DOCUMENT RESUME

ED 045 391 24t SE 010 101

AUTHOR Yost, Michael, Jr.
TITLE The Effect on Learning of Post Instructional Verbal Responses to Questions of Different Degrees of Complexity. Final Report.
INSTITUTION Nova Univ., Fort Lauderdale, Fla.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-9-D-062
PUB DATE 15 Apr 70
GRANT OEG-4-70-0005-057
NOTE 170p.

EDRS PRICE EDRS Price MF-$0.75 HC-$8.60
DESCRIPTORS *Achievement, Energy, Force, *Instruction, *Physics, Programed Instruction, *Questioning Techniques, *Secondary School Science, Verbal Learning

ABSTRACT

Reported is a study of the effect of having seventh grade science students make overt verbal responses in written form to questions of varying degrees of complexity following sequential segments of programed instruction on Newtonian mechanics. It was hypothesized that students responding to more complex questions would have significantly higher achievements than students who responded to less complex questions. Questions were of a given level of complexity for each of three treatment groups and of different levels of complexity for the different groups. A fourth group, without questioning, read a paragraph related to the questions asked other groups. Controls were provided for ability to read and comprehend science content materials. At the .01 level, groups with questioning achieved higher than the treatment group without questioning and the latter groups achieved higher than the control group without instruction; there was a significant positive trend in achievement per unit change in question complexity. It was concluded that as the complexity of questions to which students responded increased, student achievement increased, the amount of instructional time taken increased, and the numbers of errors on the interspersed questions increased. (JM)

BR 9-D-06
PA 24
SE

ED0 45391

FINAL REPORT
Project No. 9-D-062
Grant No. OEG-4-70-0005(057)

The Effect on Learning of Post Instructional Verbal Responses To Questions of Different Degrees of Complexity.

Michael Yost, Jr.

Nova University
3301 College Avenue
Fort Lauderdale
Florida 33314

April 15, 1970

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

ED0 45391

FINAL REPORT
Project No. 9-D-062
Grant No. OEG-4-70-0005(057)

The Effect on Learning of Post Instructional Verbal Responses To Questions of Different Degrees of Complexity

Michael Yost, Jr.

Nova University
3301 College Avenue
Fort Lauderdale
Florida 33314

April 15, 1970

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

The Effect on Learning of Post Instructional Verbal Responses To Questions of Different Degrees of Complexity

SUMMARY

The purpose of this research was to determine the effect on the learning of science subject content of having students make overt (verbal) responses to questions of varying degrees of complexity following sequential segments of instruction. It was hypothesized that students who responded to the more complex questions would have significantly higher achievements than students who responded to the less complex questions.

Five groups of seventh grade students were involved in the study (N=196). Four groups completed a programmed instruction sequence in Newtonian Mechanics and one group acted as the control. Questions were interspersed every twentieth frame (approximately) of the programmed instruction materials for three of the treatment groups and the fourth treatment group read a paragraph related to the questions. The questions were of a given level of complexity for any one treatment group but were of a different level of complexity for different treatment groups. A question was considered more complex than another if it required the respondent to utilize a greater number of parts of the subject content in formulating an answer. A series of judges were used to verify the differences of complexities of the questions. Immediately following the completion of the instructional materials, student achievement on the subject content relevant to the questions (relevant) and not relevant to the questions (incidental) was assessed. The treatment group that completed the instruction materials without the interspersed questions scored significantly higher ($p<.01$) than the control group on both achievement measures. The groups that completed the instruction materials and responded to the interspersed questions scored significantly higher ($P<.01$) than the group that completed the instructional materials and did <u>not</u> respond to questions. There was a significant ($P<.01$) positive trend that described the amount of change in achievement (relevant and incidental) per unit change of question complexity. The change in achievement was accompanied by a similar positive trend in the time needed to complete the instruction materials and by an negative trend in the number of errors made in responding to the interspersed questions. Within each treatment group there was a negative correlation between the number of errors made in responding to questions and achievement, while across treatment groups there was a positive relationship between the number of errors made and achievement. The differences in achievements and the number of errors made were attributed to a general practice effect and inspection behaviors exhibited by students in the learning process.

ACKNOWLEDGEMENTS

In the preparation of this document, I received the assistance of many persons to whom I am grateful. I am especially indebted to those mentioned herein.

Dr. Abraham S. Fischler, as teacher and as sponsor of this study, has been a source of guidance and a sympathetic critic during its progress.

Dr. John M. Flynn has been a continuous source of feedback and evaluation in the development of the proposal and this, the final document.

Dr. William A. Love has provided guidance with regard to the statistical analyses.

Dr. Fletcher G. Watson of Harvard University and Dr. Peter Niiler of the Nova University Oceanographic Research Center have acted as outside evaluators on the dissertation committee and have provided helpful advice and criticism.

The teachers and administrators of Overbrook Regional Junior High School, Lindenwold, New Jersey, provided the necessary environment in which to carry out the study.

Finally, I wish to express my deep gratitude to my wife, Miriam, and our children, for their unfailing support, assistance and encouragement throughout the entire study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

Problem, Theory, and Related Research

Introduction

The interplay of language and human conceptual learning, though complex, is a fundamental factor that pervades most, if not all, of education. In a classroom, the learning of a concept is usually interpreted as a student's mastering of conceptual principles and, by some method, being able to verbalize them for a teacher. In teaching for concept learning, students are given some form of instruction or instructional material that may include lectures, readings, working with apparatus (laboratory-type classes), etc. One set of activities common to nearly all instructional systems is periodic assessment or review of knowledge gained as a part of the instructional process. These complex activities may be used as a form of student evaluation or as a method of assisting students in the learning process. They may, for example, take the form of review questions placed throughout the subject content, short quizzes, student-teacher conversations, or classroom discussions. Writers on teaching methods (e.g., Inlow, 1965: Bruner, 1966) and curriculum specialists (e.g., Oliver, 1965; Wiles, 1963) consider these complex activities (verbalizations) an important, integral part of the learning process.

Investigators have identified many of the relationships that exist between the questions or statements used to elicit verbalization and their learning outcomes. These include: the effect of placing questions to produce verbalization either before or after instructional sequences (Rothkopt, 1966); the amount of intervening subject content between verbalizations (Frase, 1968); the effect of verbalization in the instructional sequence with technical and non-technical subject content (Fry, 1960); the parts of the subject content with which the questions producing verbalization are involved (Frase, 1967); and the types of responses that students use in verbalization (Goldbeck and Campbell, 1962). Through these studies and others, we have acquired information on the relationship between many of the factors involved in questioning to produce verbalization, and the learning outcomes for students.

The investigations reported in the literature on the effect of verbalization have been principally involved with the effect of question placement in the instructional sequence to produce verbalization and the types of subject content involved. Few investigations have been concerned with the characteristics of the questions that elicit the verbalization even though the characteristics of the questions may affect learning outcomes. Several studies reported in the literature have dealt with asking questions about difficult aspects of the subject content and with requiring certain types of responses, but none were found that dealt with the characteristics of the questions themselves. The characteristics of the questions used to elicit verbalization might be an important factor in learning since they may enter into the learning outcomes each time a question is asked. A more precise knowledge of the relationship between the characteristics of questions used to elicit verbalization and the learning outcomes could provide a service to education.

Teachers, authors, lecturers, and curriculum developers, among others, at some time in their professional activities utilize questions or statements to stimulate their students to review or spend time thinking about the subject content with which they are working. This is a relatively common teaching-learning activity. As a result of such activities, students who participate have been shown to learn more of the subject content that they were studying than students who did not participate in the activity (Rothkopf, 1966). There is an obvious need for a greater understanding of the relationship that may exist between the characteristics of questions used to elicit verbalization and the amount of learning that takes place as a function of this verbalization. Since questions to produce verbalization are so frequently used in education as a teaching-learning tool, experimental findings as to the value of questions having given characteristics may be applied across a number of educational environments.

The Problem

The purpose of this research was to determine the effect on the learning of science subject content of having students make overt (verbal) responses to questions of varying degrees of complexity following sequential segments of instruction.

Several groups of students were taught, utilizing the same instructional sequence with questions of a different degree of complexity interspersed in the content of each group. The questions were of a given degree of complexity for any one group, but of different degrees of complexities for different groups. The questions were placed at the same location in the subject content for all groups. It was expected that students who responded to the more complex questions would have different behaviors than the students who responded to the less complex questions. It was anticipated that students who responded to the more complex questions would: read the subject content more carefully, spend more time rereading portions of the subject content, and read more carefully the corrective feedback associated with each of the questions than students who responded to the less complex questions. The differences in the behaviors of students responding to questions of different complexities were expected to be seen in a number of measures. Students who responded to the more complex questions were expected to score higher on an achievement test given immediately following instruction, to make more errors in their responses, and take more time to complete the instructional materials than students responding to the less complex questions. It was anticipated that positive correlations would be found between the number of errors made in the verbalizations and the final achievement scores, and between the final achievement scores and the time needed to complete the instructional materials.

Definitions

Verbalization: To express the answer to a question in words, with pictures, or with numbers. In this study, students verbalized by responding to questions by writing the answers on paper.

Relevant Subject Content: That subject content which is a part of, is named in, or is an acceptable answer to the questions that follow segments of instruction.

Factor: A characteristic of subject content that can be measured or described independently of other characteristics within the subject content.

Level: A subset or subdivision of a factor.

Differences in complexity of questions: One question is more complex than another when it involves either more factors and/or more levels per factor. A question is more complex than another if it involves more factors than another question or questions, regardless of the relevance of the additional factor or factors to the solution or answer. In terms of levels, a question becomes more complex as more levels are introduced, as more members are introduced within a level, and/or as the ratios between the number of members in two or more levels increase.

For example, in subject content dealing with levers, the masses or forces involved, the positioning of the masses relative to the fulcrum, and the angle through which the lever turns are factors. Higher ratios, numbers, or magnitudes of masses and lever arm lengths, and the number of degrees of angles of rotation of the lever are levels within each of the factors.

Hypothesis

Seventh grade students were randomly assigned to work in five treatment groups. The students in four of the treatment groups received the same training (programmed instruction in mechanics) and different treatments (each group responded to questions of a different complexity that were interspersed within the programmed instruction). One treatment group acted as the control group and received neither training nor treatment. The hypotheses dealt with differences in the achievement of students in the treatment groups as a function of the differences in treatment that each received. The hypotheses were stated in the null form.

To describe the facilitative effects of verbalization on achievement, achievement was measured in terms of the subject content with which the verbalization was involved (relevant content) and the subject content with which the verbalization was not involved (incidental content). It appeared necessary, since learning was concerned with both relevant and incidental subject content, to evaluate the effects of the experimental treatment on these two measures collectively as well as separately. Each child's ability to read and comprehend science content materials was statistically controlled for in the hypotheses and analyses.

> H_1: There is no relationship between the degree of complexity of questions producing student verbalization and the amount of increase in student achievement on relevant and incidental subject content when reading ability is statistically controlled.

The preceding hypothesis was tested in the multivariate case with incidental and relevant content considered collectively in a trend analysis. The following two hypotheses were tested as univariate subsets of H_1. The second and third hypotheses (H_2, H_3) somewhat overlap H and are intended to establish the individual relationships of relevant and incidental subject content learning to the treatment in the univariate case. H_2 and H_3 will be tested only if H_1 is rejected.

> H_2: There is no relationship between the degree of complexity of questions producing student verbalization and the amount of increase in student achievement on relevant subject content when reading ability is statistically controlled.

> H_3: There is no relationship between the degree of complexity of questions producing student verbalization and the amount of increase in student achievement on incidental subject content when reading ability is statistically controlled.

Although it was established to some degree in the preceding hypotheses, the effect of both the instructional sequence used (training) and the verbalization (treatment) on student achievement of relevant and incidental subject content should be established. The fourth hypothesis (H_4) evaluated the effect of the training, and the seventh hypothesis (H_7) evaluated the effect of the treatment (verbalization) on relevant and incidental science content achievement.

> H_4: There is no difference in the achievement of two groups of students on relevant and incidental subject content when one group receives no treatment and no training and the other group receives training but no treatment, and reading ability is statistically controlled.

The fifth and sixth hypotheses (H_5 and H_6) were univariate subsets of H_4. Since H_4 was not statistically independent of H_1 and H_7, H_5 and H_6 were handled as discussion hypotheses.

> H_5: There is no difference in the achievement of two groups of students on incidental subject content when one group receives no treatment and no training and the other group receives training but no treatment, and reading ability is statistically controlled.

> H_6: There is no difference in the achievement of two groups of students on relevant subject content when one group receives no treatment and no training, and the other group receives training and no treatment, and reading ability is statistically controlled.

H_7: There is no difference in the achievement of groups of students on relevant and incidental subject content when one group receives training and no treatment, and three other groups receive both training and treatment, and reading ability is statistically controlled.

The eighth and ninth hypotheses (H_8 and H_9) were univariate subsets of H_7. Since H_7 was not statistically independent of H_1 and H_4, H_8 and H_9 were handled as discussion hypotheses.

H_8: There is no difference in the achievement of groups of students on incidental subject content when one group receives training and no treatment, and three other groups receive both training and treatment, and reading ability is statistically controlled.

H_9: There is no difference in the achievement of groups of students on relevant subject content when one group receives training and no treatment, and three other groups receive both training and treatment, and reading ability is statistically controlled.

It was anticipated that as the complexity of the questions used to elicit verbalization increases that the number of errors made in verbal responses would increase, and the use made of the information pertaining to the correct answer (corrective feedback) would also increase. With the greater use of corrective feedback, it was expected that achievement would also increase.

H_{10}: There is no relationship between the number of errors made in verbal response to questions and student achievement on incidental subject content.

H_{11}: There is no relationship between the number of errors made in verbal response to questions and student achievement on relevant subject content.

It was anticipated that responding to questions of varying degrees of complexity within the instructional materials may produce differences in student behaviors. As a result of the differences in behaviors, students may have taken either more or less time to complete the instructional materials.

H_{12}: There is no difference in the amount of time taken by groups of students to complete instructional materials when the instructional materials for each group contain questions of different degrees of complexity.

The Literature Review*

A number of studies were found in the literature dealing with the effect on learning of placing questions within an instructional sequence to produce student verbalization. Questions placed either before or after segments of an instructional sequence have, in general, produced facilitative effects on learning (Rothkopf, 1966). Also, this facilitation has been shown to occur (Rothkopf and Bisbicos, 1967; Rothkopf and Coke, 1966) without any apparent detrimental effects on student-learning activities. Frase (1967) has shown that when verbalization preceded an instructional segment (pre-segment), the increased learning was in terms of specific, question-relevant material and not in terms of enrichment or incidental material. When, on the other hand, verbalization followed an instructional segment (post-segment), both question-relevant and enrichment or incidental subject content learning was facilitated (Ausubel, 1962, Hershberger, 1964).

For both pre-segment and post-segment verbalization, the relative position of the answers (within the subject content) to the questions had no effect. That is, whether the answers were near to or far from (in terms of the amount of intervening content) the questions that produced the verbalization, there were no differences in the amounts of facilitation that were provided (Frase, 1968).

The facilitating effects of verbalization on learning has also

* An annotated bibliography of more complete readings in this subject area follows the bibliography.

been shown to be consistent over a range of ages (Jeffrey, 1963; Norcross, 1957: Wittrock, 1963). For example in the range of ages three through nine, students who verbalized as a part of instruction showed consistently and relatively uniform amounts of improvement in learning at each age level over other students who underwent the same instruction but did not experience verbalization as a part of instruction.

There are essentially two forms of verbal response that students may make to a question. These forms are covert and overt, and both have been shown to have facilitative effects on student learning (Goldbeck and Campbell, 1962; Holland, 1960; Michael and Maccoby, 1953, 1961). Since covert responses are internalized, they are difficult to analyze except in terms of their effect in some later criterion measure (Evans, Glaser, and Homme, 1960; Silverman and Alter, 1960). Overt responses are more easily evaluated than covert response, and may be divided into oral and non-oral types of responses. As the terms imply, an oral response is one that is spoken, and a non-oral response is one that is written, checked, drawn, etc. When criterion measures were administered immediately following an instructional sequence, it was found that covert and overt responses had equally facilitating effects on learning (Goldbeck and Campbell, 1962; Holland, 1960; Kanner and Sultzer, 1956; Silverman and Alter, 1960). Overt responses did, however, produce superior performance on criterion measures when the subject content utilized a technical or specialized vocabulary (Fry, 1960; McGuire, 1955, 1961; Myer, 1960) and when there was a time delay in the administration of the criterion measure

(Goldbeck and Campbell, 1962; Krumboltz, 1964; Krumboltz and Weisman, 1962).

Superimposed on these forms of responses, covert and overt (oral and non-oral), are two modes of responses. The first, or student-constructed-response (SCR) is one in which a student verbalizes by making his choice of words or sentences from his own vocabulary or the lesson content. The second, or investigator-constructed-response (ICR), is one in which a student does not utilize his own vocabulary, but instead chooses from a number of alternate answers prepared by the investigator. With subject content utilizing a non-technical, non-specialized vocabulary, there was no significant difference in the amount of learning facilitated by either SCR and ICR (Coulson and Silberman, 1960; Frase, 1968; Williams, 1965), but in subject content utilizing a technical, specialized vocabulary, significantly more learning occurred as a result of SCR than ICR (Cummings, Allena, and Goldstein, 1962).

There are several points on which there is less than complete agreement about the characteristics of questions used to produce verbalization, the types of responses required of students, and their relationship to the subject content being learned. Skinner (1958), for example, contends that students should have low error rates in responding to statements, and that high error rates reflect a fault in the development of the statements themselves and not in the abilities of students. There appears to be ample evidence to warrant the questioning of this contention in terms of post-segment verbalization. By revealing errors in student understanding through the use of interspersed questions, it is possible to minimize the difference between the intended and the actual output of a lesson through a process of feedback. To be of value, the questions producing effective verbalization have been shown to need certain characteristics. For example, Eigen and Margulies (1963), Holland (1965), and Krumboltz (1964) have found that by asking questions about essential, relevant lesson content rather than trivia, learning was facilitated. Also, Cummings, Allana, and Goldstein (1962), Eigen and Margulies (1963) have found that a question should require a "trenchant answer both in form and content so that errors in responses are not obscured by errors of measurement or simple lack of recognition (p. 52)." These studies begin to point to what may be considered the value of the difficulty or complexity of questions. Goldbeck and Campbell (1962) and Eigen and Margulies (1963) investigated the effect of asking questions of students (verbalization) about the difficult aspects of lessons, where the probability of an error was substantial, and found that this facilitates learning more than when only trivial responses are required. These two studies give additional strength to the thesis that the degree of complexity of questions that elicit verbalization will contribute to the facilitation of learning.

In the studies cited, a form of feedback was used to correct a student's misunderstanding. Once a student was aware of an error, he was provided a source of information to correct his error. Several feedback methods have been utilized, each of which has facilitated learning. These methods have included a student rereading sections of the subject matter germane to his error (Hershberger, 1963, 1964, and Hershberger et al, 1965); being given the correct answer (Myer, 1962; Peterson, 1960); being given the correct answer in a complete sentence (Bivens, 1965; Krumboltz and Bonawitz, 1962); or being given the correct answer with an accompanying explanation (Bryon and Rigney, 1965).

A natural extension of the work cited would be to determine the effect on student learning of asking questions of varying degrees of complexity about relevant aspects of subject content. More specifically, the question to be answered is: What will be the effect of post-instructional verbal responses to questions of different degrees of complexity (1) on the learning of subject content that is relevant to the questions producing verbalization and (2) on the learning of subject content that is not relevant to the questions producing verbalization?

This study was designed to make maximum use of the findings of other studies so that differences that might arise as a function of the treatment (questions of different complexities), although small, may be identified more clearly. The subject content used was technical in nature and post-segment questions with student-constructed responses were used to facilitate learning. The correct answer and an explanation of its derivation (feedback) was provided following each question used to elicit the verbalization. Achievement was assessed immediately following the completion of the learning experience and was done in terms of the subject content that was relevant and irrelevant (incidental) to the questions producing verbalization. It was anticipated that differences in achievement that would arise as a function of the treatment, and their relation to the other variables that were measured as a part of the study, could be identified and explained.

