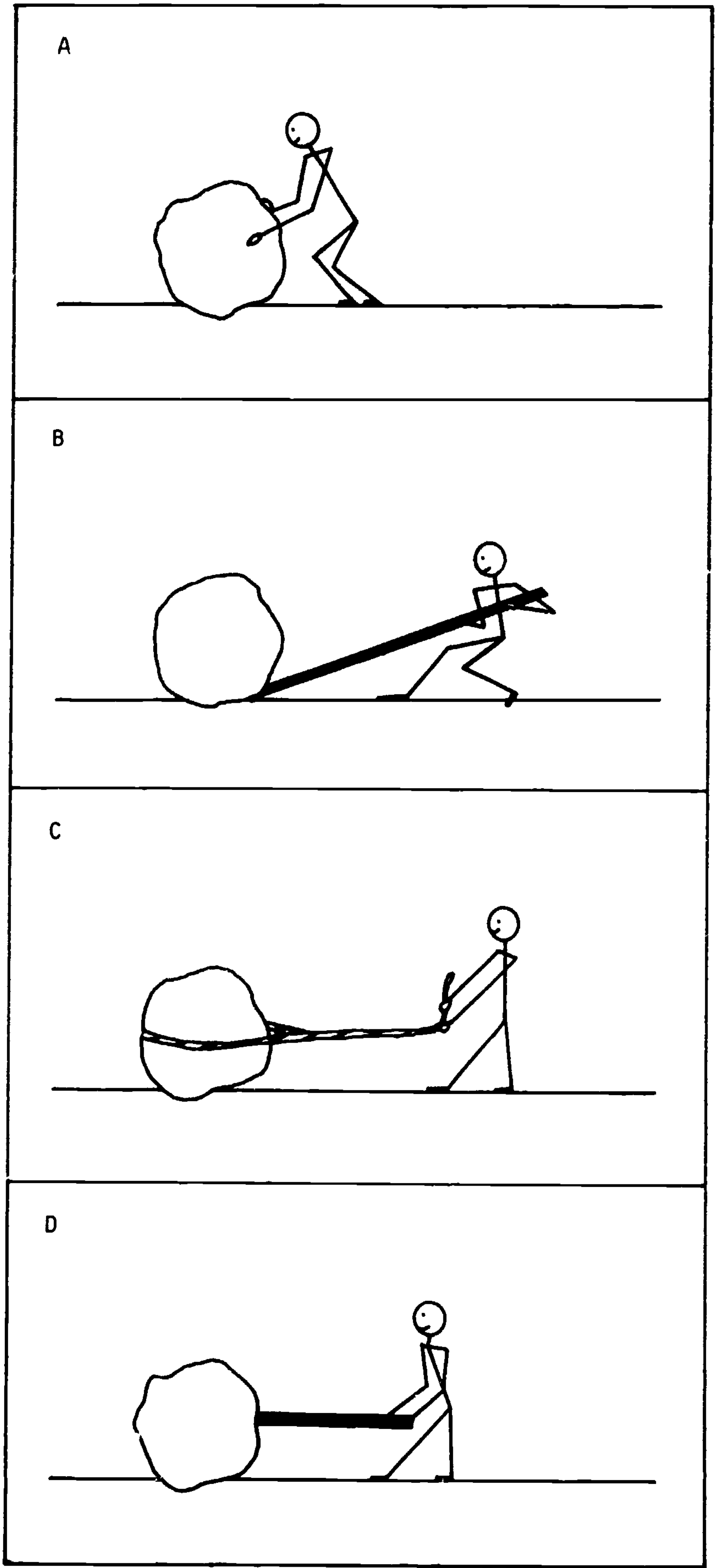
C



D

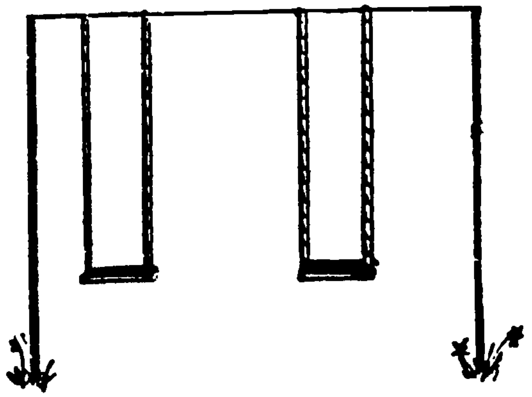


5. Which picture shows the way of getting the rock loose that takes the least force?



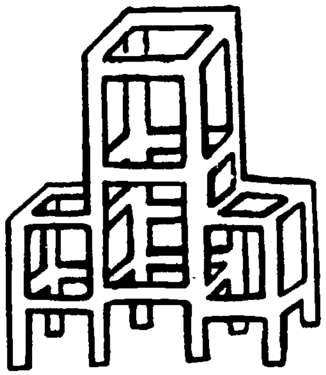
6. Which picture shows an inclined plane?

A

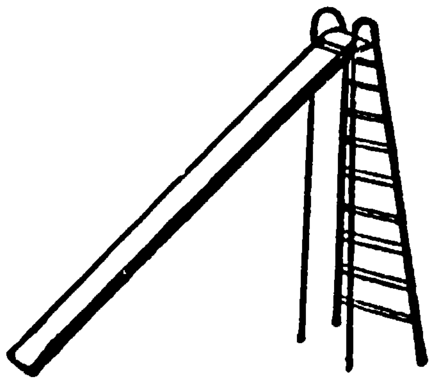


B

C

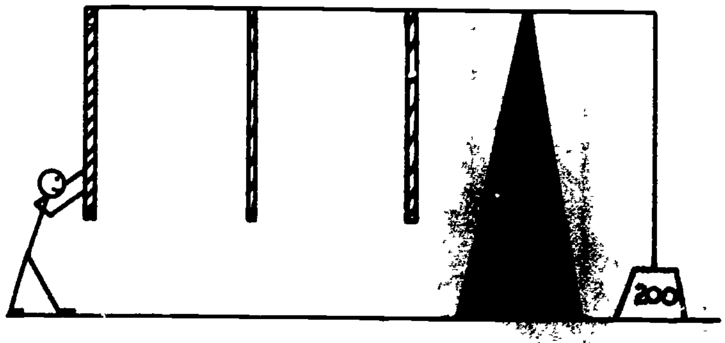


D

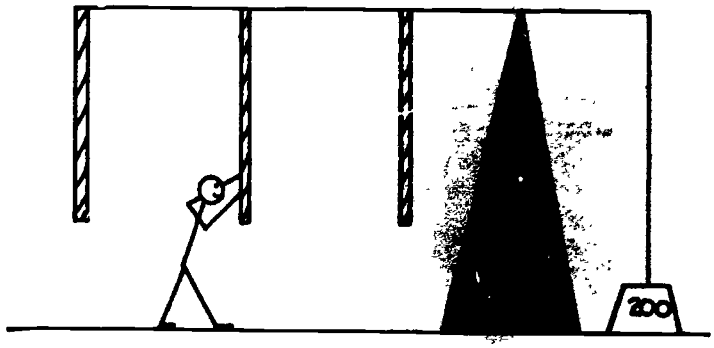