Assumptions

Several assumptions were made regarding the execution and interpretation of the study. These were:

1. Stratified random assignment of students to treatment groups insures that all groups are equal in all ways that are critical to the study.
2. The extra-experimental influences will be the same on all treatment groups, i.e., the four experimental and one control.
3. If differences between the treatment groups as indicated by the criterion measure do exist, they are attributable to the differential effects of the experimental conditions.
4. The experimental conditions vary only on the dimensions indicated.

Chapter II

Procedures

Two pilot studies were carried out prior to the execution of this, the main study. The first pilot study utilized 10 children from the University apartment complex, and the objectives were to evaluate the readability of the instructional and test materials and to determine the time needed by students to complete each of the materials. The second pilot study, which used 30 seventh grade students, was conducted in Thomas Jefferson Junior High School in Dade County, Florida. This pilot study led to further refinement of the wording of portions of the instructional materials and test questions, the deletion of certain test items, and the determining of efficient methods for handling classroom procedures. As a result of the preplanning and of the experience gained through the two pilot studies, the execution of this, the main study, proceeded very smoothly.

The Sample

The sample for the study came from a regional school district in a suburb of Camden, New Jersey. The three communities composing the regional school district have a combined population of approximately 20,000 . In the past 10 years these communities have had a continuous growth of housing tracts, shopping centers, and schools characteristic of a rapidly expanding suburban community. Employment representation within the communities favors blue-collar and white-collar workers, with a modest representation from the professional ranks.

The junior high school (grades 7, 8, and 9) in which the study was conducted has a student population of approximately 900. There were 14 seventh grade class sections with an average of 24 students per section. Eight of the 14 class sections were randomly chosen to participate in the study. Complete data was obtained from 193 of the 196 students in the eight class sections (See Table 1).

These sections were taught science one period a day (42 minutes), five times a week, at five different periods of the seven-period school day. For ease of explanation, the class sections will be referred to by the numbers one through eight.

The execution of the study involved four male science teachers, each of whom taught two of the class sections. The teachers were each given an honorarium for cooperating in the study. All of the teachers are certified by the state of New Jersey to teach science in the seventh grade. Three of the teachers hold master's degrees and are tenured in their positions, and one teacher holds a bachelor's degree and will be eligible for tenure in the 1970-71 school year.

Table 1

Number of Boys and Girls in Class Sections

Class Section	Number of Boys	Number of Girls	Total
1	12	14	26
2	12	10	22
3	12	9	21
4	14	11	25
5	17	10	27
6	10	12	22
7	12	12	24
8	11	15	26
	100	93	193

Prior to the execution of the study, the boys and the girls in each of the class sections were randomly assigned to one of five groups. Each of the five groups was composed of approximately equal numbers of boys and girls from each of the eight class sections. Table 2 contains the number of students from each of the class sections in each of the groups, and the total number of students in each group. Following the establishment of the five groups, each group was randomly assigned to receive one of the experimental treatments.

Table 2

Number of Students of Each Treatment from Each Class Section

Class Section	Treatment Groups 1	2	3	4	5	Total
1	4	5	6	4	7	26
2	4	5	5	3	5	22
3	5	3	5	4	4	21
4	4	4	5	7	5	25
5	6	7	5	4	5	27
6	5	5	3	5	4	22
7	5	4	6	5	4	24
8	6	5	5	5	5	26
						193
Number of Boys	21	18	20	19	20	
Number of Girls	18	20	20	18	19	
Total Number Per Treatment Group	39	38	40	37	39	

In the spring of 1969, the students of the sample completed the Henmon-Nelson Test of Mental Ability (form G). These I.Q. scores were obtained by the investigator from the permanent record folders of the students. The mean I.Q. for the students in the sample was 108.6 and the standard deviation was 10.8. This mean was approximately one half a standard deviation above the national mean for this test. This tends to characterize the students in the sample as being somewhat higher in mental ability than average students across the country. The means and standard deviations of I.Q. for the students in each of the treatment groups appears in Table 3. A single-classification analysis of variance was performed. The results were:

$F = .22$

Degrees of Freedom = 4 and 188

Probability ($P < .9$)

No significant differences were found between the means of the mental ability test measures for the five treatment groups.

Table 3

Means and Standard Deviations of I.Q.'s for Students in Treatment Groups

	Treatment Groups					All Groups Combined
	1	2	3	4	5	
Means	108.3	108.9	107.7	108.2	109.8	108.6
Stand. Dev.	10.9	10.6	10.8	11.0	11.1	10.8

Development of Instructional Materials

The design of the study required that one treatment group act as the control group and the four other treatment groups act as the experimental groups. The instructional materials were the same for each of the four experimental treatment groups. Materials to produce verbalization were interspersed at the same points in the instructional materials for each of the experimental treatment groups. The materials were of a given degree of complexity for any one treatment group, but of different degrees of complexities for different treatment groups. Treatment group one acted as the control, and underwent neither instruction nor questioning. Treatment group two had instruction without questioning and read explanations of the subject content relevant to the questions asked treatment groups three, four, and five. Treatment groups three, four, and five responded to questions dealing with the same subject topics but of different complexities (See Figure 1).

Figure 1

Experimental Design for the Use of Instructional and Treatment Materials

Treatment Groups	1	2	3	4	5
Training (Instructional Materials)	None	Each of these groups received the same instructional materials			
Treatment	None	Read explanations	Least complex questions	More complex questions	Most complex questions

The instructional materials for the study were a modification of Chapter 9 of the Mechanics Unit by TEMAC Programmed Learning Materials, which is entitled Levers: Moment of Force, Clockwise and Counterclockwise Rotation. Conditions of Equilibruim (Encyclopaedia Britannica Press, 1964).

One of the program modifications involved the changing of all measurements from the metric to the British system, and a second modification involved placing the correct answer in the response blank for each frame. By having the correct response in the blank of each frame, students were required to overtly respond (verbalize) only to questions of the experimental treatment. This was a method of attempting to insure that any differences that appeared in achievement at the conclusion of the instructional sequence would be due to the experimental treatment alone. The third modification involved the placement of questions to produce verbalization at every 20th (approximately) frame. The questions were placed at points in the instructional materials that the investigator considered "natural breaks". These were points where the instruction materials changed from one topic area to another. The minimum number of frames between questions was 14 and the maximum number of frames between questions was 28. The programmed instruction materials contained 250 frames with questions placed at 13 points in the materials.

A panel of five experts in the field of education evaluated the questions that were used following the segments of instruction of the experimental treatment groups. The qualifications of the experts appears in Appendix A. Each expert independently assessed whether the questions did, in fact, differ in complexity, meet the criteria set down in the definitions, involve subject content relevant to the preceding segment of instruction (content validity), and maintain a seventh grade reading level. Each expert was given two sets of questions (three questions per set - one question per treatment group) that could be used to follow each of the 13 segments of instruction, the definitions for question complexity, and the subject content that preceded each set of questions. Appendix B contains a set of the instructions that were given to the judges.

Three or more of the five experts agreed that all but one of the sets of questions met the preset standards on the complexity, validity, and reading level (See Appendic A). The one unacceptable set of questions was modified, resubmitted to the experts, and found to meet the criteria.

The experts evaluated 26 sets of questions. The questions in the pairs of sets were essentially parallel forms dealing with the same content. One set of questions, of the two sets evaluated to be used following each segment of instruction, was randomly chosen to be used as a part of the instructional materials. Portions of the remaining sets of questions were used as a part of the final evaluation of relevant subject content. The final draft of the instructional materials, with the questions of each of the complexities interspersed, appears in Appendix C. Note that the instructional materials used by an individual student contains questions of a single complexity, and not of all complexities as appears in the appendix.

The format of the instructional materials was the same for the students in each of the four experimental treatment groups. Each question was printed on a single sheet of paper. The questions for each treatment group were then collated into the pages of the programmed instruction materials that were common to all treatments.

The questions for treatment groups three, four, and five were printed on one side of a page with the correct answers printed on the reverse side. The correct answer to each question was preceded by an explanation of the reasoning used to obtain that answer. The explanation also included the reasoning used to obtain the answers to the questions found at the same point in the instructional materials for the two other treatment groups who answered questions. Treatment group two did not answer questions, but did read the explanations of how answers were derived for the questions of treatment groups three, four, and five. This technique of having all of the experimental treatment groups who received instruction read the same combined explanation was a method of equalizing the amount of information given all groups.

Development of Evaluation Instruments

Two achievement tests and one science content paragraph comprehension test were developed and used as a part of the study. Both of the achievement tests were 25 item multiple choice type tests. One of the tests (Test I) tested achievement on relevant subject content. The items for the test were randomly selected from the sets of items judged by the panel of experts that were not used in the instructional materials. The second achievement test (Test II) tested incidental content achievement. The items for this test dealt with the topics of the programmed instruction material not covered by the questions that had been developed and interspersed in the programmed instruction materials. The content validity of these questions was established by a panel of five seventh and eighth grade science teachers.

The members of the teacher-panel were given a set of the instructional materials, the questions, and were asked, "Can seventh grade students obtain sufficient information from this programmed instruction sequence to answer these questions?" (See Appendix E.) The members of the panel independently evaluated each question. Each of the questions used on the test was seen as acceptable by at least four of the five panel members.

The items for the science content paragraph comprehension test were taken directly from the paragraph comprehension sections of Stanford Achievement Tests. The items were taken from alternate forms of the Intermediate II and the Advanced test batteries. All items used involved the physical, biological, or earth sciences. A total of 50 multiple choice items were used in the test. (See Appendix F for copies of all tests.)

Design

Three weeks (five school days per week) were needed for the execution of the study. Two of the three weeks were used in preparing students for the study (pre-experimental treatment), and one week was used for the experimental treatment and data collection. Students were not informed that they were participating in an experimental study, but instead were told that they were "trying out" some new instructional materials. They were told that the grades they earned would count toward their marking period grade and that they should work as conscientiously as possible.

The investigator was present in the school for the last two of the three weeks to coordinate the activities of teachers and students, and to collect and distribute the experimental materials. Students inquiring as to the investigator's presence were told that he was responsible for the trying out of the new instructional materials.

Prior to the study, the cooperating classroom teachers were given a set of instructions that described the students' and teachers' parts in the study (See Appendix D). The instructions included a timetable of classroom activities and the behaviors expected of the teachers during each of the classroom activities. The teachers were not told the hypotheses or objectives of the study until all of the data had been collected.

The execution of the study was dependent upon the students' ability to work with linear programmed instruction materials. In the first two weeks of the study (pre-experimental treatment), all students in the sample worked in Encyclopaedia Britannica Educational Corporation's programmed text in astronomy. The purpose of this two weeks' work was to give students experience at working with programmed instruction materials. Executing the study without this introduction to programmed instruction would have introduced another variable to the experimental situation that may have had a confounding effect in interpreting the outcomes of the study. On the last class day of the pre-experimental treatment, all students in the sample were given an astronomy test. The test covered the subject content of the first two chapters of the programmed astronomy text. The test grades were not used as a part of the study, but only to insure that all students in the study had the same pre-experimental treatment.

On the tenth class day, a Friday, the students in the sample completed the science content reading comprehension test. They were given the entire science class period (42 minutes) in which to complete the test. Neither the teachers nor the students were given the results of this test.

The experimental treatment began on the eleventh class day - a Monday. For the first three days of the experimental treatment, the students in each class who had been randomly assigned to treatment group one (the control group) were sent to the library to continue working in their programmed astronomy texts. On the fourth and fifth days of the experimental treatment, the students from treatment group one returned to their respective science classes to complete the two achievement tests.

The students who had been randomly assigned to treatment groups two through five in each of the eight science classes worked on the experimental instruction material in their assigned classrooms during their regular science period. That is, some of the students from each of the eight science classes were working in each of the four treatment groups. Students in the science classrooms were not permitted to talk or move about.

The science teachers maintained an accounting of the amount of time each student worked on the instructional materials of the study. Teachers noted the time when the classes of students began working and finished working each period. If a student left class or was otherwise interrupted from his class work, the teacher noted the number of minutes and subtracted this from his total work time for the day. If a student completed his work prior to the end of his last class period, his working time for that day was noted.

It was expected that an average seventh grade student would not spend all of his class time working uninterruptedly on the instructional materials. The investigator, in his classroom visits, saw students who occasionally looked up from their work, looked around the room, and then returned to their work. This was a type of "break" that students took from the instruction. Nothing was said to students taking this type of "break", but students who tended to "daydream" (unnecessarily long breaks) were instructed to return to their work. The stimulus used to have "daydreaming" students return to work took the form of a softly spoken personal conversation rather than a reprimand that may have affected the performance of that student or other students in the class.

Students were not permitted to remove the instructional materials from the classroom. This procedure insured a more accurate accounting of individual student working time, reduced the probability of students interacting with others on the subject content, and reduced the probability of outside influences entering into the experimental outcome.

Following the completion of the instructional materials, each student - including the control group - completed the two achievement tests. Test number one was on relevant subject content and test number two was on incidental content. Students were free to take as much time as they needed to complete the two tests. To insure against the possibility that taking one test might serve as a learning experience and thus influence scores on the second test, the order in which the tests were taken was reversed for half of the classes. That is, class sections one, three, five and seven took test one followed by test two and class sections two, four, six, and eight took the tests in the reverse order. Following the administration of the two tests, the science teachers returned to their regular curriculum materials.

The investigator and a science teacher evaluated the verbal responses (written answers) made by students to the questions that were interspersed in the programmed instruction materials. The science teacher had no knowledge of the objectives of the study, did not know the students involved in the study, and except for evaluating student responses, the investigator and the science teacher read each question with its accompanying feedback and correct answer, and discussed what minor modifications of the stated answers they would accept as a correct answer. Since many of the questions required mathematical or scientific answers, there was little ambiguity as to what was an acceptable, correct answer. The investigator and the teacher independently evaluated each student's responses and the correlation (product moment) between the scores of their evaluations was .97. Following the initial, independent evaluations of the responses, the investigator and ... her collectively reevaluated those responses on which they had ... nally disagreed. Both of the evaluators found that they had mi ... ed some of the students' responses. The score obtained for ... student following the joint reevaluation of questions on which ... valuators did not initially agree was the one used in the data ... is of the study.

The investigator collected all instruct ... terials and tests. Student responses on tests were key punched, ... ed, and machine scored. Figure 2 lists the treatment given ... data collected from the students in each of the treatment gr...

Figure 2

Experimental Design

Treatment Groups	1	2	3	4	5
Training	None	Each of these groups will receive the same instructional materials			
Treatment (questions to produce verbalization)	None	None	Least complex	More complex	Most complex
Error Count (number of errors made in verbal responses)	N.A.*	N.A.*	Error count	Error count	Error count
Tests	I. Test on relevant subj. content. II. Test on incidental subj. content. Both tests were administered to all groups.				
Covariate	All groups were given a reading test that acted as a covariate in the statistical analyses.				
Time	N.A.*	The amount of time taken by students to complete the instruction materials was recorded.			

*N.A. not applicable in this treatment.

Chapter III

Statistical Analyses

Item Analyses and Test Reliabilities

Guertin (1965) has suggested that a total test score, which is made up of contributions from all the individual items of the test, tends to suffer from "anemia" (p. 376). That is, a portion of the total score variance of a test is made up of a number of components other than those originally intended to be a part of the test. "Total score is composed of: variance due to many smaller group-factors scattered among the test items, to unique factors measured only by a single item, and unreliability (Guertin, 1965, p. 376)." Richardson (1936) has shown that the correlation between a test item and the composite score for a test "gives an indication of the extent to which the item measures what the test as a whole measures (p. 71)." Richardson (1936) continues by stating that the rejection of items that have a relatively low correlation with the composited test score raises the average intercorrelation of the remaining items and develops a test of higher reliability. Flynn (1969) developed an algorithm for use with a computer that was based on the work of Guertin, Richardson, and others (Kuder and Richardson, 1937) in the field of item and test reliability. Flynn's program allows an investigator to evaluate the correlations between each of the test items and the composite score, and to delete those items having low correlations with the composite score. The resultant test contains fewer items than the original test, tends to have a higher reliability coefficient, and contains items that tend to measure student ability in the same subject area. Following the deletion of test items, the program treats those items that were deleted as a new test to determine whether the deleted items, in fact, measure a characteristic of the subject content different from the test from which they were deleted. The program grades each student's responses, calculates the number and percentage of students making correct responses to each item, calculates the correlation (point biserial) between each item and the total test score, and calculates the Kuder-Richardson (20) (Kuder and Richardson, 1937) reliability for the test. After completing these calculations, the program then identifies the lowest item-composite correlation, eliminates this item from all subjects' scores, and repeats the above calculations for what may be considered a new test having one less item than the original test. The item elimination process is terminated when preset criterion values for the item-composite correlation or the Kuder-Richardson (20) reliability coefficient is reached.

The item analyses and reliability coefficients for each of the three tests (Reading, Test I, Test II) were calculated using the responses of all students in the sample (N-193). The equation [Kuder-Richardson (20)] used to calculate the reliability was:

$$r_n = \frac{n}{n-1} \cdot \frac{\sigma_t^2 - \Sigma pq}{\sigma_t^2}$$

where n is the number of items in the test, σ_t^2 is the variance of the scores on test t, Σpq is the sum of the item variances, and

Table 4

Items Retained Following Item Analysis of Reading Test

Item	r	Percent Correct	Item	r	Percent Correct
1	0.25	55.	26	0.31	60.
2	0.37	91.	27	0.41	67.
3	0.37	86.	28	0.37	78.
4	*****		29	0.30	48.
5	*****		30	*****	
6	0.30	52.	31	0.50	60.
7	*****		32	0.31	64.
8	0.37	71	33	0.37	66.
9	0.28	48.	34	0.37	31.
10	0.29	40.	35	0.33	48.
11	0.42	68.	36	0.37	64.
12	0.36	87.	37	0.43	66.
13	0.34	47.	38	0.45	73.
14	0.40	67.	39	0.27	78.
15	0.41	41.	40	0.41	76.
16	0.39	84.	41	0.39	54.
17	0.39	62.	42	0.31	45.
18	0.43	66.	43	0.48	85.
19	*****		44	0.37	83.
20	0.37	50.	45	0.42	86.
21	0.39	77.	46	0.37	90.
22	0.25	27	47	0.49	73.
23	0.27	86.	48	0.45	80.
24	0.31	64.	49	0.46	70.
25	0.35	63.	50	0.42	70.

***** Indicates the item has been eliminated as a part of the item analysis.

Table 5

Items Retained Following Item Analysis of Achievement Test I

Item	r	Percent Correct
1	0.47	33.
2	0.34	29.
3	0.33	55.
4	0.43	86.
5	*****	
6	0.37	43.
7	0.33	36.
8	0.37	79.
9	0.33	53.
10	0.45	74.
11	*****	
12	0.36	46.
13	0.56	71.
14	*****	
15	0.36	55.
16	0.36	65.
17	*****	
18	0.40	40.
19	0.40	63.
20	*****	
21	*****	
22	0.46	75.
23	0.37	45.
24	0.42	74.
25	0.34	72.

***** Indicates the item has been eliminated as a part of the item analysis.

Table 6

Items Retained Following Item Analysis of Achievement Test 2

Item	r	Percent Correct
1	0.49	55.
2	0.47	65.
3	*****	
4	0.50	58.
5	0.41	67.
6	0.39	45.
7	0.25	57.
8	*****	
9	0.53	55.
10	0.40	39.
11	0.48	57.
12	0.37	47.
13	*****	
14	0.43	51.
15	*****	
16	0.48	45.
17	0.46	46.
18	*****	
19	0.50	72.
20	0.33	60.
21	*****	
22	0.33	37.
23	*****	
24	0.35	59.
25	0.36	42.

***** Indicates the item has been eliminated as a part of the item analysis.

r_{tt} is the reliability coefficient of test t. Tables 4, 5, 6, and 7 contain the information pertaining to the outcome of the item analyses. Although it was not the criteria for rejection of items from the test, the proportion of students getting each answer correct was also reported. A point-biserial correlation cutoff value of .25 was used in rejecting test items. That is, items having a correlation with the composite score of less than .25* were eliminated.