7. Which rope would you pull so tha: it is easiest to lift the weight?

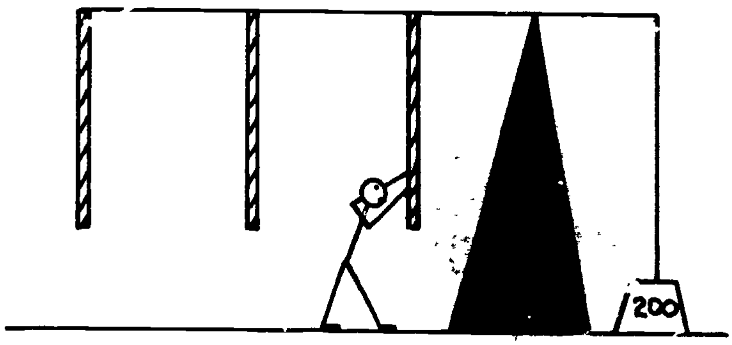
A



B



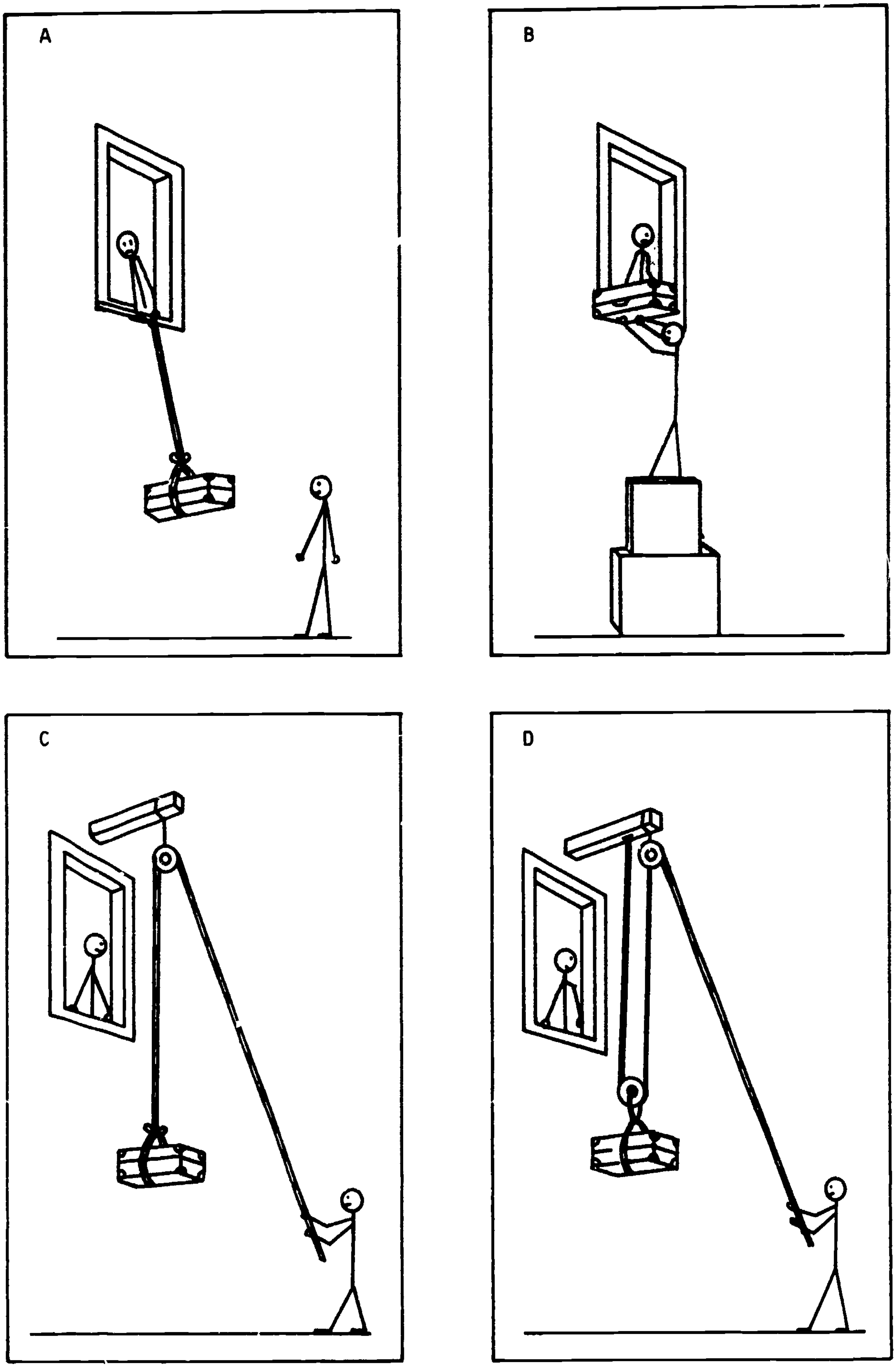
C



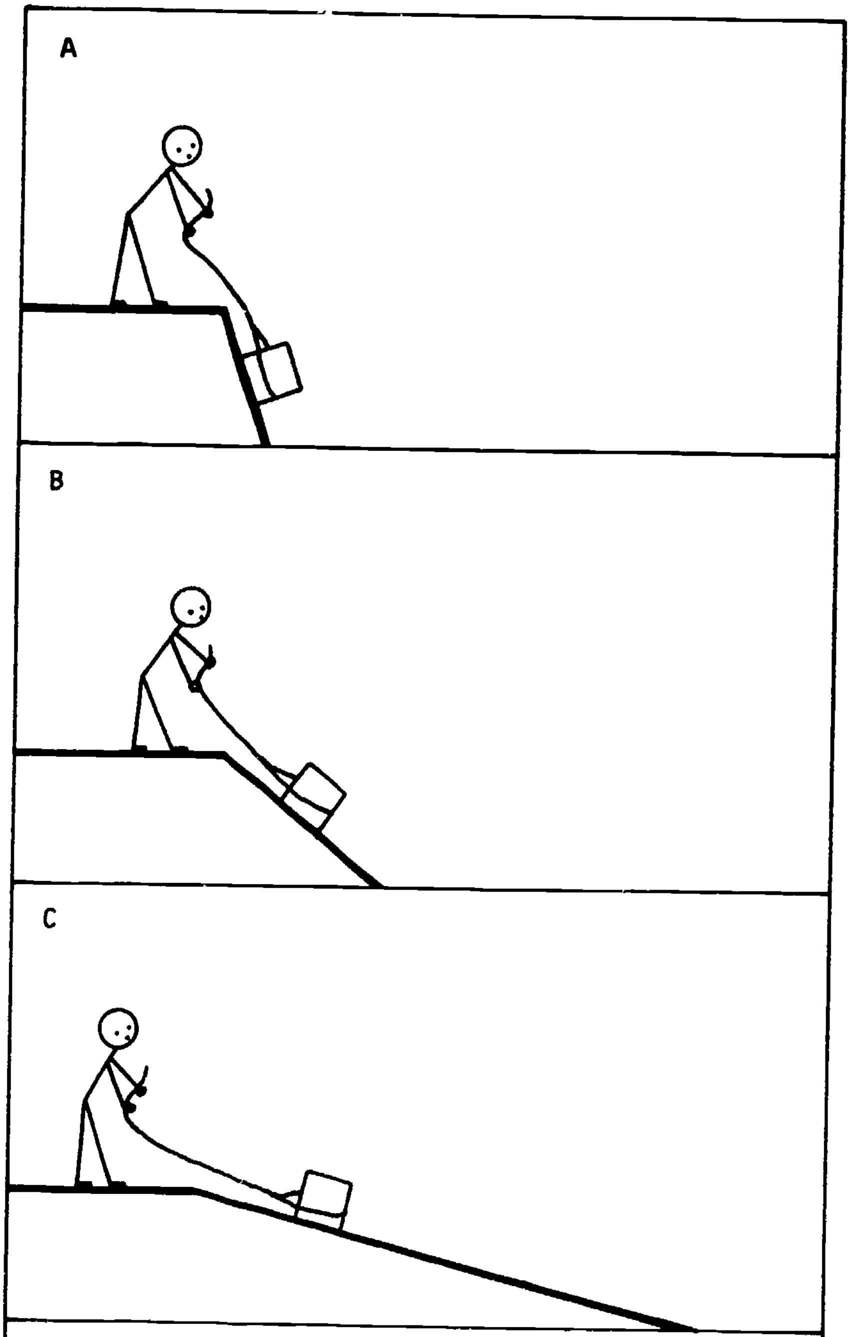
D

It's the same for all three.