Table 7

Kuder-Richardson Coefficients for the Reading and Achievement Tests

	Reading	Test 1	Test 2
N	193	193	193
Initial Number of Items in Test	.50	25	25
Final Number of Items in Test	45	19	18
Mean of Final Test	29.6	10.8	9.8
S.D. of Final Test	7.5	4.1	3.6
K.R. R for Final Test	.88	.77	.79

Hypothesis Testing

The boys and girls in each of the eight science classes were assigned to work in one of five treatment groups. The students worked independently - without the aid of the classroom teacher or other students - and at their own rates.

The student responses in each of the five treatment groups were gathered from eight different science classes (See Table 2). This

* Point biserial correlations of .25 or more with a sample of this size have a chance probability of less than .01.

pooling of student responses from different classes to form treatment groups tended to make the analyses less sensitive to differences that may have arisen as a result of an individual teacher, classroom, or class period of the school day and more sensitive to the experimental treatment. The number of students in treatment groups one through five were 39, 38, 40, 37, and 39 respectively.

The analysis of the data for the first nine hypotheses was performed with a computer program designed to perform multivariate (two or more criteria) and univariate (single criteria) analyses of variance and covariance (Clyde, Cramer, and Sherrin, 1966). For three of the hypotheses (H_1, H_4 H_7), the analysis was a multivariate analysis of covariance which used relevant and incidental content achievements as the criteria, and science content reading achievement as the covariate. The multivariate test of significance with the computer program used Wilks' lambda criterion and Rao's F test approximation.

Table 8

Means and Standard Deviations of Criteria and Covariates

Treatment Group		Relevant	Incidental	Reading
1	Mean	6.8	6.4	28.7
	S.D.	2.4	2.4	8.3
2	Mean	8.6	9.4	31.0
	S.D.	2.7	3.2	6.7
3	Mean	10.7	10.0	27.8
	S.D.	3.3	3.3	7.7
4	Mean	13.0	11.1	30.5
	S.D.	2.8	3.0	6.5
5	Mean	14.9	12.4	30.3
	S.D.	3.1	3.0	7.8

The matrix of contrasts for the three multivariate hypotheses was:

H_1 =	0.0	-3.	-1.	1.	3.
H_4 =	1.	-1.	0.0	0.0	0.0
H_7 =	0.0	3.	-1.	-1.	-1.

The rows of the matrix represent the contrasts or comparisons, and the columns represent the treatment groups. Since the sum of the products of the elements in any two of the rows does not equal zero, the contrasts are not orthoginal or independent; i.e., for H_1 and H_4 :

$$(0.0)\ (1.0) + (-3.0)\ (-1.0) + (1.0)\ (0.0) + (3.0)\ (0.0) \neq 0.0$$

This relationship also holds true for H_1 and H_7, and H_4 and H_7. Since these contrasts or comparisons are not independent, it follows that they are retesting at least a portion of the same variance. When repeated tests of significance are made between the same variances or combinations of the variances, the probability value associated with all of the tests except the first is dubious or open to question (Winer, 1962; Hays, 1966). The reasoning behind the question is that an investigator testing 20 non-significant relationships would expect one of the 20 tests to be significant at the .05 level by chance alone. The first and most important hypothesis of the study (H_1) was tested first, and the probability value associated with the test was reliable. The other two multivariate hypotheses (H_4 and H_7) were then tested. The F values associated with these two tests were reported, and were discussed in light of the findings of H_1.

The first hypothesis was tested using an equally spaced linear trend in a multivariate analysis of covariance. The analysis tested the linear trend of the amount of change of achievement per unit change of complexity. The linear trend was tested across treatment groups two, three, four, and five, and the differences in the complexities of treatments were assumed to be evenly incremented. The results of the test for H_1 were:

$F = 62.2$

Degrees of freedom = 2 and 148

Probability $P < .01$

The second and third hypotheses (H_2 and H_3) were a subset of H_1 and used a univariate analysis of covariance. The second hypothesis (H_2) tested the linear trend of the amount of change of relevant achievement per unit of complexity, and the third hypothesis (H_3) tested the linear trend of the amount of change of incidental achievement per unit of complexity. Tables 9 and 10 contain the information pertaining to the testing of these hypotheses, and Figure 3 is a graph of the means of the two achievement variables versus complexity for the four treatment groups.

Table 9

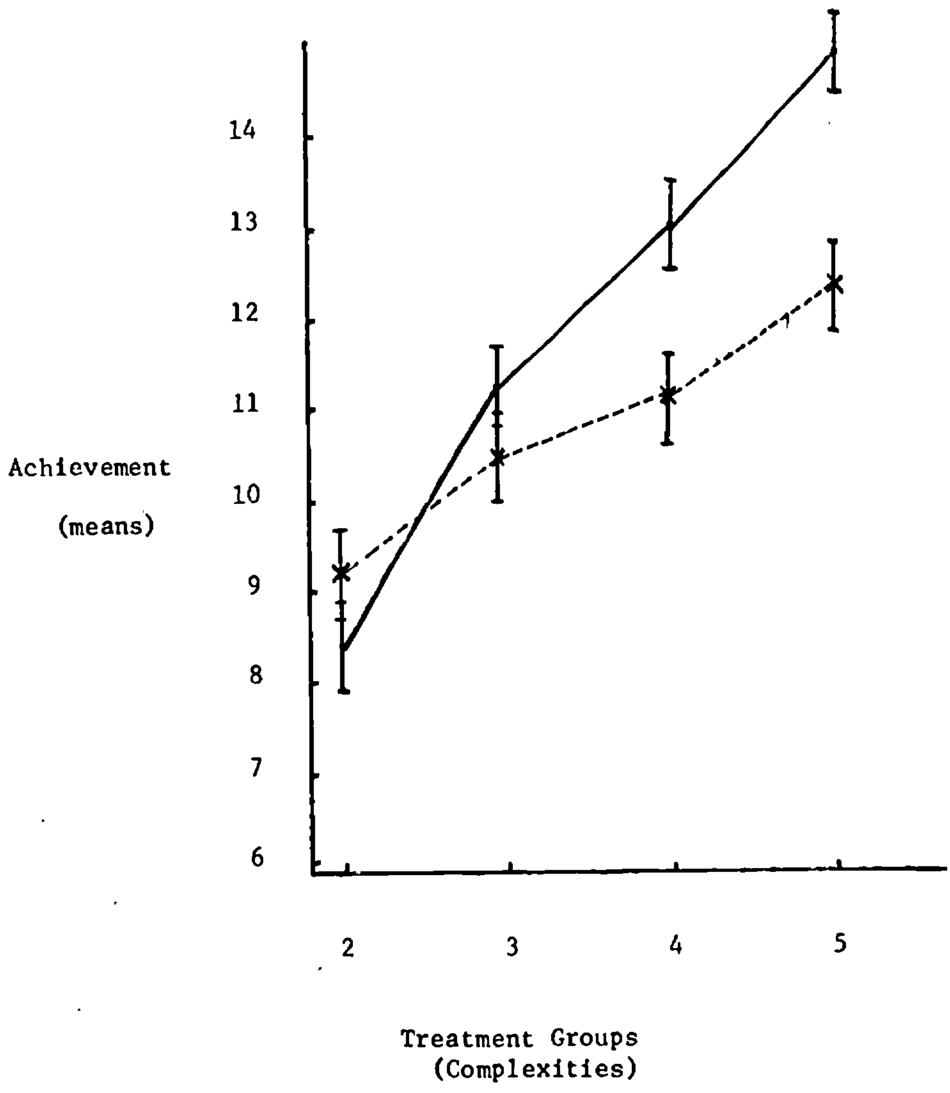
Means and Significance Test Information for H_2

Treatment Group	2	3	4	5
Adjusted Mean	8.4	11.2	12.9	14.8
Unadjusted Mean	8.6	10.7	13.0	14.9
N	38	40	37	39

$F = 124.6$
Degrees of Freedom = 1 and 149
Probability $P<.01$

Figure 3

Graph of Adjusted Means of Relevant and Incidental Achievement Versus Complexity



. = relevant achievement means

x = incidental achievement means

I = standard error of the mean

Table 10

Means and Significance Test
Information for H_3

Treatment Group	2	3	4	5
Adjusted Mean	9.2	10.4	11.0	12.3
Unadjusted Mean	9.4	10.0	11.1	12.4
N	38	40	37	39

F = 23.75
Degrees of Freedom = 1 and 149
Probability $P<.01$

The fourth hypothesis (H_4) was tested using a two-group multivariate analysis of covariance. The two achievement scores were used as the criteria and the reading score as the covariate. The comparison was between treatment groups one and two and was a test of the effect that the programmed instruction materials had on the learning (achievement) of the subject content: that is, was there any difference in the achievements of the two groups of students when one group had no instruction and the other group had the programmed instruction? An F value of 9.85 with 1 and 74 degrees of freedom was obtained. Had this two-group comparison been orthoginal to the other comparisons, the probabilitv value would have been $P < .01$.

The contrast for H_7 tested for differences between the achievement of treatment group two and the average achievement of treatment groups three, four, and five. The question asked was: Did differences exist between the achievement of two groups of students when both groups experienced the same programmed instruction sequence but one group responded (verbalized) to questions interspersed in the programmed instruction? Thirty-eight students (group two) experienced only the programmed instruction and 116 students (groups three, four, and five) experienced the programmed instruction with questions interspersed. An F value of 45.7 with 1 and 151 degrees of freedom was obtained. Had this test been made independently of the other tests, the probability value would have been $P < .01$. For discussion purposes, additional information relevant to H_8 and H_9 was obtained from Tables 2 and 3 and Figure 3.

The regression equations developed to test H_{10} and H_{11} were, by Scheffe's (1959) definition, an analysis of covariance. The tenth hypothesis (H_{10}) tested the relationship between the number of errors made in responding to the questions interspersed in the subject content and relevant subject content achievement, and H_{11} tested the relationship between the number of errors made in responding to the questions interspersed in the subject content and incidental subject content achievement.

The questions used to test H_{10} and H_{11} were of the same mathematical form with the appropriate variables substituted into the equations to test the hypotheses. The equations and procedures used were those described by Ward and Bottenberg (1963) for use when the covariate (achievement) may be affected by the treatment.*

---

* The regression equations and their associated significance tests were completed using a computer program developed by Veldman (1967), with modifications as suggested by Love (1968).

The three equations used were:

Equation 1 $Y = B_3X^{(3)} + B_4X^{(4)} + B_5X^{(5)} + B_6X^{(6)} + B_7X^{(7)} + B_8X^{(8)} + e$

Equation 2 $Y = B_3X^{(3)} + B_4X^{(4)} + B_5X^{(5)} + B_9X^{(9)} + e$

Equation 3 $Y = B_1U + B_9X^{(9)} + e$

B's = Standardized regression coefficients developed by the computer to minimize the value of e, the error term.

Y = criterion vector, of dimension $n(n_3 + n_4 + n_5$: n_3 = number of students in treatment group 3; n_4 = number of students in treatment group 4; n_5 = number of students in treatment group 5) which has as its elements the number of errors made in responding to the interspersed questions.

$X^{(3)}$ = 1 if corresponding element in Y is from a student in treatment group 3, and 0 otherwise.

$X^{(4)}$ = 1 if corresponding element in Y is from a student in treatment group 4, and 0 otherwise.

$X^{(5)}$ = 1 if corresponding element in Y is from a student in treatment group 5, and 0 otherwise.

Z = achievement score vector. Incidental achievement in equations used with H_{10} and relevant achievement in equations used with H_{11}.

m_3 = mean achievement of students in treatment group 3. H_{10} = 10.05, H_{11} = 10.75

m_4 = mean achievement of students in treatment group 4. H_{10} = 11.16, H_{11} = 13.05

m_5 = mean achievement of students in treatment group 5. H_{10} = 12.41, H_{11} = 14.92

$X^{(6)}$ = $(X^{(3)} * Z - m_3X^{(3)})$, a vector which has as elements the differences between the mean achievement score for treatment group 3 ($m X^{(3)}$) and achievement scores of students ($X^{(3)} * Z$) who were in treatment 3; and 0 for students in treatment groups 4 and 5.

$X^{(7)} = (X^{(4)} * Z - m_4 X^{(4)})$, a vector which has as elements the differences between the mean achievement score for treatment group 4 $(m_4 X^{(4)})$ and achievement scores of students $(X^{(4)} * Z)$ who were in treatment 4; and 0 for students in treatment groups 3 and 5.

$X^{(8)} = (X^{(5)} * Z - m_5 X^{(5)})$, a vector which has as elements the difference between the mean achievement score for treatment group 5 $(m_5 X^{(5)})$ and achievement scores of students $(X^{(5)} * Z)$ who were in treatment group 5; and 0 for students in treatment groups 3 and 4.

$X^{(9)} = X^{(6)} + X^{(7)} + X^{(8)}$

$U = X^{(4)} + X^{(5)} + X^{(6)}$, a unit vector.

A two-step procedure was used in testing both H_{10} and H_{11}. The first step tested whether the amount of change in making errors per unit change in achievement was the same for all three treatment groups. Equation 1 was used as the full regression model. By letting $B_6 = B_7 = B_8 = C$ (a constant) in equation 1, equation 2 was obtained and served as the reduced regression model. The results of the two tests appears in Table 11 and indicates that the amount of change in making errors per unit change in achievement (both relevant and incidental) was not significantly different, over the range of measured scores, for all three treatment groups. This is better known as homogenity of regression.

The second step of the procedure was carried out after it was determined that there was homogenity of regression. The second step tested whether the regression lines, shown in the preceding test to be parallel, were coincident. Equation 2 was used as the full regression model. Letting $B_3 = B_4 = B_5 = 0$ (zero) in equation 2, equation 3 was obtained and served as the reduced regression model. The results of this test appear in Table 12 and indicate that the regression lines are not coincident. The mean number of errors made by students in each of the treatment groups appears in Table 13, and a graph of the mean number of errors made by students versus the mean achievement (relevant and incidental) for students in each of the treatment groups appears in Figures 4 and 5.

Table 11

Test of Homogenity of Regression for H_{10} and H_{11}

	H_{10}	H_{11}
R Squared, Full Model	.4559	.4513
R Squared, Reduced Model	.4556	.4503
R Squared Difference	.0003	.0010
F Ratio	.027	.109
Degrees of Freedom	2 and 110	2 and 110
Probability	$P < .97$	$P < .89$

Table 12

Test of Coincidence of Parallel Regression Lines for H_{10} and H_{11}

	H_{10}	H_{11}
R Squared, Full Model	.4556	.4503
R Squared, Reduced Model	.1048	.0881
R Squared Difference	.3508	.3621
F Ratio	36.09	36.89
Degrees of Freedom	2 and 112	2 and 112
Probability	$P < .01$	$P < .01$

Table 13

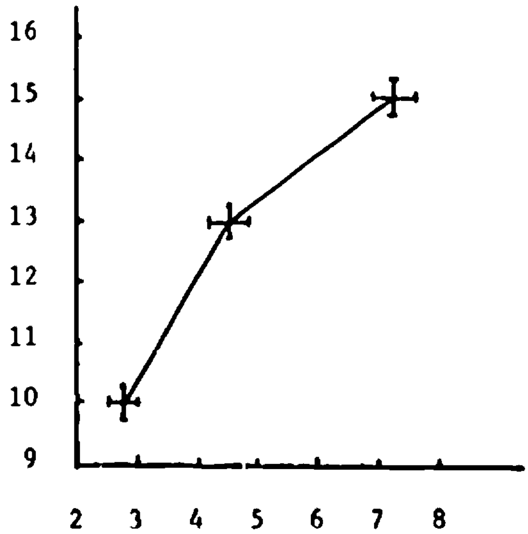
Number of Errors Made in
Responding to Questions

Treatment Group	3	4	5
Mean Number of Errors	2.8	4.6	7.2
Standard Deviation	2.0	2.9	2.7

Figure 4

Graph of Relevant Content Achievement Versus the
Number of Errors Made in Verbal Responses

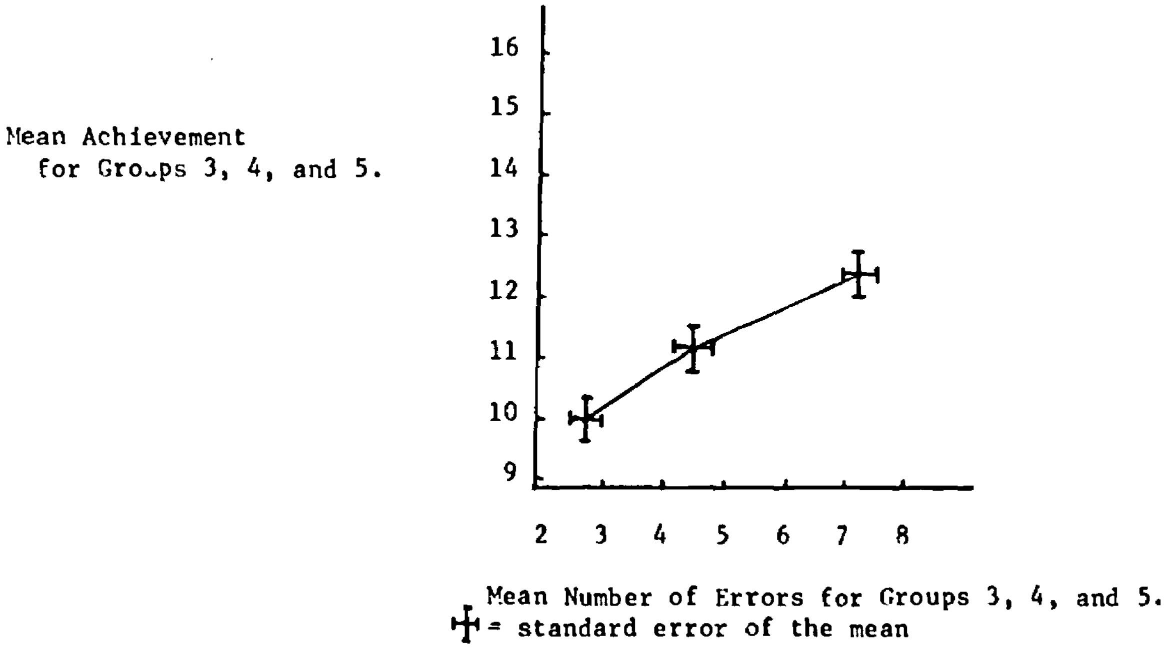
Mean Achievement
for Groups 3, 4, and 5.



Mean Number of Errors for Groups 3, 4, and 5.
= Standard error of the mean

Figure 5

Graph of Incidental Achievement Versus the Number of Errors Made in Verbal Responses



The amount of time that was taken by the students of treatment groups two, three, four, and five was measured. The means and standard deviations of the times for each treatment group appears in Table 14. To test H_{12}, a single classification analyses of variance was performed with this data. The results were:

$F = 5.56$

Degrees of Freedom 3 and 150

Probability $P < .01$

A graph of the time taken to complete the instructional materials versus the complexity of questions appears in Figure 6.

Table 14

Time Needed to complete
the Instructional Sequence

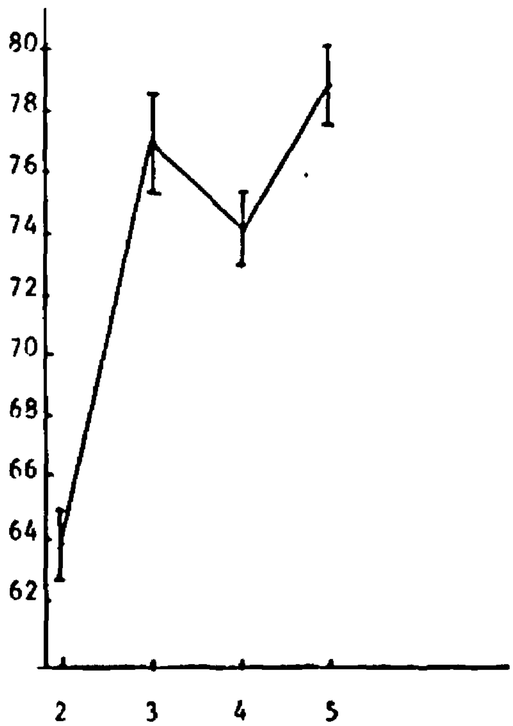
Treatment Group	2	3	4	5
Mean	63.7	76.6	73.9	78.4
Standard Deviation	15.5	22.1	13.5	16.1
N	38	40	37	39

F = 5.56
Degrees of Freedom = 3 and 150
Probability P < .01

Figure 6

A Graph of the Time Taken to
Complete the Instructional Materials
Versus the Complexity of Questions

Mean time used to complete instruction (in minutes)



Treatment Groups
(complexities)

I = standard error of the mean

Chapter IV
Discussion and Conclusions

Experimental Design

The experimental design was developed: (1) To measure those variables that were related to the hypotheses (relevant and incidental achievement, time taken to complete the instructional materials, and errors made in verbal responses); (2) To measure and statistically control for a variable (reading ability), and; (3) To gather the data so that error variance had a chance of occurring with equal probability in all treatment groups. Other variables, such as student interaction or student-teacher interactions which may have biased the experimental outcomes, were looked for by the investigator in his classroom visits and in his conversations with the teachers. No potentially biasing or otherwise disruptive events were detected by the investigator. Therefore, it is assumed that none occurred, or if they did occur, that they affected each of the treatment groups equally.