8. Which picture shows the way of lifting the trunk that takes the least force?



9. Which ramp makes it easiest to raise the weight?

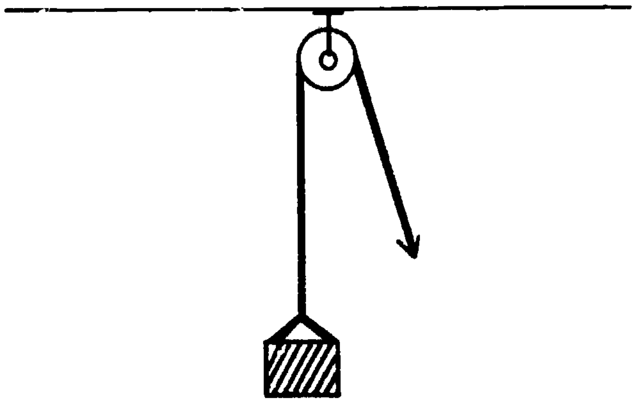


D

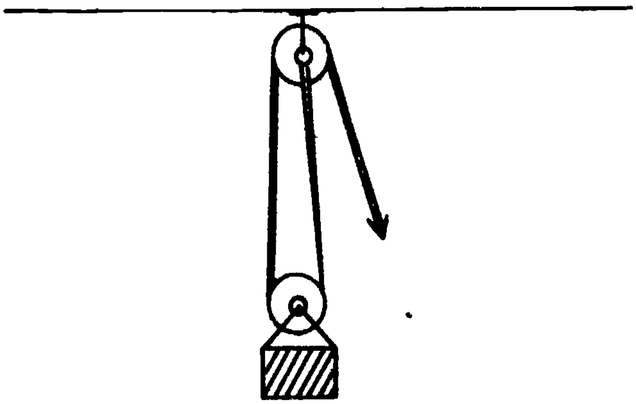
It's the same for all three.

10. Which pulley system makes it easiest to lift the weight?

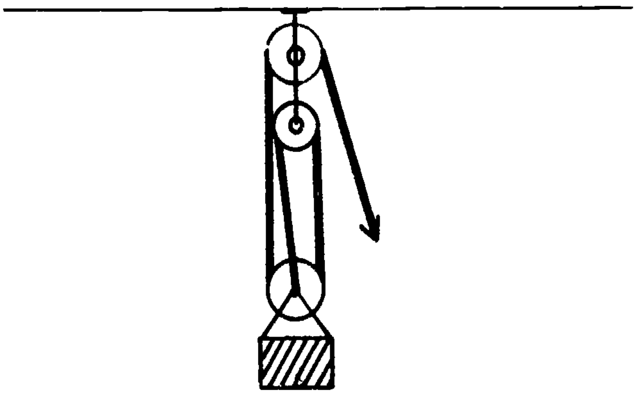
A



B



C



D

It's the same for all three.

APPENDIX A-6

PERFORMANCE TEST

Lever Station

"This Smurf has to move this heavy weight to the top of the platform. He is too little to do it without some help."

1. "Show me how the Smurf could use one of these to make the job easier."

   Correct use ____

   Other response (Specify) ____

2. "Which one would make it easiest to lift the weight to the platform?" (Check which effort arm was selected.)

   Longest ____

   Medium ____

   Shortest ____

3. "Why did you choose that one?" ____

4. "What is this called?" ____

133

Inclined Plane Station

"This Smurf has to move this weight to the top of the platform. He is too little to do it without some help."

1. "Show me how the Smurf could use one of these to make the job easier."

   Correct use _____

   Other response (Specify) ______________________________

2. "Which one would make it easiest to lift the weight to the platform?" (Check which ramp was selected.)

   Longest _____

   Medium _____

   Shortest _____

3. "Why did you choose that one?" ______________________________

4. "What is this called when it is used like this?"

   ____

Pulley Station

"This Smurf has to move this heavy weight to the top of the platform. He is too little to do it without some help."

1. "Show me how the Smurf could use one of these to make the job easier."

   Correct use _____

   Other response (Specify) ______________________________

2. "Which one would make it easiest to lift the weight to the platform?" (Check which pulley system was selected.)

   Three _____

   Two _____

   One _____

3. "Why did you choose that one?" ______________________________

4. "What is this called?" ______________________________

135

APPENDIX A-7

VISITING TEACHER QUESTIONNAIRE

NAME OF SCHOOL____________________________ GRADE_____ TEACHER________________________

If additional space is required for your answers, please use the back of the page.

1. Briefly list up to five reasons you had for deciding to bring your class to The Franklin Institute Science Museum.

2. Number the reasons cited above in order of importance. (1 is most important)

3. Which of these purposes were fulfilled by your recent visit? (list by number)

4. If any were not fulfilled, please list them and explain why you think this is so.

5. Do you think the visit had any educational outcomes for your students?

   a. If so, list them.

   b. What specific student behaviors provided evidence of these outcomes?