Reliability and Validity of the Measuring Instruments

Three tests were developed by the investigator for use in the study. Initially, the two achievement tests each contained 25 items and the reading test contained 50 items. The item analysis procedure used with the student responses on the tests reduced the number of items in the tests to 18, 19, and 45. This procedure resulted in three tests, each of which was composed of items that were significantly correlated with the composite score for that test. This meant that the items retained following the item analysis were related in that they measured student ability in essentially the same subject or ability area. Those items that were deleted from the tests had low correlations ($r < .25$) with the composite scores and tended to measure some characteristic of the subject matter other than what was measured by the final test. When items are deleted from a test, there is a possibility that a specific student ability or knowledge that was measured only by a deleted item goes unmeasured. The probability of deleting an item that measures a unique ability or knowledge that is independent of the other knowledges or abilities measured by the test appears low if one considers the interrelatedness of the concepts involved in the subject content and reading materials of the study.

The item deletion process of the item analysis can also be viewed in another way. For each of the three tests, as the items having low correlations with the composite score were deleted, the reliability coefficient for that test increased. The process of attempting to increase the value of the reliability coefficient of a test is consistent with good test development (Guilford, 1965) in that it reduces the portion of the total test score variance that is error variance. When the deleted items from each test were treated as a new test and reanalyzed, the resulting reliability coefficients were low ($R < .2$). This indicated that none of the tests contained subtests or sets of items that measured a common characteristic that had been removed from the original test in the item elimination process.

The Effect of the Covariate in the Analyses.

The first nine null hypotheses of the study required the use of a covariance statistical design. The covariate measure used in the analyses was the science content paragraph comprehension. This type of statistical design meant that the results of the analyses should be interpreted as though the students in each of the treatment groups could read and interpret science content equally well - as measured by the science content paragraph comprehension test. This statistical design was used to control for differences in reading ability which avoided problems that may have arisen in interpreting the results of the study. Each student's performance in the study was highly dependent upon his ability to read and interpret science subject content. If, by chance, the students in one of the treatment groups was able to read significantly better than the students in another treatment group, and reading ability was not statistically controlled for, the interpretation of differences in achievement would have been difficult, if not impossible, to interpret.

The adjustment made in the criterion scores as a result of the covariate score was so small that when one of the statistical tests (H_1) was completed, both with and without the covariate, the results of the analyses were significant at the same level. The amount of adjustment made in a criterion score by a covariate is a function of two numbers (Snedecor, 1962). The first number is the raw score regression coefficient for the criteria regressed on the covariate (.217 for relevant achievement and .109 for incidental achievement). The second number is the difference between the grand mean of the covariate scores and the mean covariate score for the cell in which the criteria are being adjusted. The equation used was:

$$Y_{adj} = Y_c - B\ (X_g - X_c)$$

Table 15

Group Means Associated With H_4

	Treatment Group 1 (Did not work with programmed instruction)	Treatment Group 2 (Worked with programmed instruction)
Adjusted mean on incidental achievement	6.66	9.12
Adjusted mean on relevant achievement	6.69	8.68
N	39	38

The statistical test for H_4 was not independent of the other statistical tests of the study. Therefore, the probability value associated with the test for H_4 was not representative. Based on the magnitude of the F-ratio and the magnitude of the differences in the means of the criteria, the null hypothesis that there was no difference in the achievements of the two groups was rejected. The alternate hypothesis that the treatment group that completed the programmed instruction materials (Group 2) had significantly higher achievement (relevant and incidental) than the group that did not complete the programmed instruction materials (Group 1) was accepted.

The rejection of H_4 was not unexpected. The programmed instruction materials, with the exception of the changes introduced by the investigator, were developed and published by a reputable company (Encyclopaedia Britannica Press) for use by seventh grade students.

where,

Y_{adj} = adjusted cell mean of the criterion score for cell c.

Y_c = unadjusted cell mean of the criterion score for cell c.

B = raw score regression coefficient.

Xg = grand mean of covariate scores.

X_c = mean score of covariate for cell c.

The grand mean of the covariate scores (reading) was 29.67. The cell mean scores of reading were relatively uniform (see Table 8) and small differences existed between each of the cell means and the grand mean. The small differences between the cell means and the grand mean and the low value of the regression coefficients resulted in small adjustments in the cell means of the criteria (see Tables 9 and 10). The covariate was a necessary precaution, although not needed in the final analyses.

The Hypotheses

The hypotheses will not be discussed in the order in which they were stated or tested. The investigator believes that a more meaningful interpretation of the hypotheses may be obtained by first discussing H_4 and H_7 which dealt with the overall affects of the experimental treatment and training, and then discussing H_1, H_2, H_3, H_{10} H_{11}, and H_{12} which contain more specific information on the effect of the experimental treatment.

The fourth hypothesis (H_4) was a test of the effect on achievement of completing the programmed instruction materials. The F-ratio for the statistical test was 9.85 and the mean values for the two criteria measures appear in Table 15.

Since the instructional materials are used elsewhere in similar educational situations to raise student achievement levels, there is little reason to believe that these same materials would not raise achievement levels in the environment of this study.

The object of H_7 was to determine whether students who experienced both the programmed instruction and the questions had achievements (relevant and incidental) that were significantly different than students who experienced only programmed instruction. The F-ratio associated with the statistical test was 45.76, and the mean achievements for the two treatments are shown in Table 16.

Table 16

Group Means Associated With H_7

	Treatment Group 2 (Programmed instruction only)	Treatment Groups 3,4,& 5 (Programmed instruction and questions)
Adjusted mean on incidental achievement	9.22	11.27
Adjusted mean on relevant achievement	8.45	12.98
N	38	116

The statistical test for H_7 was not independent of the other statistical tests of the study, and the probability value associated with the statistical test was not representative. When considered collectively, the magnitude of the F-ratio, and the magnitude and direction of the differences in the means ot the achievement measures reveal what appears to be a definite relationship. Therefore, the null hypothesis was rejected. The alternate hypothesis accepted was that the treatment group that completed the programmed instruction materials and responded to questions scored significantly higher on relevant and incidental subject content achievement than the group that completed only the programmed instruction materials.

The experimental results leading to the rejection of H_7 were similar to the findings of Hershberger (1963), Hershberger and Terry (1964), and Estes (1960) who investigated the effect on learning of placing questions within the instructional materials to which students were required to respond. Although these studies differed from the present study in that they were not conducted in the science subject content area, and the objectives of the studies were not to evaluate the effect of complexity of post-instructional questions on achievement, the results of the studies were comparable. That is, when students were required to respond to questions that were placed in the instructional materials, learning was facilitated.

The two alternative hypotheses associated with H_4 and H_7 suggest that: (1) Students can learn (score higher on an achievement test) by reading programmed instruction materials: (2) Students who respond to questions of different complexities placed within programmed instruction materials learn more than students who read only the programmed instruction materials. It would appear that developers and publishers of programmed instruction materials should make greater use of the research findings cited and confirmed by this study. Some publishers of programmed instruction texts do embed questions in their instruction materials. For example, the I.B.M. programmed instruction series in

FORTRAN (1965) makes use of a "picture book" and a "question book". Periodically in the programmed instruction series, students are sent to one of these books to answer questions. When the questions are answered, the student returns to the programmed text. TEMAC programmed learning materials in their 1969 series uses a "supplemental panels" booklet in the same way as the I.B.M. series uses picture books.

Although the degree of complexity of questions used to elicit verbalization was relatively constant for one treatment group, questions were of different complexities for each of the treatment groups. The hypotheses of the study were not dependent upon equal increments of complexity between the questions used at a given point in the instructional materials with any two groups. The intent was to utilize 13 sets (three questions per set - one for each level of complexity or treatment group) of questions that varied in complexity within each set. When all of the sets of questions were considered, there was a mean difference in the complexity of questions used with each of the treatment groups. The complexity of the questions used by the treatment groups, therefore, formed an ordinal scale.

For purposes of making multivariate statistical tests of H_1, H_2, and H_3, it was necessary to assume that the ordinal scale of the questions was an interval scale. Lacking more complete knowledge of the nature of the ordinal scale, the investigator chose the simplest type of interval scale and assumed that the increments in complexity were equal. That is, the average difference in the complexities of the interspersed materials used with any two adjacent treatment groups (2,3,4, and 5) were equal. The statistical testing for the linear trend in H_1, H_2 and H_3 were therefore considered appropriate knowing that non-equal increments in complexity could still lead to the identification of a positive trend across treatments if, in fact, it was present. The first null hypothesis (H_1) was: There is no relationship between the degree of complexity of questions producing student verbalization and the amount of increase in student achievement on relevant and incidental subject content when reading ability is statistically controlled. The statistical procedures used with H_1 tested for a linear relationship between the achievement variables and question complexity for the four experimental treatment groups. The statistical test yielded an F-ratio of 62.2 ($P < .01$) and the null hypothesis (H_1) was rejected. Since the F-ratio was large, and the graphs of the two achievement variables versus the complexities of questions indicated a positive trend (Figure 3), tests for higher order curve fitting was not done. The statistical test with which the null form of H_1 was rejected was a directional test. The rejection of the null hypothesis with this test implies that a positive trend best describes the relationship between the variables. Therefore, the alternate hypothesis accepted is that there is a positive relationship between the amount of change in student achievement (relevant and incidental) and the complexity of questions.

The positive trend associated with H_1 also described the relationship between each of the achievement variables and the complexity of questions. The F-ratio for the test of H_2 was 124.60 ($P < .01$) and for H_3 was 23.75 ($P < .01$). The null form of H_2 and H_3 were rejected. The rejection of the null hypotheses with this test implies that a positive trend best describes the relationship between the variables for H_2 and H_3. Therefore, the two alternate hypotheses accepted were: (H_2) there is a positive relationship between the amount of change in student achievement (relevant) and the complexity of questions, and; (H_3) there is a positive relationship between the amount of change in student achievement (incidental) and the complexity of question.

A graph of the means of achievement versus complexities of questions for the four treatment groups tested by H_2 and H_3 appears in Figure 3. The two sets of points that were plotted come close to forming the straight line that statistically best describes the relationship between the variables (assuming unit increments in complexity). The two lines have different slopes. The difference in slopes may be interpreted as saying that the rate of increase in relevant achievement per unit of question complexity is greater than the rate of increase in incidental achievement per unit of question complexity. This difference may have been due to students reading and working with a larger number of relevant subject content tasks (frames and questions) than incidental subject content tasks, or that more time was spent working on relevant subject content than on incidental content. It appears that students may have spent more time working on and had more experience working with the relevant subject content than the incidental subject content since the relevant content appeared in both the programmed instruction and the interspersed questions, and the incidental content appeared only in the programmed instruction.

At the lowest level of complexity (Figure 3), the two lines that describe the two sets of points appear to meet or cross. Students positive trend best describes the relationship between the variables for H_2 and H_3. Therefore, the two alternate hypotheses accepted were: (H_2) there is a positive relationship between the amount of change in student achievement (relevant) and the complexity of questions, and; (H_3) there is a positive relationship between the amount of change in student achievement (incidental) and the complexity of question.

A graph of the means of achievement versus complexities of questions for the four treatment groups tested by H_2 and H_3 appears in Figure 3. The two sets of points that were plotted come close to forming the straight line that statistically best describes the relationship between the variables (assuming unit increments in complexity). The two lines have different slopes. The difference in slopes may be interpreted as saying that the rate of increase in relevant achievement per unit of question complexity is greater than the rate of increase in incidental achievement per unit of question complexity. This difference may have been due to students reading and working with a larger number of relevant subject content tasks (frames and questions) than incidental subject content tasks, or that more time was spent working on relevant subject content than on incidental content. It appears that students may have spent more time working on and had more experience working with the relevant subject content than the incidental subject content since the relevant content appeared in both the programmed instruction and the interspersed questions, and the incidental content appeared only in the programmed instruction.

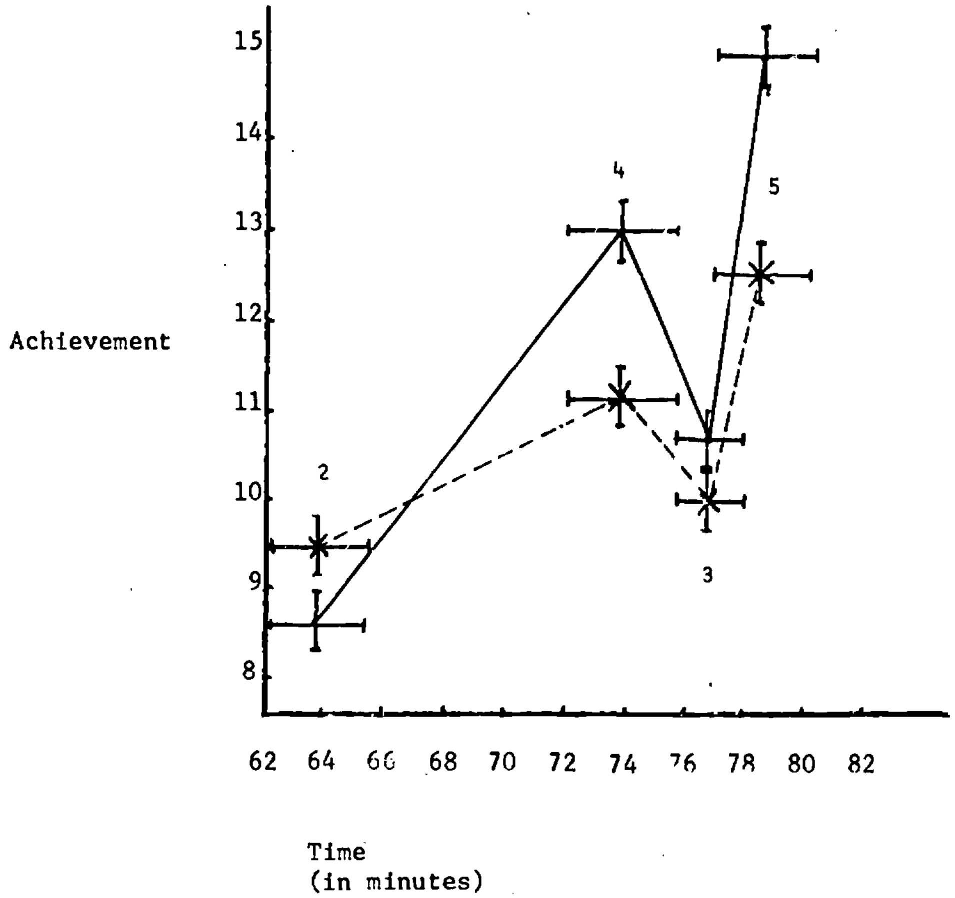
At the lowest level of complexity (Figure 3), the two lines that describe the two sets of points appear to meet or cross. Students working with the programmed instruction materials at the lowest level of complexity did not answer questions. Students in this treatment group read explanations of the subject content that were dealt with in the questions answered by other treatment groups. That is, students in treatment group 2 read the programmed instruction materials and read explanations of the relevant subject content while students in the other treatment groups read the programmed instruction materials and responded to questions. As a result of the difference in experience, an investigator might expect to obtain smaller differences in the two types of achievement for treatment group two than with the other treatment groups.

The null hypothesis (H_{12}) that there were no significant differences in the amount of time taken by students in the treatment groups to complete the instructional materials was rejected ($F = 5.56, P<.01$). As the complexity of questions used to elicit verbalization increased, the amount of time taken by students to complete the instructional materials tended to increase (see Figure 6). Figure 7 is a graph of the mean relevant and incidental achievements versus the time needed to complete the instructional materials. The points on the graph, with one exception, indicate a positive relationship between the achievement variables and the time spent working on the instructional materials. As the amount of time taken by the students in the treatment groups increased, achievement increased. The exception, treatment group three, took longer to complete the instructional materials than the students in treatment group four, but scored lower on the achievement tests than the students in treatment group four. As was shown with H_1 the increments in achievement across treatments was relatively uniform. The increments in time across treatment groups was not even. When the times and achievements of the treatment groups were considered collectively, the students of treatment group three took more time than was expected to complete the instruction materials. Since the only difference in treatment was the complexity of questions, it is assumed that this difference in time was a function of the differences in questions. A possible explanation of the differences in the time taken to complete the instructional materials was obtained from the students who participated in the pilot studies. Following the collection of data in the pilot studies, several students from each of the treatment groups were casually asked their opinions of the questions for which they wrote answers. One reply of the students in treatment group three was, "The questions were so easy that I didn't trust you and I had to reread each question several times to make sure I answered it correctly." It should be noted at this point that the questions to which treatment group three responded were the least complex of the questions to which students responded. This is a possible explanation of why the time-achievement point on the graph does not fit more closely to the positive trend of the other points on the graph.

The information available from Figures 3, 6, and 7 can be evaluated collectively. Figure 3 and the hypothesis associated with it indicate that as the complexity of questions used to elicit verbalization increases, student achievement increases. Figure 6 shows that as the complexity of questions used to elicit verbalization increases, the amount of time spent working on the instructional materials also increases. A coupling of these findings (Figure 7) indicates that the

Figure 7

A Graph of Achievement Versus Time



. = Relevant achievement
x = Incidental achievement
= Standard error of the mean
Numbers represent the treatment groups in which the date was gathered.

increases in achievement across groups was accompanied by an increased amount of time spent working which, in turn, was a function of differences in the complexities of the questions. The more complex questions generated student behaviors or attitudes that were time-consuming, and allowed students to score higher on the two achievement measures. This study was not designed to determine what types of student behaviors took place as a function of questions of different complexities or of the effect of different behaviors ... ter achievement.

The same two-step procedure wa used for testing both H_{10} and H_{11}. The first step tested for homogeneit of regression and the second step tested for the coincidence of the regression lines. The variables involved in the testing of H_{10} were err scores and incidental achievement, and the variables involved in the testing of H_{11} were error scores and relevant achievement. F-ratio for the test of homogeneity of regression for H_{10} was (P < .97) and for H_{11} was .109 (P < .89). Both H_{10} and H_{11} pas test for homogeneity of regression. This meant that over the observed range of achievement that the amount of change in error score p of achievement score (both relevant and incidental) were si ntly different for all three treatment groups. The second ste e testing procedure determined whether the three regression lin oincident or whether the three treatments were equally ef (produced the same number of errors) over the observed ran hievements. The F-ratio for the test of coincidence of re lines for H_{10} was 36.09 (P < .01) and for H_{11} was 36.8 (P < .01). Both H_{10} and H_{11} failed the second test. This meant that there were differences in the mean number of errors made by students in the three treatment groups. Following the testing of H_{10} and H_{11}, the regression weights, the correlations between the predictor variables, and the correlations between the predictor variables and the criterion variable were checked to determine whether all of the predictors in the mathematical models were predicting a portion of the criterion variance. The null hypothesis (H_{10}) that there is no relationship between the number of errors made in verbal response to questions and student achievement on incidental subject content was rejected. The null hypothesis (H_{11}) that there is no relationship between the number of errors made in verbal response to questions and student achievement on relevant subject content was also rejected.

In earlier statistical tests and discussion, it was shown that both relevant and incidental achievement increase positively across treatment groups. As the complexity of questions used to elicit verbalization increases, both relevant and incidental achievement increases. Similarly, the mean number of errors made by students in responding to the interspersed questions increased across treatment groups. The mean number of errors made by the students of treatment groups three, four, and five were 2.8, 4.5, and 7.2, respectively. A graph (Figure 5) of the mean number of errors made in responding to questions versus the mean incidental achievement for treatment groups three, four, and five has a positive slope. This positive slope indicates that the increase in incidental achievement across groups was accompanied by an increase in the number of errors made in making verbal responses to questions placed in the instructional materials. Therefore, since H_{10} was rejected, the alternate hypothesis accepted is that across treatment groups there is a positive relationship between the number of errors made in responding to questions and student achievement on incidental subject content.

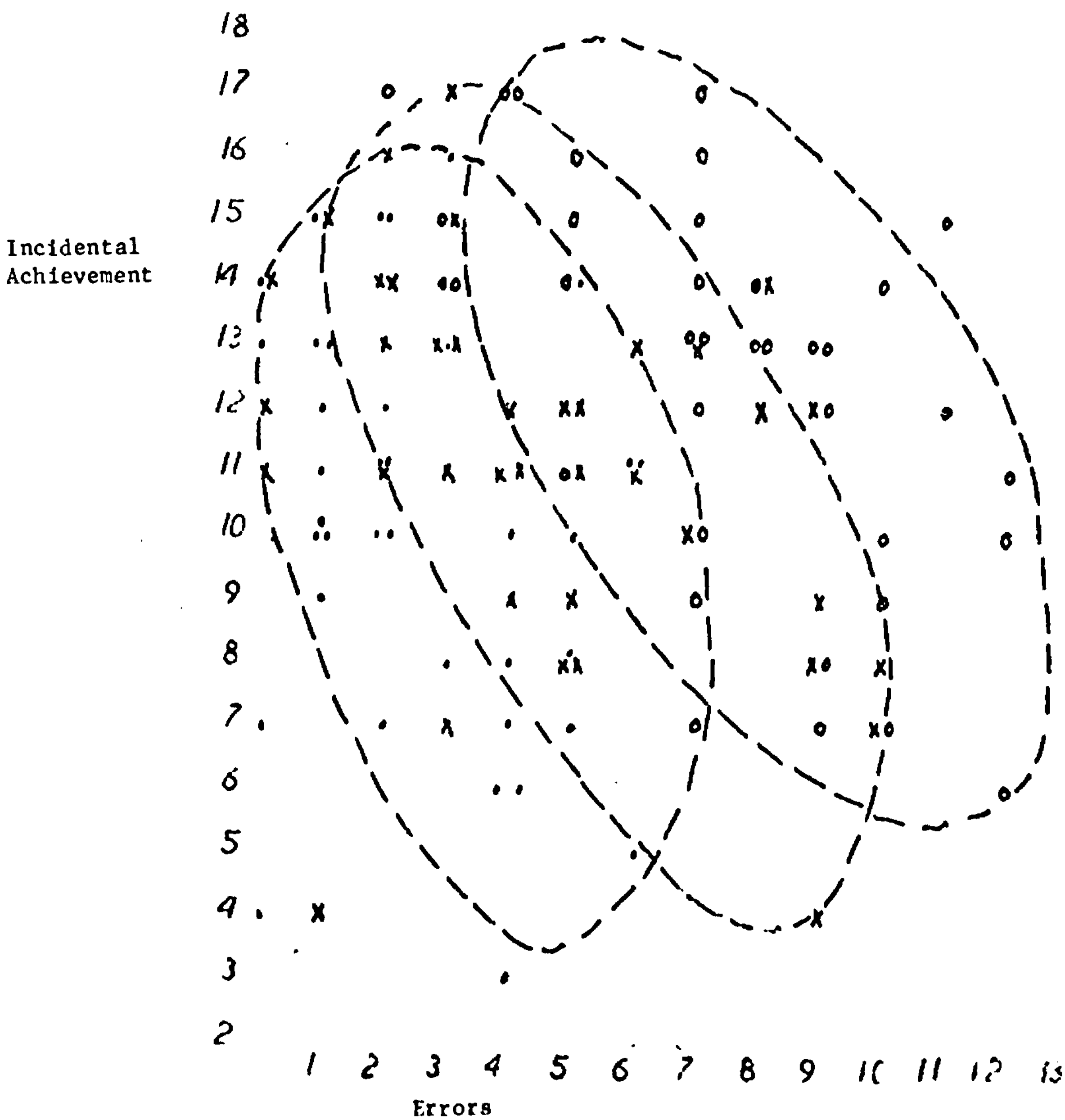
A graph (Figure 4) of the mean number of errors made in responding to questions versus the mean relevant achievement for treatment groups three, four, and five also has a positive slope. The positive slope indicates that an increase in relevant achievement across groups is accompanied by an increase in the number of errors made in making verbal responses to questions placed in the instructional materials. Therefore, the accepted alternate hypothesis to H_{11} is that across treatment groups there is a positive relationship between the number of errors made in responding to questions placed in the instructional materials and student achievement on relevant subject content.

What might be considered a "traditional" negative correlation (product moment) between achievement and the number of errors made in responding to questions did exist in each of the three treatment groups. The average correlation between the number of errors and incidental achievement was -.39*, and between the number of errors and relevant achievement was -.33.* This meant that within each of the three treatment groups, those students making the fewest errors when working on the instructional materials tended to be high achievers and those students making the largest number of errors when working on the instructional materials tended to be low achievers. It should be noted that there were no significant differences in the mean I.Q.'s of students in the treatment groups. For the three treatment groups, the average correlation between I.Q. and relevant achievement was + .47, and between I.Q. and incidental achievement was + .50.* The correlations indicate that students having higher I.Q.'s tend to make fewer

*Average correlations were derived by transforming the r values to Fisher's Z coefficients, calculating the average Z coefficient, and transforming the average Z coefficient to the average r (Guilford, 1965).

Figure 8

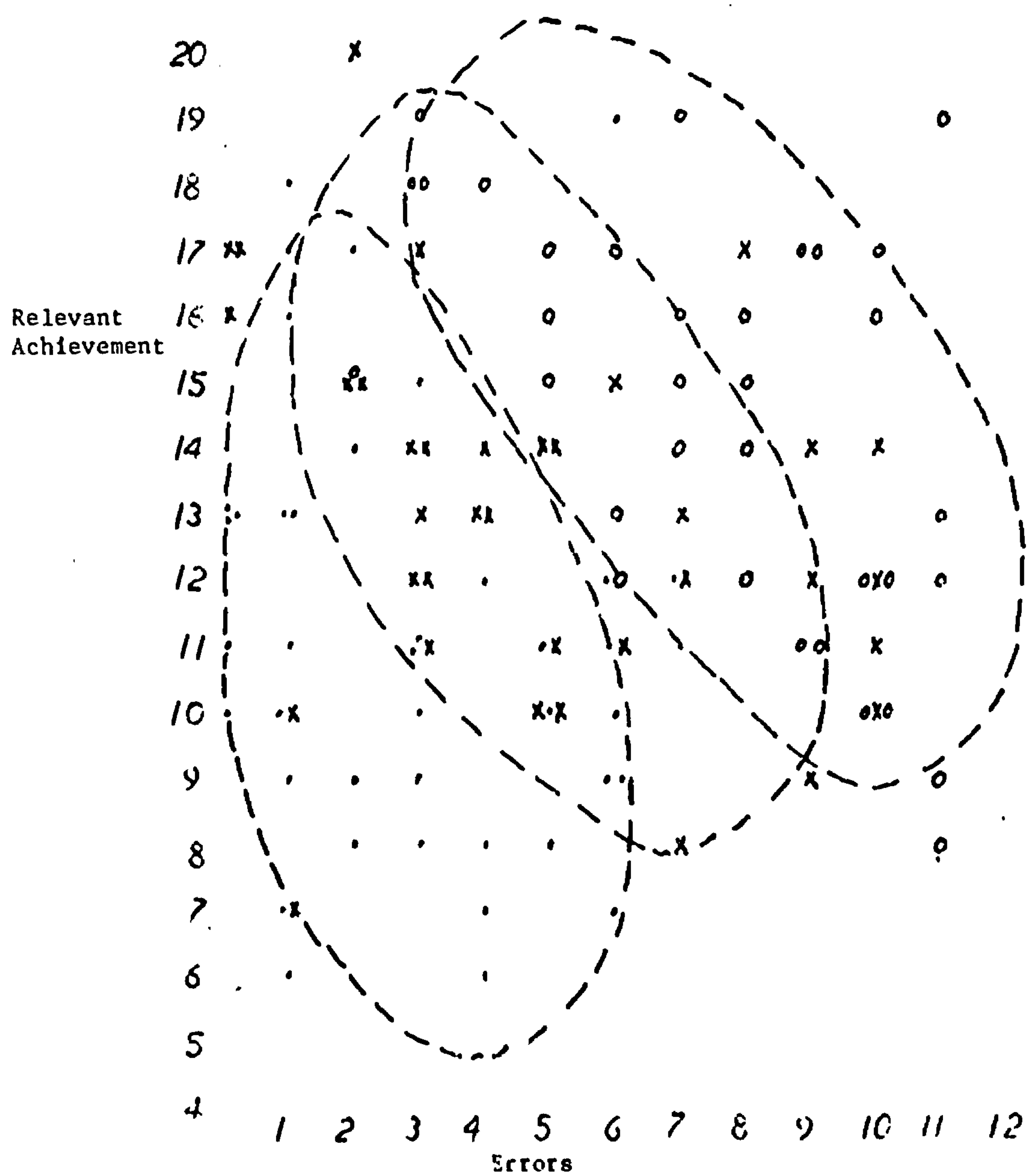
Scatter-gram of Incidental Achievement Versus Errors



• = Group 3
x = Group 4
o = Group 5

Figure 9

Scatter-gram of Relevant Achievement Versus Errors



• = Group 3
x = Group 4
o = Group 5

errors in the learning process and score higher on achievement measures, while students with lower I.Q.'s tend to make larger numbers of errors in the learning process and score lower on achievement measures.

The relationship between errors and achievement within treatment groups and across treatment groups can be seen graphically in Figures 8 and 9. Figure 8 is a graph of incidental achievement scores versus error scores for students in treatment groups three, four, and five, and Figure 9 is a graph of relevant achievement scores versus error scores for the same students. The treatment group to which each of the plotted points belongs appears in the legends of the graphs. The investigator enclosed the points from each of the three treatment groups in a figure. With few exceptions (the exceptions can be seen on the graphs), the points from each treatment group are enclosed within their own figure and there are relatively equal numbers of points from each of the treatment groups not falling within their own figure. Each of the figures is an approximate representation of a treatment group's scores on these two measures. In both of the graphs, the major axis (the longest axis) of each of the figures has a negative slope. This negative slope is characteristic of the negative correlation that exists between the number of errors made in responding to questions in the instructional materials and achievement.

The three figures in each of the graphs orient themselves in similar patterns. The figure that represents treatment group four is to the right and above that of treatment group three, and the figure that represents treatment group five is to the right and above that of treatment group four. Since the centers of the figures represent the approximate means of the two variables for each treatment group, it can be seen that across treatments, as achievement increases, the number of errors made also increases. From the graphs, it can be seen that within treatment groups there is a negative relationship between errors and achievement, and that across treatment groups there is a positive relationship between errors and achievement.

The negative correlation between errors and achievement found within the treatment groups is consistent with the early findings of Skinner (1958), who found that with programmed instruction, students who have high error rates in responding to frames tend to have low achievements. It should be remembered that Skinner's studies differed from this study in format and objectives. In Skinner's experiments, students learned by using a programmed instruction text. In this study, students learned by using a programmed instruction text and, in addition, were required to respond to questions placed throughout the text. Skinner then studied the amount of learning that took place as a function of the programmed texts, while this study dealt with the amount of learning (achievement) that took place as a function of the programmed instruction plus that which was a function of the questions built into the programmed text. Also, Skinner required students to respond to each programmed frame. In this study students read the correct answer that had been placed in the response blanks and responded only to the experimental questions. Skinner's studies and this study have, however, had a similar finding. Both studies have found that within a treatment, high error rates in responding during the instructional sequence are associated with low achievement. A possible explanation of this relationship appeared in this study. There was a negative correlation between I.Q. and errors and a positive correlation between I.Q. and achievement. High ability (I.Q) students tended to have high achievement and make low numbers of errors during instruction while low ability students tended to have low achievement and make high numbers of errors during instruction. These results were obtained in the three treatment groups that responded to questions. It would appear that within a given treatment, the negative relationship between errors and achievement is more of a function of the students' ability than low achievement is a function of high error rates alone.

Possible Explanations of the Differences in Achievement

With the exception of the control group, all students in the sample had the same instruction and, had it not been for the interspersed questions, would have been expected to have similar achievements. The interspersed questions required students to utilize the information they obtained through reading the programmed text materials and, as a result, achievement was facilitated. The students who responded to the more complex questions utilized a larger number of the characteristic parts of the instructional materials (factors) and had a larger number of experiences in working with the factors than the students who responded to the less complex questions. This introduces an experience and/or a practice factor into the study. The students responding to the more complex questions utilized a larger number of factors of the instructional materials (experience) in responding to questions and had a larger number of uses of each factor (practice) than did the students responding to the less complex questions. A number of studies have been reported in the literature that dealt with the effect of practice (Callatine and Warren, 1955; Bowne and Haygood, 1959) and experience (Morrisett and Hovland, 1959; Kersh, 1958) on learning and have shown that within limits, both experience and practice have facilitative effects on achievement. In this study, practice and experience were closely tied, and it is difficult to separate the amount of effect that each had on the experimental outcome. It is suspected that the increased amounts of practice and experience obtained by the students in the treatment groups that responded to the more complex questions contributed to the differences in their achievements.

When students worked with the pre-experimental programmed instruction materials in astronomy, the investigator observed that they worked through the pages sequentially and seldom, if ever, referred to the preceding pages. When students worked on the programmed portion of the experimental materials, the investigator observed that they worked in much the same way except when they encountered one of the interspersed questions. When a question was encountered, students tended to reread portions of the programmed text. Several students were observed rereading frames after they had answered the questions and checked the answers. Students were also observed rereading the questions, the answers, and the corrective feedback that accompanied the answers. These rereading behaviors are referred to by Rothkopf (1963, 1965) as one of a number of behaviors known as inspection behaviors.

According to Rothkopf's theory of inspection behaviors, the interspersed questions act as the stimuli and the students respond by exhibiting behaviors that are responsible for higher achievements. As an extrapolation from Rothkopf's reasoning to the present study, one might consider that more complex questions bring into play a larger number of stimuli (subject content factors) than less complex questions. If the greater number of factors (stimuli) brought into play a larger number of inspection behaviors, then these behaviors may have been responsible for the differences in achievement across treatment groups. The fact that students not only had higher achievement scores but also took more time to complete the instruction material supports this reasoning.

Rothkopf's inspection behaviors also give a possible explanation of why the interspersed questions facilitated relevant achievement more than incidental achievement. The students responding to more complex questions had an opportunity to direct their inspection behaviors towards a larger number of specific parts of the subject content (factors) than the students responding to the less complex questions. The inspection behaviors may have amounted to a sorting process. To find information in the text that was related to a factor, students would be required to sort (read) the preceding subject content until they found their answer. In the sorting process, students would review both the relevant and incidental subject content and as a result both relevant and incidental subject content achievement would be facilitated. As might be expected, since the inspection behaviors were topic specific, relevant subject content achievement would be facilitated more than incidental subject content achievement.

What types of inspection behaviors might be generated by a student finding that he had incorrectly answered one of the interspersed questions? One option is that a student may ignore the question and continue on in the instruction materials. In this study, it appears that if this occurred, it occurred with low frequency since students who made the largest number of errors tended to take longer to complete the instructional materials and had the highest achievements. Although a few students may have reacted negatively to the questions, the achievement and time measures indicate that in general just the reverse must have happened.

Student achievement may have been facilitated by a number of different types of behaviors or activities that followed the recognition of an incorrect answer. First, as in the studies by Meyer (1960) and Peterson (1960), students may have read the correct answer and obtained sufficient information to facilitate achievement. Second, students may have obtained information from the supplemental information (corrective feedback) that accompanied the correct answer. The supplemental information in this study was the explanation of how the correct answer was derived. Bryon and Rigney (1956), and Bryon, Rigney, and Van Horne (1957) have found that the feeding back of supplemental information (information relevant to the question but not including the correct answer) in response to an incorrect answer was sufficient to facilitate learning. Third, on finding they had given an incorrect answer, students may have reread sections of the instructional materials applicable to the question. Hershberger (1963, 1964) and Hershberger and Terry (1965) found that by permitting students to reread sections of the instruction materials germane to the questions asked, learning was facilitated. Through each of these three activities or behaviors, student would have been able to obtain the information necessary to facilitate learning.

The event that may initiate any of the three behaviors is a student's recognition that a response is incorrect. Since the students in the treatment groups responding to the most complex questions made the largest number of errors, one would expect the students in the same treatment groups to exhibit the largest number of these learning-centered inspection behaviors or activities. The students in the treatment groups responding to the most complex questions also had the highest achievement and spent the greatest amount of time completing the instructional materials. Therefore, there is a high probability that the facilitation of achievement was a function of one or more of these three behaviors or activities. Further, the differences in the achievement of the treatment groups may be a function of the differences in the number of these behaviors exhibited by the students in each of the treatment groups.

Summary

Assuming that the methodological and statistical design of the study were effectively carried out, there were four major findings of the study. These were:

1. As the complexity of the questions to which students responded increased, student achievement on relevant and incidental subject content increased.
2. As the complexity of questions to which students responded increased, the amount of time taken by students to complete the instructional materials increased.

3. As the complexity of questions to which students responded increased, the number of errors made by students in responding to the interspersed questions increased.

4. As the number of errors made by students in responding to the interspersed questions increased (across treatments), relevant and incidental subject content achievement increased.

These findings answered the basic questions asked in this study and, when considered collectively, give a relatively complete picture of what occurs when students are required to respond to questions of different complexities that are interspersed in the instructional materials. The students in the treatment groups that responded to the more complex questions took longer to complete the instructional materials, made more errors in their responses to the interspersed questions, and had higher achievements than students in the treatment groups that responded to the less complex questions.

In the development of the questions that were interspered in the subject content to produce student verbalization, the investigator utilized his own logic as to which factors came into play in answering the questions. The panel of judges, in evaluating the relative complexities of questions developed by the investigator, in effect agreed that there was a high probability that students would utilize the logic of the investigator and use a given number of factors or levels in formulating answers. The questions were formulated so that answering the least complex questions tended to involve the manipulation of a single factor while the answering of more complex questions tended to involve the manipulation of more than one factor. The students in the treatment groups that responded to the more complex questions had a larger number of experiences in working with selected portions of the subject content (factors) than students in the treatment groups that responded to the less complex questions. All students in the sample who worked with the programmed instruction materials had an equal opportunity to obtain information pertaining to the subject topic, but those students responding to the more complex questions were required to reuse or apply a greater number of the characteristic parts of the subject content. The increased number of experiences of working with the subject content introduced a practice or experience effect that was responsible for some or all of the differences in the achievements of the treatment groups.

Along with having higher achievements, the students in the treatment groups who responded to the more complex questions took more time to complete the instructional materials. That is, students responding to the more complex questions were required to go through a larger number of steps (manipulate a larger number of factors) in deriving the answers, and this is expected to have taken the additional student time.

There are a number of other student behaviors that may have occupied a portion of the additional time taken by the students in the treatment groups responding to the more complex questions that could have contributed to the differences in achievement. These behaviors may have included rereading the question, the answer, the corrective feedback, or the portions of the programmed text applicable to the question.

Making errors in responding to the interspersed questions might also have caused students to begin reading the programmed text materials more carefully the first time, which also would have consumed additional time.

Within each of the treatment groups, there was a negative correlation between the number of errors made in responding to the interspersed questions and the achievement variables. This correlation was accompanied by a positive correlation between achievement and I.Q., and a negative correlation between I.Q. and the number of errors made in responding to the interspersed questions. Since these correlations appeared as a result of each of the treatments, it was argued that the negative correlation between achievement and errors was more of a function of I.Q. than achievement was a function of the number of errors alone.