6. Describe any other behaviors on the part of your students which you feel were inspired by the visit to the Museum.

7. Did you conduct preparatory activities in the classroom prior to the museum visit?

If so, describe them briefly.

8. Did you conduct follow-up activities in the classroom after the museum visit?

If so, describe them briefly.

9. Please share with us any additional thoughts you might have about the benefits of a school trip to a science museum.

APPENDIX B-1

DESCRIPTION OF KALLIROSCOPE

Two interactive exhibits, called kalliroscopes, were built by Cambridge artist Paul Matisse. They were placed back-to-back in a large free-standing unit. Each device was filled with a mixture of cleaning fluid and aluminum flakes and could be operated by turning a wheel or pushing hot and c ld switches. The movement of the liquid resulted in patterns resembling eddies and turbulence.

The label above the turbulence wheel said: "Try spinning [it] at different speeds. The patterns of flow depend partly on how fast the different parts of the liquid are moving past one another. You should be able to create smooth, flowing, large eddies and turbulence."

The label at the side of the "hot and cold" machine explained that "warm fluids rise, while cool ones sink. The resulting motion causes typical patterns that are very similar to the fluid in hot gas from a volcano (photograph of Mt. St. Helens) or the liquid device at your left." The label asked visitors to try to match the patterns they created with the device to the pattern in the photograph of Mt. St. Helens.

APPENDIX B-2

PLANETS AND MOONS TALK

Imagine you are a space traveler from another galaxy on a scientific expedition. Your mission is to collect information about the SURFACE FEATURES and conditions on the planets of this solar system. What would you need to observe about the planets, with their differences in size, make-up, and surface temperatures, to get this information? Working out from the sun, let's take a closer look at several of these planets and moons.

MERCURY, the planet closest to the sun, is constantly in ferocious heat. It has no atmosphere, and it is likely that at one time Mercury was twice as large as it is now. Scientists think that the sun caused half of the mass of this planet to evaporate. The lighter elements escaped and left a planet that is made mostly of metals.

Examining Mercury from space is hard because of the glare from the sun. The surface of this rocky planet looks like it has been bombarded by comets and asteroids. There are craters everywhere.

A year on Mercury is only 88 Earth days long. A solar day which is 24 hours on Earth is 176 days long. This means that for Mercury to go around the sun it takes 88 days and then for it to turn around on its axis it takes 176 days.

Our next stop is VENUS. This is a planet of great mystery because of the dense fog that surrounds it. The layers of clouds that cover this planet are 35 miles thick. The clouds that cover Earth are only 10 miles thick. This dense atmosphere scatters light just like a heavy fog does. The dense atmosphere of Venus traps the sun's heat. It helps to build up furnace-like conditions. It is hotter on Venus than on any other planet in this solar system.

If you could stand on the surface of Venus, you would find it is very flat. Scientists think that this planet doesn't have any mountains or valleys. A day on Venus is 243 Earth days long. Venus turns in the opposite direction from most of the other planets, so you would see the sun rise in the west every 117 days.

The next planet from the sun is the EARTH. It seems to be blue and white when you look down on it. We know more about this planet than other planets and so we think there are more forms of life here. We don't really know for sure. From space we see white swirling patterns of clouds. As we get closer we find a much wider variety of surface features than has been found on the two previous planets.

Most of the planet Earth is covered by large bodies of water. There are also chains of mountains, deep canyons, and great, brown flatlands. Many features have been shaped by wind and water. We don't see many craters; however, the Earth's crust is broken by earthquakes and volcanoes. This is a clue that Earth is still very much an active planet.

MARS is the next planet out from the sun. It has been of great interest to the inhabitants of Earth who are in constant search for other life forms in this solar system. Like Earth, there is a diversity to this planet. As an observer, you will find three distinct types of landscapes. There is a plains region, like on Earth. There is an area full of craters, and there is a third terrain that has been described as "chaotic." Scientists think that the plains area was caused by an eroding force such as wind which blanketed the area with dust. The south polar region is where the craters are found. The largest crater found on Mars is 300 miles across, or about the distance between Boston and Philadelphia on the planet Earth. The chaotic region covers hundreds of thousands of square miles. It has been thought that frozen dirt below the surface withdrew, like a glacier, and the surface collapsed. Unlike Earth, no water has been found on Mars.