Across treatment groups, there was a positive relationship between achievement and the number of errors made in responding to the interspersed questions. As the complexity of the questions to which the students in the treatment groups responded increased, the mean number of errors and the mean achievement also increased. The number of errors made in responding to the interspersed questions and the achievements were a function of the treatments that the groups received. It was hypothesized that the differences in achievement were a result of inspection behaviors which were a function of responding to questions of different complexities. The investigator observed some inspection behaviors in his classroom visits that seemed related to response behaviors that have been shown to facilitate achievement.

As the complexities of questions to which students in the treatment groups responded increased, achievement increased, the amount of time spent working on the instruction materials and questions increased, the number of errors made in responding to the questions increased. The data analysis, knowledge of the experimental design, and personal observations lead the investigator to conclude that by asking more complex questions as a part of the instructional sequence, higher relevant and incidental achievement occurs. This higher achievement is a result of additional experience or practice obtained by students as a function of responding to questions, and may have been due to inspection behaviors that were a function of the questions to which the students responded and the errors made in responding to interspersed questions.

Suggestions for Further Research

This study was designed to ask certain basic questions. It was hoped that the experimental findings, while not answering all questions, would shed some light on the relationship of the complexity of questions to which students respond as a part of the instructional materials to achievement. As is the case with many research studies, though some questions are answered, at least tentatively, many more investigations need to be carried out in order to further refine our knowledge.

There are two areas of investigation that could act as extensions of the present study and further refine the knowledge obtained in the present study. The first area of investigation involves determining the effect of selected inspection behaviors or combinations of inspection behaviors on the facilitation of learning. The inspection behaviors of interest are: (a) Reading only the correct answer after responding to a question; (b) Reading an explanation of the derivation of an answer to a question, and; (c) Rereading portions of the learning materials relevant to the question after responding to a question.

The second area of investigation involves the effect of responding to highly complex questions on the facilitation of learning. More specifically, when questions used to elicit verbalization become highly complex (contain large numbers of factors and levels), is learning facilitated? In the present study, the investigator initially suspected that the most complex questions might not facilitate learning to the same degree as did the questions of the next lower level of complexity. This, however, was not true. Further investigations would attempt to determine whether responding to highly complex questions as a part of the instruction materials facilitates learning, and if they do not, at what level of complexity do questions begin to lose their effectiveness in facilitating learning.

Bibliography

Ausubel, D. P. The psychology of meaningful verbal learning. New York: Grune and Stratton Co., Inc., 1963.

Bivens, L. W. Feedback complexity and self-direction in programmed instruction in plane geometry. In Hershberger, W. A., & Terry, D. F. Delay in self-testing in three types of programmed text, Journal of Educational Psychology, 1965, 56, 22-30.

Bottenberg, R. A., & Ward, J. H. Applied Multiple Linear Regression. Tech. Documentary Report PRL-63-6, March, 1963.

Bruner, J. S. On knowing: Essays for the left hand. Cambridge: Belknap Press, 1966.

Bryon, G. L., & Rigney, J. W. An evaluation of a method for shipboard training in operations knowledge. In Hershberger, W. A., & Terry, D. F. Delay of self-testing in three types of programmed text. Journal of Educational Psychology, 1965, 56, 22-30.

Clyde, D. J., Cramer, E. M., & Sherin, R. J. Multivariate statistical programs. Miami, Florida: Biometric Laboratory, University of Miami, 1966.

Coulson, J. E., and Silberman, H. F. Results of an initial experiment in automated teaching. Teaching machines and programmed learning. Lumsdaine, A. A., and Glaser, R. (Eds.), Washington, D. C.: N.E.A., 1960. Pp. 452-468.

Cummings, E., Allana, L., & Goldstein, S. Effect of overt and covert responding on two kinds of learning tasks. Tech. Report No. 620919, 1962, Center for Programmed Instruction, New York.

(Eds.) Temac programmed learning materials - general science series. Chicago: Encyclopaedia Britannica Educational Corporation, 1964.

(Eds.) FORTRAN for the IBM 1130. Endicott, N.Y.: 1965.

Evans, J. L., Glaser, R., and Homme, L. E. A preliminary investigation of variation in the properties of verbal learning sequences of the "teaching machine" type. In Lumsdaine, A. A., and Glaser, R. (Eds.), Teaching machines and programmed learning: A source book. Washington, D. C.: N.E.A., 1960. Pp. 446-451.

Eigen, L. D., & Margulies, S. Response characteristics as a function of information level. Journal of Programmed Instruction, 1963, 2, 45-54.

Flynn, J. M. Effects of unit size and content upon programmed instruction and prediction on learner success. H.E.W. grant no. OEG-4-8-070072-0016-057. Nova University, 1969.

Bibliography, Cont'd.

Frase, L. T. Learning from prose material: Length of passage, knowledge of results, and position of questions. Journal of Educational Psychology, 1967, 58, 266-272.

Frase, L. T. Effect of question location, pacing, and mode upon retention of prose material. Journal of Educational Psychology, 1968, 59, 244-249.

Fry, E. B., A study of teaching machine response modes. In Lumsdaine, A. A., and Glaser, R. (Eds.), Teaching machines and programmed learning. Washington, D. C., N.E.A., 1960. Pp. 469-474.

Goldbeck, R. A., & Campbell, V. N. The effects of response mode and response difficulty on programmed learning. Journal of Educational Psychology, 1962, 53, 110-118.

Guertin, W. H. Updating the internal criterion of item analysis. Psychological Reports. 1965, 16, 376-378.

Guilford, J. P. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1965.

Hershberger, W. A. Learning via programmed reading. Tech. Report No. AIR-C-28-7/63-TR5, July 1963, American Institute for Research, Palo Alto, California.

Hershberger, W.A.Self-evaluational responding and typographical cueing: Techniques for programming self-instructional reading materials. Journal of Educational Psychology. 1964, 55, 288-296.

Hershberger, W. A., & Terry, D. F. Delay of self-testing in three types of programmed text. Journal of Educational Psychology, 1965, 56, 22-30.

Holland, J. G. Program design and use. In L. Van Atta (Chmn.) Automated teaching methods in undergraduate instruction and educational research, a paper presented at A.P.A. Association meeting, Chicago, September, 1960.

Holland, J. G. Response contingencies in teaching-machine programs. Journal of Programmed Instruction, 1965, 3, 1-8.

Inlow, G. M. Maturity in high school teaching. Englewood Cliffs, New Jersey: Prentice-Hall, 1965.

Jeffrey, W. E. The effects of verbal and non-verbal responses in mediating an instrumental act, Journal of Experimental Psychology, 1963, 45, 327-333.

Kanner, J. H., & Sulzer, R. L. Overt and covert rehearsal of 50% versus 100% of the material in film learning. Air Force Personnel and Training Research Center, Training Aids Research Laboratory, Unpublished Staff Memorandum TARL 56-12, 1956.

Krumboltz, J. D. Factors affecting difficulty level and criterion performance. Programmed Instruction, 1964A, 3, 4-5.

Krumboltz, J. D. The nature and importance of the required response in programmed instruction. American educational Research Journal, 1964B, 4, 203-209.

Krumboltz, J. D., & Bonawitz, B. The effect of receiving the confirmation response in context in programmed material. Journal of Educational Research, 1962, 55, 472-475.

Krumboltz, J. D., & Weisman, R. G. The effect of overt versus covert responding to programmed instruction on immediate and delayed retention. Journal of Educational Psychology, 1962, 53, 89-92.

Kuder, G. F., & Richardson, M. W. The theory of the estimation of test reliability. Psychometrika, 1937, 2, 151-160.

Love, W. A. Personal communication, Nova University, Fort Lauderdale, Florida, 1968.

Michael, D. N., & Maccoby, N. Factors influencing verbal learning from films under varying conditions of audience participation. Journal of Experimental Psychology, 1953, 46, 411-418.

Michael, D. N., & Maccoby, N. Factors influencing the effects of student participation on verbal learning from films: Motivation versus practice effects, "feedback", and overt versus covert responding. In Lumsdaine, A. A. (Ed.). Student response in programmed instruction: A symposium. Washington, D. C.: National Academy of Sciences - National Research Council, Pub. No. 943, 1961.

McGuire, W. J. Studies of factors influencing the effectiveness of demonstrational films for teaching a motor skill: III. Serial position and proximity of reward within the film. Air Force Personnel and Training Research Center, Training Aids Research Laboratory. NOTE TARL-LN-55-51, December, 1955.

McGuire, W. J. Effects of serial position and proximity to "reward" within a demonstration film. In Lumsdaine, A. A. (Ed.). Student response in programmed instruction: A symposium. Washington, D.C.: National Academy of Sciences - National Research Council, Pub. No. 943, 1961.

Myer, S. R. Report on the initial test of a junior high school vocabulary program. In Lumsdaine, A. A., & Glaser, R. (Eds.). Teaching machines and programmed learning. Washington, D. C. : N.E.A. 1960, 229-246.

Bibliography Cont'd.

Norcross, K. J., & Spiker, C. C. The effect of type of stimulus pretraining on discrimination performance in preschool children. Child Development, 1957, 28, 79-84.

Oliver, A. I. Curriculum improvement, a guide to problems, principles, and procedures. New York: Dodd, Mead, and Company, 1965.

Peterson, J. C. The value of guidance in reading for information. In Lumsdaine, A. A., & Glaser, R. (Eds.). Teaching machines and programmed learning. Washington, D. C.: N.E.A. 1960, 52-58.

Richardson, M. W. Notes on the rationale of item analysis. Psychometrika. 1936, 1, 69-76.

Rothkopf, E. Z. Some theoretical and experimental approaches to problems in written instruction. In Krumboltz, J. D. (Ed.). Learning and the Educational Process, Chicago: Rand McNally, 1965, 193-221.

Rothkopf, E. Z. Learning from written instructive material: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 341-150.

Rothkopf, E. Z., & Bisbicos, E. E. Selective facilitative effects of interspersed questions on learning from written materials. Journal of Educational Psychology, 1967, 58, 56-61.

Rothkopf, E. Z., & Coke, E. U. Variations in phrasing, repetition intervals, and the recall of sentence material. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 86-91.

Scheffe, H. The Analysis of Variance. New York: John Wiley & Sons, Inc. 1959.

Silverman, R. E., & Alter, M. Note on the response in teaching machine programs. Psychology Reports, 1960, 7, 496-499.

Skinner, B. F. Teaching machines. Science, 1958, 128, 969-977.

Snedecor, G. W. Statistical Methods applied to experiments in agriculture and biology. Ames, Iowa: The Iowa State University Press, 1962.

Veldman, D. J. FORTRAN programming for the behavioral sciences. New York: Holt, Rinehart, and Winston, Inc., 1967.

Wiles, K. The changing curriculum of the American high school. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

Williams, J. P. Effectiveness of constructed-response and multiple-choice programming modes as a function of test mode. Journal of Educational Psychology, 1965, 56, 111-117.

Wittrock, M. C. Response mode in the programming of kinetic molecular theory concepts. Journal of Educational Psychology, 1963, 54, 89-93.

Appendix A

Question Evaluations By Panel of Experts

A. Qualifications of the experts on the panel.

1.	Dr. A. S. Fischler	Ed. D. in science education from Columbia University. Author of an elementary and a junior high science textbook series.
2.	Dr. Joseph Lipson	Ph. D. in physics from the University of California. Developer of I. P. I. science and math for elementary schools.
3.	Dr. John M. Flynn	Ed. D. in educational psychology from the University of Florida. Project Director, An Analysis of the Role of the Teacher in an Innovative Prototype School.
4.	Edward R. Simco	M. S. in physics from the University of Pittsburgh. Former physics teacher and chairman of a junior high school science department. Candidate for a Ph. D. in science education.
5.	Robert A. Lehman	M. Ed. in science education from the University of Oklahoma. Former junior high school science teacher and candidate for a Ph. D. in science education.

B. Responses of the Experts.

Judges		1	2	3	4	5
		VRC	VRC	VRC	VRC	VRC
1	A	XXX	X	000	XXX	XXX
	B*	XXX	X	000	XXX	XXX
2	A*	XXX	XXX	OXX	XXX	XXX
	B	OXX	OXX	OXX	OXX	OXX
3	A	OXX	XX	XXX	XXX	XXX
	B*	XX	XX	XXX	XXX	XXX
4	A*	OOX	XXX	XXX	XXX	XXX
	B	OXO	X	XXX	XXX	XXX
5	A	XXX	XXX	XXX	XXX	XXX
	B*	XXX	XXX	XXX	XXX	XXX
6	A*	XXX	XXX	XXX	XXX	XXX
	B	XXX	XXX	XXX	XXX	XXX
7	A*	XXX	XX	XXX	XXX	XXX
	B	XXX	XX	XXX	XXX	XXX
8	A	OXX	XXX	X X	XXX	XXX
	B*	XXX	XXO	XXX	XXX	XXX
9	A*	XXX	XXX	XXX	XXX	XXX
	B	XXX	XXO	XXX	XXX	XXX
10	A	XXX	XX	XX	XXX	XXX
	B*	XXX	XXX	XXX	XXX	XXX
11	A*	XXX	XXX	XXX	XXX	XOX
	B	XXX	XXX	XXX	XXX	XOX
12	A*	XXX	XXX	XXX	XXX	XXX
	B	XXX	XXX	XXX	XXX	XXX
13	A*	XXX	XXX	XXX	XXX	XXX
	B	XXX	XXX	XXX	XXX	XXX

V = Content Validity
R = Reading
C = Complexity

X = Acceptable on this criterion.
O = Not acceptable on this criterion.
= Questionable acceptance on this criterion.

* = Sets of questions that were interspersed in the text materials.

Appendix B

Instructions to Judges

Nova University
Social and Behavioral Sciences Center
3301 College Avenue
Fort Lauderdale, Florida

Dear Sir:

I would like to take this opportunity to thank you for your willingness to act as a judge of the questions that I anticipate using as a part of my thesis. This letter is intended to serve as an introduction to the format of the accompanying materials and to set down the criteria on which you will judge the questions.

The accompanying package of materials contains a 250-frame programmed learning sequence that has been divided into sections. The programmed materials are identical to those that will be used by students in the study, and you will note that all of the frames have the answers typed into the response spaces. Following each of the sections you will find two sets of questions with three questions in each set. Eventually, one of each of the two sets of questions will be used as a part of the instructional sequence, and portions of the other set will be used in the final student evaluation. Please evaluate each set of questions on three points or criteria. These points are: (1) Are the words and sentence structure used in the questions and the accompanying answers consistent with the abilities of seventh grade students; (2) Are the concepts involved in answering the questions included in the preceding sections of the programmed instruction sequence, and; (3) Do the three questions in each set vary sequentially in complexity as set down in the following definitions.

The second question in each set is designed to be more complex than the first, and the third question in each set is designed to be more complex than the second. The second and third questions are expected to be increasingly more complex than the first, but the increment or difference in complexity between questions in one set or different sets is not expected to be the same.

A question is defined as being more complex than another when it involves either more factors and/or more levels per factor. A factor is a characteristic of subject content or topic that can be measured or described independently of other characteristics within the subject content or topic. In terms of factors, a question is more complex than another if it involves more factors than another question or other questions, regardless of the relevance of the additional factor or factors to the solution or answer. A level is a subset of a factor. In terms of levels, a question becomes more complex as more subsets are introduced, as more members are introduced within each subset, and/or as the ratios between the numbers of members in two or more subsets increases. Also, a problem in which a comparison is made of two or more factors or levels will be considered more complex than a problem in which factors or levels are considered separately. For example, in subject content dealing with levels, the masses or forces involved, the positioning of the masses relative to the fulcrum, and the angle through which the lever turns are factors. Higher ratios, numbers, or magnitudes of masses and lever arm lengths, and the number of degrees of angles of rotation of the lever are levels within each of the factors. A comparison of the clockwise and counter-clockwise moments (factors) in a lever is considered more complex than the derivation of either set of moments themselves.

There are then three points on which you will evaluate each set of questions. These are: (1) reading level; (2) content validity, and; (3) levels of complexity. A check list is provided on the first question of each set for your convenience. Please mark the acceptableness of each set of questions for each criteria on the check list. If you feel that you would like to make other constructive criticisms, please feel free to write comments at points that you feel are appropriate.

I appreciate your taking time from a busy schedule to perform this task for me. If you are interested, I would be happy to send you a copy of the experimental results when they are completed.

Sincerely,

Michael Yost, Jr.

Appendix C

Programmed Instruction Text and the Interspersed Questions

Name ______________________________
last first

Teacher ______________________________

Section ______________________________

MECHANICS

Please do not place marks in the blocks that appear below. They are for use by the teacher only.

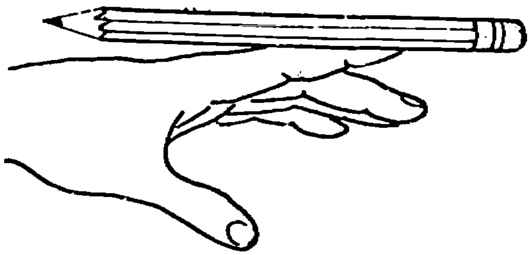
Instructions:

You are being asked to read this booklet and, at times, to answer questions about the subject content. It is important that you pay very close attention to what you read, and that you do not move ahead until you understand what you have already read. This booklet is very much like the one on astronomy that you have recently been working in. You will not be required to fill in the blanks of this booklet as you did in the astronomy booklet. Instead, you will read each sentence with the correct answer typed in the blank. Every 15 or 20 statements you will find a page containing a single question or statement. Carefully read the question or statement. If it is a question, answer it, and draw a circle around your answer. If it is a statement, read it, and continue on. Questions have their answers on the back of the page on which they are found. Do not read the correct answer before writing down your own answer, and do not change your answer after you have read the correct answer.

After reading these instructions you may begin working. Please use the booklet conscientiously and follow the directions at all times. You may work at your own speed, and you should try not to waste any of your time. As soon as you complete the booklet, see your teacher. He will give you two short quizzes to see how much of the science material you can remember and use.

Do you now have any questions about what you will be doing while working with this booklet? If you have any questions, ask them at this time. Your teacher will answer any questions that you may have at this time, but will not answer questions once you begin working.

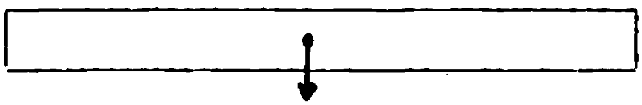
1. You have learned that an object like your pencil will __balance__ at the point of its center of gravity.



Balanced

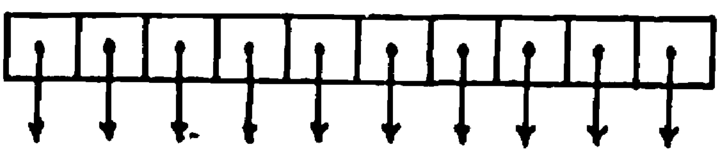
2. All of the weight of an object seems to be concentrated at its __center of gravity__.

C of G



Yard Stick

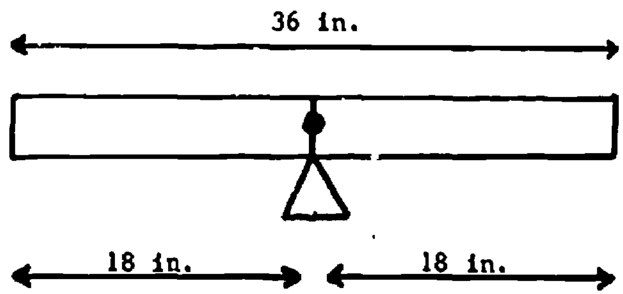
3. Actually, the __weight__ of the yard stick is distributed over its entire length.



Weight Distributed

4. But, the yard-stick acts as if its weight were concentrated at its center.

5. A yard-stick at the 18-inch mark will be exactly balanced.

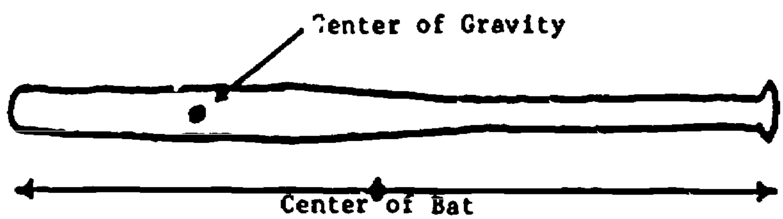


6. The number of units of length on the left will be equal to the same number of units of length on the right .

7. This point where the yard-stick will be exactly balanced is called its center of gravity.

8. If we support the yard-stick at its center of gravity it will be exactly balanced.

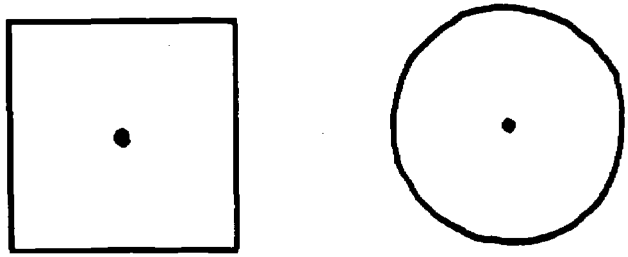
9. The center of gravity of an irregular object like a baseball bat will not be located at its center .



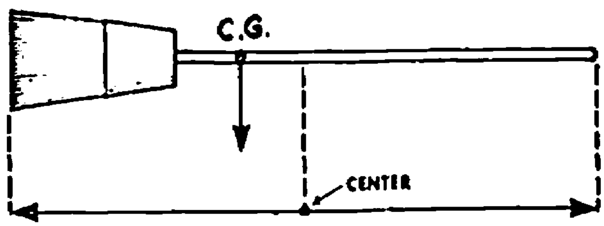
10. The entire weight of the baseball bat seems to be concentrated at the thick end.

11. Generally, every object acts as if its weight were concentrated at some single point.

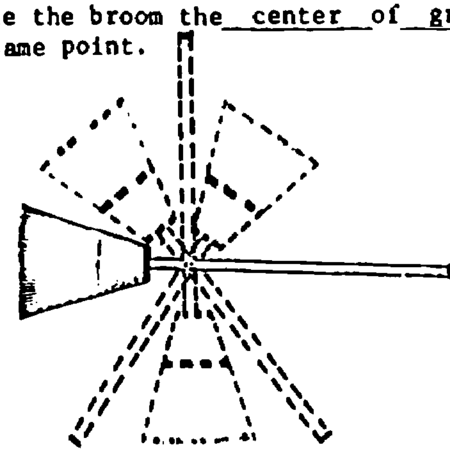
12. A regularly shaped object like a cube or a circular disc will have its center of gravity at the center.



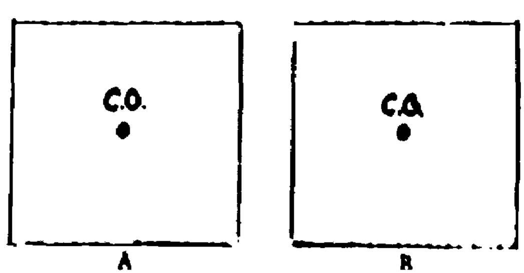
13. An object like a broom would not have its center of gravity at its center.



14. No matter how you rotate the broom the center of gravity will always be at the same point.

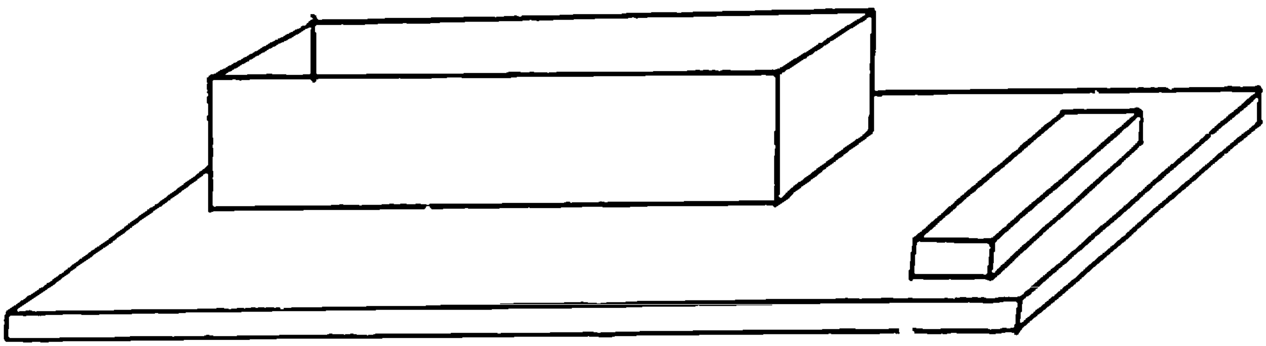


15. The two objects shown below have their centers of gravity at their centers.



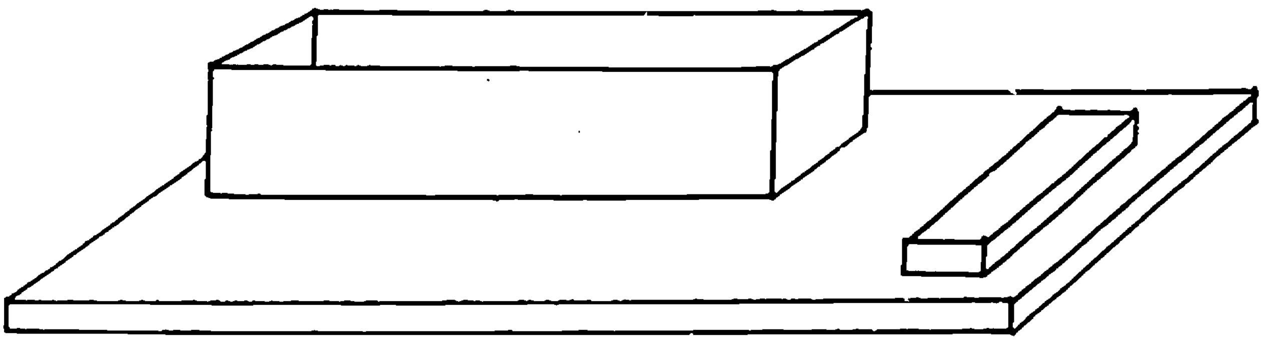
An object can be balanced on another object, or pivot point, if we place it so that there are equal amounts of weight on each side of the point on which it balances.

Suppose that you are given an empty box and a rectangular stick as shown in the drawing below. Where can you place the stick if you are to have the box balance on it and not touch the table?



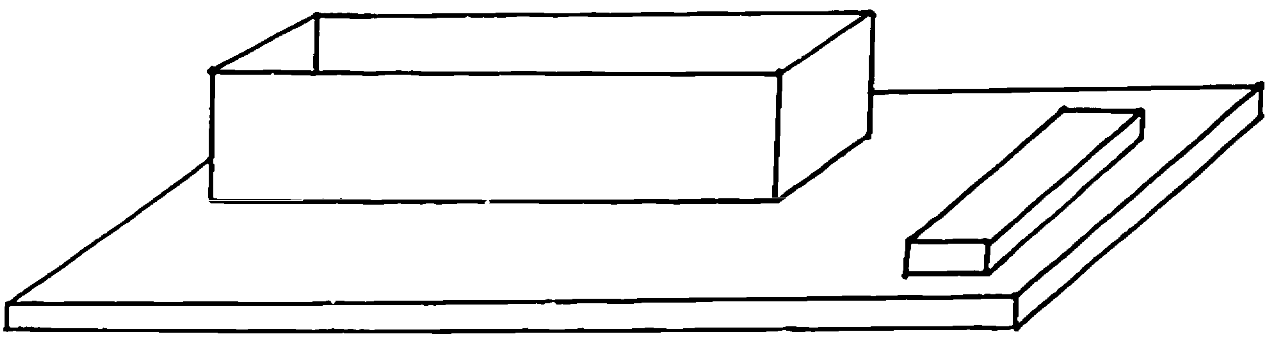
To have the box balance on the stick, place the stick so that there are equal amounts of the box on each side of the stick. That is, place the stick under the center of the box.

Suppose that you are given an empty box and a rectangular stick as shown in the drawing below. You then fill the box with sand so that the sand is even with the top of the box. In what position under the sand-filled box can you place the stick to have the box balance and not touch the table?



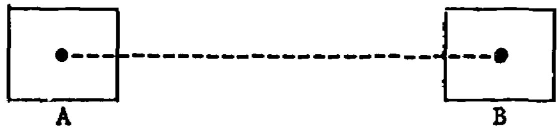
To have the box balance on the stick, place the stick so that there are equal amounts of the box and sand on each side of the stick. That is, place the stick under the center of the box.

Suppose that you are given an empty box and a rectangular stick as shown in the drawing below. The box is then filled with sand so that the sand is even with the top of the box. A small bucket of stones is placed on top of the right side of the sand-filled box. In what position under the sand-filled box and bucket of stones can the stick be placed to have them balance and not touch the table?

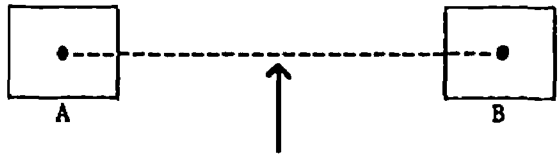


To have the box, stones, and sand balance on the stick, place the stick so that there are equal amounts of weight on each side of the stick. That is, place the stick under the box to the right of the center to allow for the weight of the stones on that side.

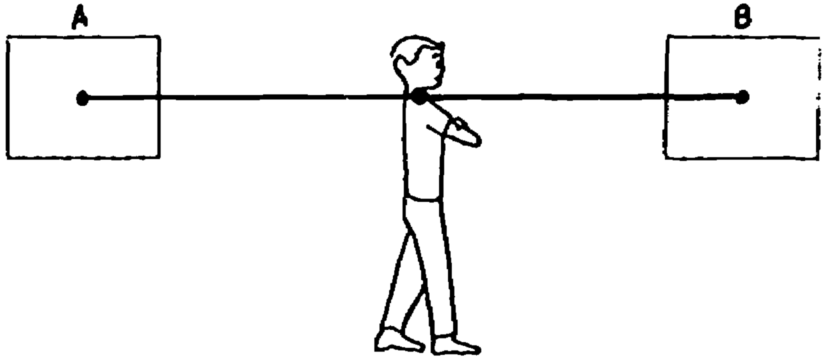
16. The center of gravity of A and B together is found by joining their centers of gravity.



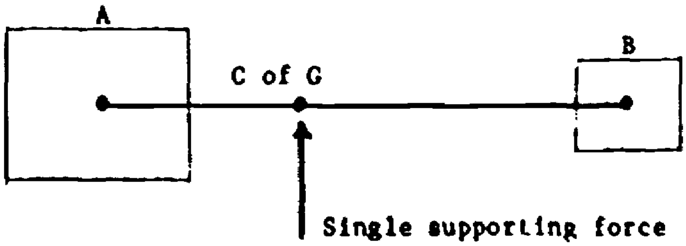
17. We can use a single supporting force at the new center of gravity to balance A and B.



18. A and B (below) are supported at the new center of gravity.

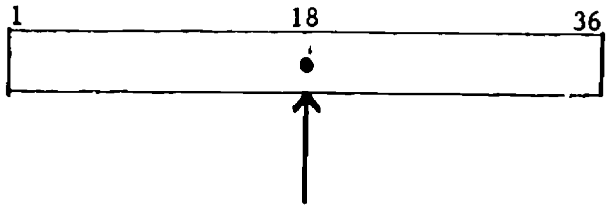


19. If A were heavier than B, their common center of gravity would not be at the center of the connecting rod.

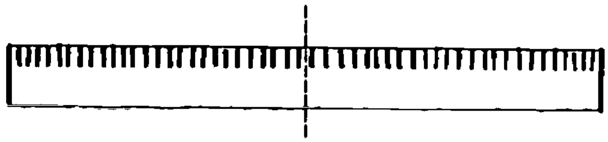


20. We can find a single supporting__force__to balance A and B.

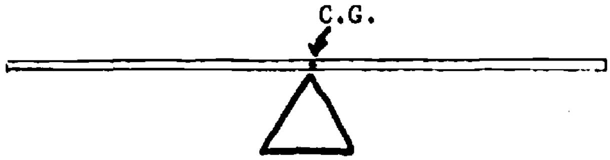
21. A yard-stick will balance at the 18-inch mark, or its__center__ of__gravity__.



22. There are 18 equal units of length on one side and__18__ equal units of length on the other side of a yard-stick.



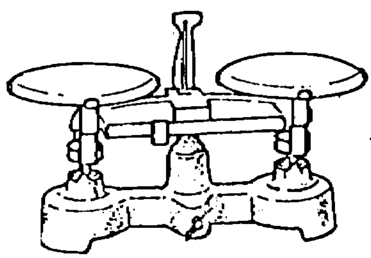
23. A LEVER is like a __yard__-stick, rod, or plank.



Supporting Force

24. A lever will balance when its supporting force is at its __center of gravity__.

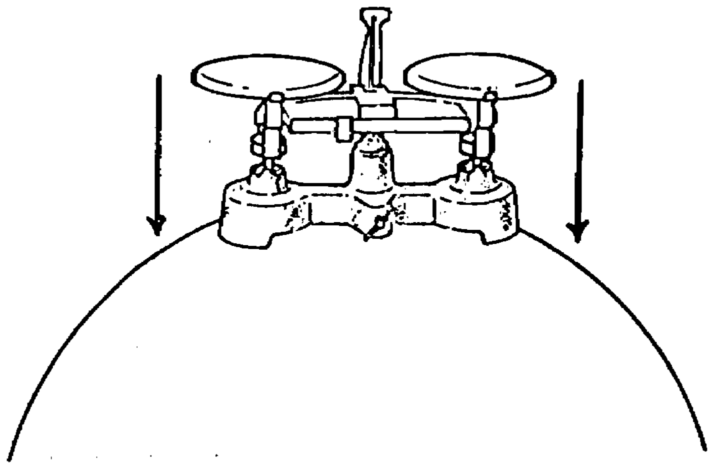
25. Perhaps you have used an equal-arm balance.



Two balanced pans are supported at (equal/unequal) __equal__ distances from a common point of support.

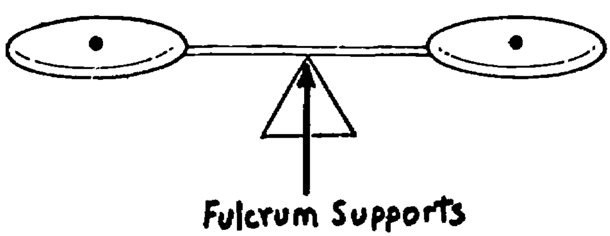
26. The balanced scale pans are also (equal/unequal) __equal__ in mass.

27. Since they are equal in mass they are attracted equally to the earth and hence have equal__weight__.



28. Mass and weight (do/do not)__do not__mean the same thing.

29. The two scale pans are supported at (equal/unequal)__equal__ distances from a common support called the FULCRUM.

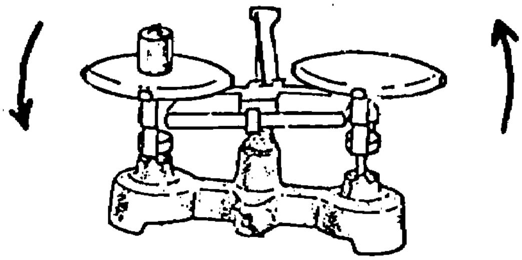


Balanced Scale Pans

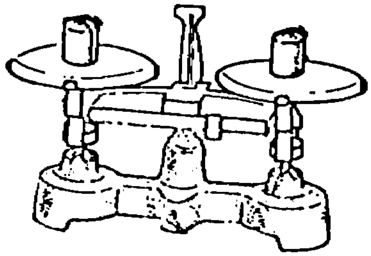
30. The common support from which the two scale pans are free to turn is called the __FULCRUM__.

31. The pans will__balance__at their common center of gravity.

32. If you should add a brass weight to one pan, the pans would not be __balanced__.

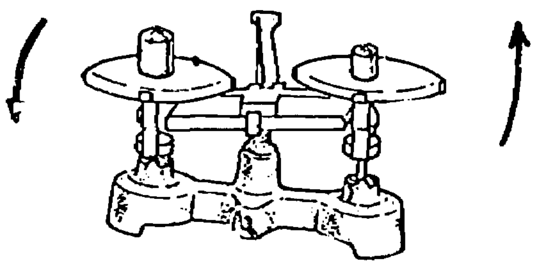


33. To restore the __balance__ we must add an equal weight to the other side.



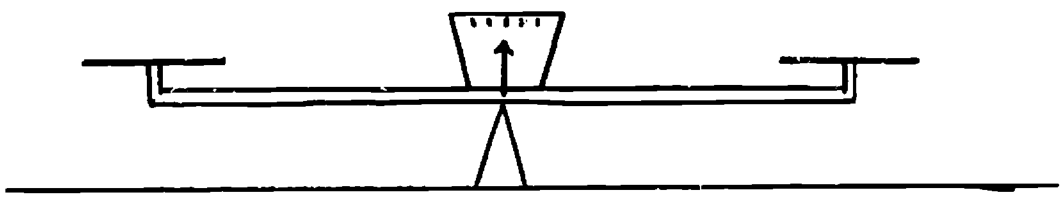
34. The scale pans __balance__ when the weights on both sides are equal.

35. The scale pans move up and down if the weights on either side are (equal/unequal) __unequal__.



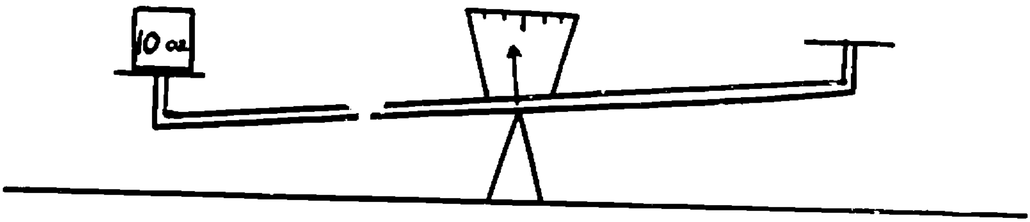
In order to have an equal-arm pan balance balance, the amount of weight on the left pan must be equal to the amount of weight on the right pan.

The picture below is of an equal-arm pan balance. If no weights are placed on either side, the pans would be ________________.



In order to have an equal-arm pan balance balance, the amount of weight on the left pan must be eaual to the amount of weight on the right pan. In this problem, the pans would be balanced.

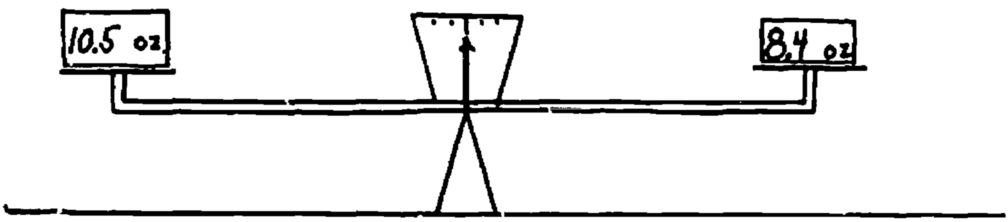
The picture below is of an equal-arm pan balance. If a 10-ounce weight is placed on the pan on one side, how much weight must be placed on the pan on the opposite side in order to have the pans balance?



In order to have an equal-arm pan balance balance, the amount of weight on the left pan must be equal to the amount of weight on the right pan.

In this problem, a 10-ounce weight must be placed on the opposite side in order to have the pans balance.

The picture below is of an equal-arm pan balance. A 10.5-ounce weight has been placed on the left pan and an 8.4-ounce weight has been placed on the right pan. <u>How much</u> weight must be placed on <u>which</u> pan in order to have the pans balance?

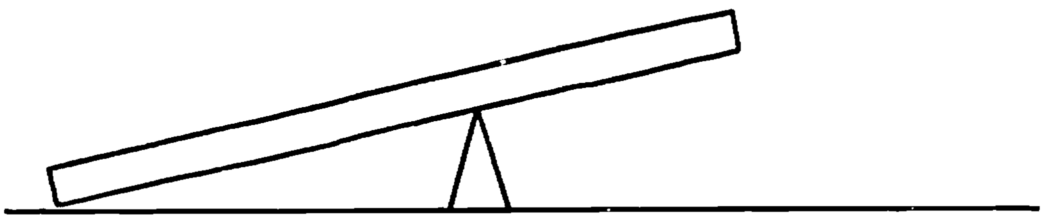


In order to have an equal-arm pan balance balance, the amount of weight on the left pan must be equal to the amount of weight on the right pan.