The last planet we will visit on this expedition is the great planet JUPITER. We think that Jupiter is a frozen sample of the original cloud of dust and gas that made up this solar system. The gases which make up Jupiter are trapped there due to very cold surface temperatures and a strong gravitational pull (2-1/3 times that of Earth). Jupiter is 11 times bigger than Earth. It would take 1,300 planet Earths to fill up a hollow ball the size of Jupiter. As you approach this immense planet, you find bands of pastel-colored clouds which spin at whirlwind speeds. Although Jupiter is so large, it turns very quickly. A day on Jupiter is only 10 Earth hours.

There is a great red spot in Jupiter's frozen atmosphere. This egg-shaped area seems to float among the other clouds. It is larger than the surface of the whole Earth. The 14 moons that surround Jupiter add to its mystery and beauty.

I hope you can take this information back to your galaxy. I hope also that you can make another trip to study Saturn, Uranus, Neptune and Pluto.

APPENDIX B-3

VERBAL TEST

Directions for Verbal Activity

INTRODUCTION

We are trying to learn more about things that interest our museum visitors and to find out if our exhibits are doing what we hope they can do.

We should like you to help us by participating in an information sharing activity. We call it a Write-a-thon. You know what a marathon is; you know what a telethon is. This is a write-a-thon. Let me tell you how it works.

DIRECTIONS TO STUDENTS

1. Each of you has received a set of six questions.

2. Would you spend two minutes on each question. Write everything that comes into your mind about that question. I will tell you when the time is up for each question.

3. You may start with any question you wish. Just be sure to move to another question when I announce the time.

4. Are there any questions about what you are going to do?

   Begin. Record the time when students start to work. At the end of two minutes direct students to go on to the next question of their choice. Record the time for Question 2. Continue to announce time and direct students to go on to questions until all six questions have been completed.

COMPLETION OF ACTIVITY

1. Collect completed activity sheets from all students.

2. Collect pencils.

3. Convene group(s) for next activity.

Verbal Test - Answer Sheet

1. If you looked down from space at other planets in our solar system, what kinds of surface features would you see that remind you of Earth?

____

____

____

____

2. Pick any two planets or moons in our solar system. Name two ways they are like each other and two ways they are different from each other.

Your First Planet or Moon is:	Alike
____________________	1. ____________
	2. ____________
Your Second Planet or Moon is:	Different
____________________	1. ____________
	2. ____________

3. You are standing on the surface of a newly discovered planet. You see what looks like smoke coming out of an opening in the surface. What does this tell you?

____

____

____

____

____

4a. As your space ship travels through the solar system, you see a rotating planet covered with gas. Describe the patterns you see surrounding this planet.

____

____

____

____

4b. Patterns made by gases that cover a planet give clues about the planet. As a scientist studying the planet you just described in the last question, what information might you get from studying these clues?

____

____

____

____

5. From what you have seen or learned about the planets in our solar system, what things are most interesting to you?

____

____

____

____

APPENDIX B-4

VISUAL TEST

Directions for Visual Activity

INTRODUCTION

We are trying to learn more about the way people view exhibits. Part of what we are examining is how people organize information in their minds as they are looking at exhibits. We should like you to help us by participating in a picture-sorting activity.

DIRECTIONS TO STUDENTS

1. Each of you will be given a packet of 13 pictures. Each picture has a number on the front.
2. We should like you to sort these pictures into 2 piles. It's up to you to decide how you will sort them. For example, if you were given a set of pictures of animals, you might sort by those animals with 2 legs and those with 4 legs, or you might do it by mammals versus reptiles.
3. After you have sorted the 2 piles, write the topic you sorted by for each pile on the line on your work sheet, and then write the numbers of the pictures that fit into each group in the spaces below. You do not have to have the same number of pictures in each pile.
4. We should like you to do this activity <u>twice</u>, each time with different topics.

Please count the pictures to be sure you have 13. You may want to spread them out in front of you as you sort.

If you have any questions, feel free to ask them.

**If a student asks: "What if I make two piles and have some left over?"

Answer: For your first sort, the rule is to use all pictures. In the second sort, you may have some left over.

(Students will work for 3 minutes on each sort.)

Visual Test - Answer Sheet

PICTURE SORTING

First Sort

File 1= ______ (topic)		File 2= ______ (topic)	
#	#	#	#
#	#	#	#
#	#	#	#
#	#	#	#
#	#	#	#
#	#	#	#

Second Sort

File 1= ______ (topic)		File 2= ______ (topic)	
#	#	#	#
#	#	#	#
#	#	#	#
#	#	#	#
#	#	#	#
#	#	#	#

APPENDIX B-5

PERFORMANCE TEST.

Directions for the Performance Test

INTRODUCTION

The purpose of the activity we are about to do is to create four different features you might see on or around a planet or moon if you were a space traveler. These features are not living things like plants. This activity will give you a choice of materials to use. You will find the materials in the brown container which the Visitor Guides will bring to you.

DIRECTIONS

a. Take the materials out of the box to see what is there.

b. You will find:
   1 board with clay on it
   1 cup with clay in it
   1 cup with cotton in it
   1 piece of aluminum foil
   1 Magic marker
   2 coffee stirrers
   a set of labels

c. You may use any of these materials to make your features.

d. After you finish each feature, take one of the sticky labels, write on it what the feature is that you have made and stick it right beside the feature. We gave you a sample label to show you where the word goes.

e. You will have about 3 minutes to make each feature. We'll let you know when it might be a good idea to go on to making another one.

REFERENCES

Ausubel, D. P. Educational Psychology: A Cognitive View. New York: Holt, Rinehart and Winston, 1968.

Borun, M. Measuring the Immeasurable: A Pilot Study of Museum Effectiveness. Washington, D.C.: Association of Science-Technology Centers, 1977.

Borun, M., & Miller, M. What's in a Name? A Study of the Effectiveness of Explanatory Labels in a Science Museum. Washington, D. C.: Association of Science-Technology Centers, 1980.

Botel, M. Botel Predicting Readability Levels. Chicago: Follett, 1962.

Bruner, J. "The Act of Discovery." Harvard Educational Review, 1961, 31, 21-32.

Eason, L. P., & Linn, M. D. "Evaluation of the Effectiveness of Participatory Exhibits." Curator, 1976, 19(1), 45-62.

Falk, J. Paper presented in session entitled "Preparing School Children for Museum Visits," at annual meeting of the American Association of Museums, Philadelphia, June 1982.

Finklestein, L. B., & Hammill, D. D. "A Reading-Free Science Test." Elementary School Journal, October 1969, 34-37.

Guthrie, J. T. "Expository Instruction Versus a Discovery Method." Journal of Educational Psychology, 1967, 58(1), 45-49.

Inhelder, B., & Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Kimche, L. "Science Centers: A Potential for Learning." Science, 1978, 199, 270-273.

King, R. G. "Educational Capabilities of Science Museums." Newsletter (Museum of Science, Boston), October 1981.

King, R. G. "Learning in Museums." Newsletter (Museum of Science, Boston), April 1979.

Lambert, L. J. "Effects of Museum Resources on the Learning of Anthropological Material in the Elementary Grades." Paper presented at annual meeting of the Central States Anthropological Society, University of Notre Dame, March 1978.

Science Curriculum Improvement Study. SCIS Teacher's Handbook. Berkeley: University of California, 1974.

Sneider, C. I., Eason, L. P., & Friedman, A. J. "Summative Evaluation of a Participatory Science Exhibit." Science Education, 1979, 63(1), 25-36.

Weiss, R. S., & Boutourline, S. "The Communication Value of Exhibits." Museum News, 1963, 43(3), 2?-27.

Wright, E. L. "Analysis of the Effect of a Museum Experience on the Biology Achievement of Sixth Graders." Journal of Research in Science Teaching, 1980, 17(2), 99-104.