In this problem, an extra weight of 2.1 ounces must be placed on the <u>right</u> pan in order to have the pans balance.

A bar or see-saw will be balanced when there are weights of the same size attached at equal distances on each side of the fulcrum. If one weight is larger than the other, then the larger of the two weights must be placed closer to the fulcrum in order to have the bar or see-saw balance.

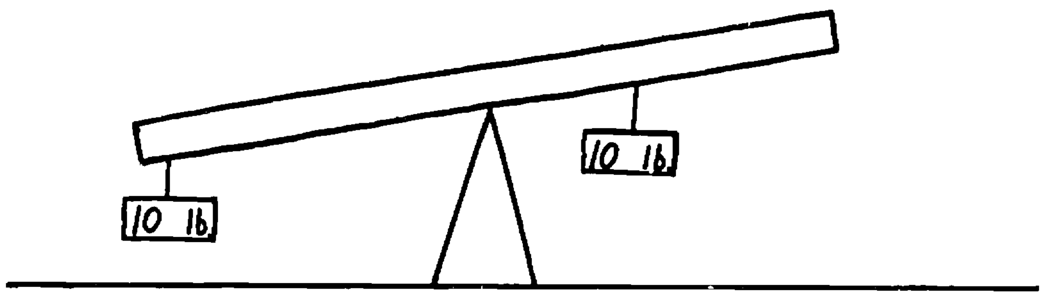
The bar pictured below is on a stand that will allow it to pivot or turn. The bar is of the same thickness all along its length. In the space below the picture, tell why the bar is not balanced.



A bar or see-saw will be balanced when there are weights of the same size attached at equal distances on each side of the pivot. If one weight is larger than the other, then the larger of the two weights must be placed closer to the pivot in order to have the bar or see-saw balance.

The bar is not balanced because there is more of the bar to the left of the pivot than to the right of the pivot or because the amount of bar to the left of the pivot weighs more than the amount of bar to the right of the pivot.

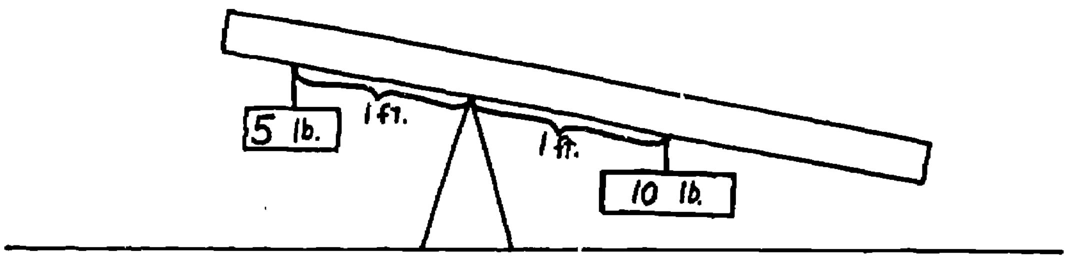
The bar pictured below is on a stand that will allow it to pivot (go up or down). The bar is of the same thickness all along its length and was balanced before the weights were attached. In the space below the picture, tell why the bar is <u>not</u> balanced.



A bar or see-saw will be balanced when there are weights of the same size attached at equal distances on each side of the pivot. If one weight is larger than the other, then the larger of the two weights must be placed closer to the pivot in order to have the bar or see-saw balance.

In this problem, the bar is not balanced because the equal-sized weights are attached at <u>different</u> distances from the point about which the lever pivots.

The bar pictured below is on a stand that will allow it to pivot or turn. The bar is of the same thickness all along its length. In the space below the picture, tell why the bar is _not_ balanced.

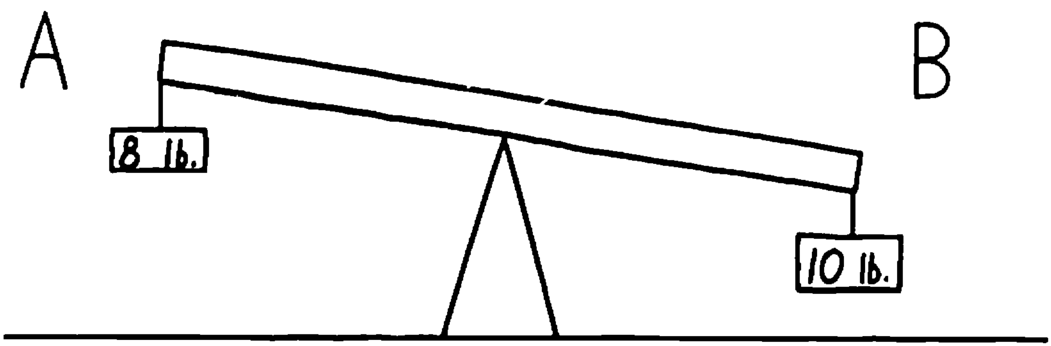


A bar or see-saw will be balanced when there are weights of the same size attached at equal distances on each side of the pivot. If one weight is larger than the other, then the larger of the two weights must be placed closer to the pivot in order to have the bar or see-saw balance.

There are two reasons why this bar is not balanced. One reason is because there is more bar to the right of the pivot than to the left of the pivot. The other reason is because the weight on the right is larger than the weight on the left even though they have the same lever arm length.

When objects of unequal weight are placed [illegible]qual distances on each side of the pivot of a lever, it wil[illegible] e balanced. The end of the lever on which the heavier object is hung will go down and the opposite end will go up. Each object pulls down with all of its weight. The difference in the weights of the two objects will be effective in causing the heavier end of the lever to go down and the lighter end to go up.

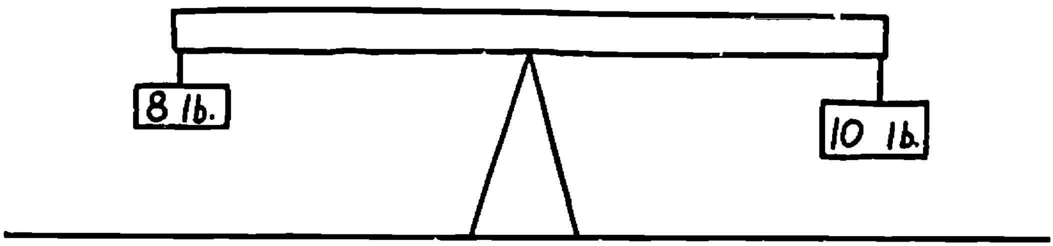
The drawing below is of a lever with arms of equal length. How much of a pull downward does the weight on end A exert?



When objects of unequal weight are placed at equal distances on each side of the pivot of a lever, it will not be balanced. The end of the lever on which the heavier object is hung will go down and the opposite end will go up. Each object pulls down with all of its weight. The difference in the weights of the two objects will be effective in causing the heavier end of the lever to go down and the lighter end to go up.

In this problem, the weight on end A exerts a downward pull of 8 pounds.

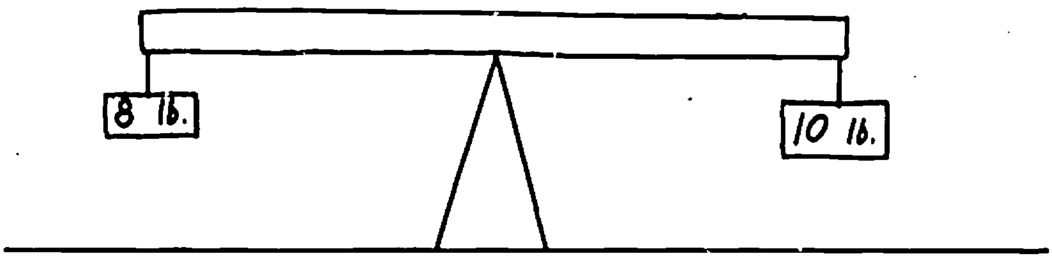
The drawing below is of a lever with arms of equal length. Place arrows on the drawing to show which end will go down and which end will go up.



When objects of <u>unequal</u> weight are placed at <u>equal</u> distances on each side of the pivot of a lever, it will not be balanced. The end of the lever on which the heavier object is hung will go down and the opposite end will go up. Each object pulls down with all of its weight. The difference in the weights of the two objects will be effective in causing the heavier end of the lever to go down and the lighter end to go up.

In this problem, the end of the lever to which the 10-pound weight is attached will go down and the end to which the 8-pound weight is attached will go up.

The drawing below is of a lever with arms of equal length. How much of the weight of the two objects is effective in causing the lever to move? Which end of the lever will go up and which end will go down?

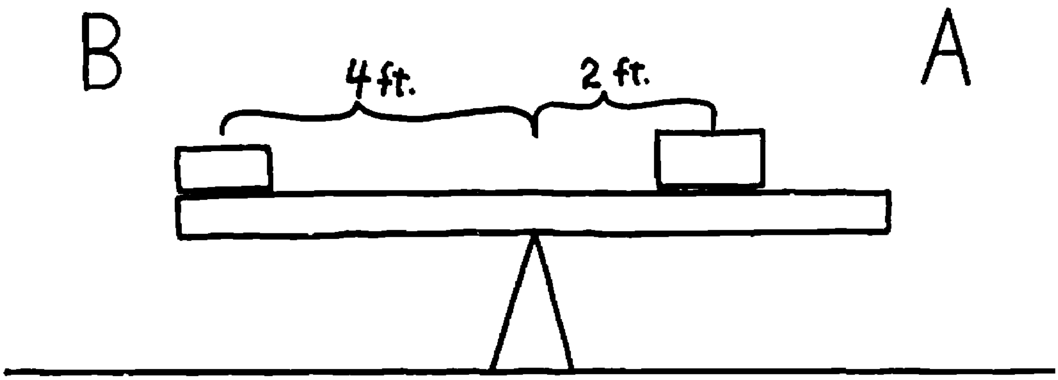


When objects of unequal weight are placed at equal distances on each side of the pivot of a lever, it will not be balanced. The end of the lever on which the heavier object is hung will go down and the opposite end will go up. Each object pulls down with all of its weight. The difference in the weights of the two objects will be effective in causing the heavier end of the lever to go down and the lighter end to go up.

In this problem, the difference between the two weights, or 2 pounds (10-8=2), acts to cause the right end to go down and the left end to go up.

To find or calculate the moment of a lever, a person must know both the weight acting on the lever arm and the length of the lever arm on which the weight acts. By multiplying these two numbers together, we get the moment of that end of the lever.

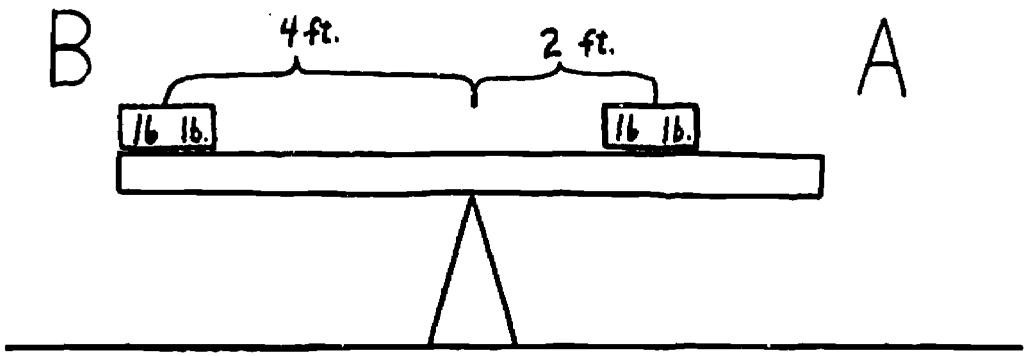
In the drawing below, which length would be used in calculating the moment for the weight on end A of the lever?



To find or calculate the moment of a lever, a person must know both the weight acting on the lever arm and the length of the lever arm on which the weight acts. By multiplying these two numbers together, we get the moment of that end of the lever.

In this problem, 2 feet would be used as the lever-arm length in calculating the moment for end A of the lever.

Find (calculate) the moment for side B of the lever shown in the drawing below.

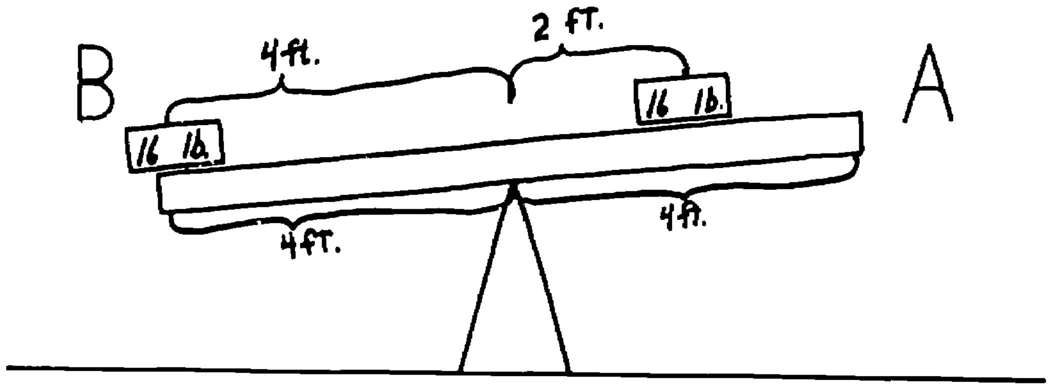


To find or calculate the moment of a lever, a person must know both the weight acting on the lever arm and the length of the lever arm on which the weight acts. By multiplying these two numbers together, we get the moment of that end of the lever.

In this problem:

Moment = weight x lever arm
Moment = 16 pounds x 4 feet
Moment = 64 pound-feet

Find (calculate) the difference in the moments of sides A and B of the lever shown in the drawing below.



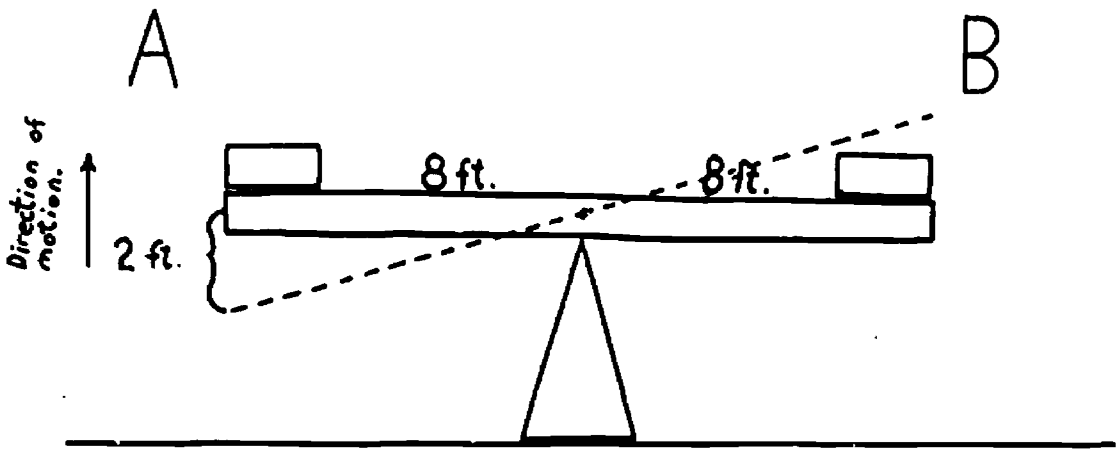
To find or calculate the moment of a lever, a person must know both the weight acting on the lever arm and the length of the lever arm on which the weight acts. By multiplying these two numbers together, we get the moment of that end of the lever.

Side B	Side A
Moment = weight x lever arm	Moment = weight x lever arm
Moment = 16 pounds x 4 feet	Moment = 16 pounds x 2 feet
Moment = 64 pound-feet	Moment = 32 pound-feet

Difference = 64 pound-feet - 32 pound-feet = 32 pound-feet.

To calculate the work done by a lever in moving an object, we multiply the object's <u>weight</u> times the <u>vertical</u> <u>distance</u> that it is moved. To calculate the moment of the same object, we multiply the object's weight times the length of its lever arm.

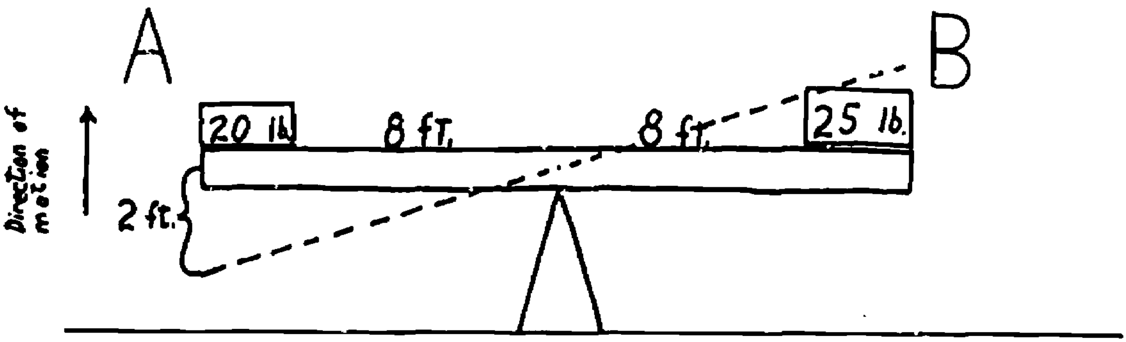
In calculating the moment for end A of the lever pictured below, which length or distance would be used?



To calculate the work done by a lever in moving an object, we multiply the object's <u>weight</u> times the <u>vertical distance</u> that it is moved. To calculate the moment of the same object, we multiply the object's weight times the length of its lever arm.

In this problem, we would use 8 feet as the lever arm length in calculating the moment of end A of the lever.

Find the amount of work done in lifting the weight that is on end A of the lever.

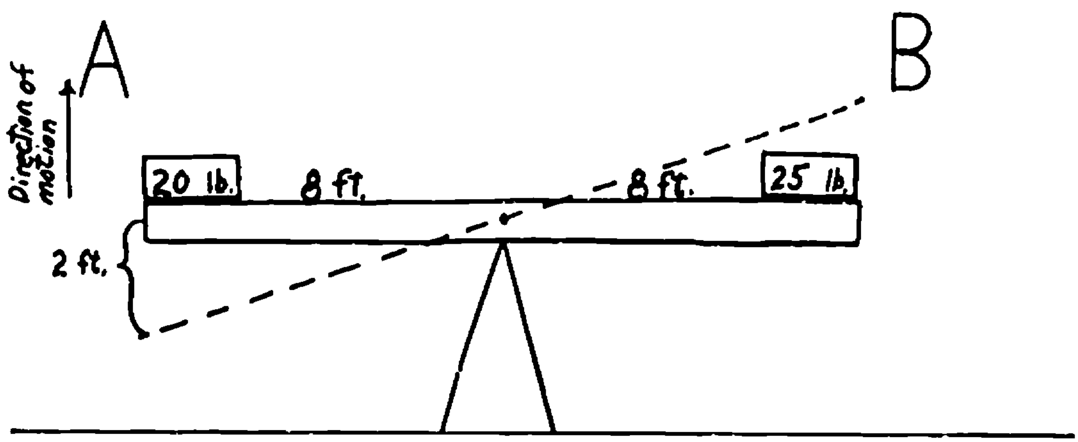


To calculate the work done by a lever in moving an object, we multiply the object's weight times the vertical distance that it is moved. To calculate the moment of the same object, we multiply the object's weight times the length of its lever arm.

In this problem, we would calculate the amount of work by:

Work = weight x distance moved
Work = 20 pounds x 2 feet
Work = 40 foot-pounds

For Side A of the lever, work is done in lifting the object. The object pushes down on the lever causing a moment. Find (calculate) the difference between the size of the moment and the amount of work done in lifting the object.



To calculate the work done by a lever in moving an object, we multiply the object's <u>weight</u> times the <u>vertical distance</u> that it is moved. To calculate the moment of the same object, we multiply the object's weight times the length of its lever arm.

In this problem, we can calculate the work done and the moments of end A of the lever by:

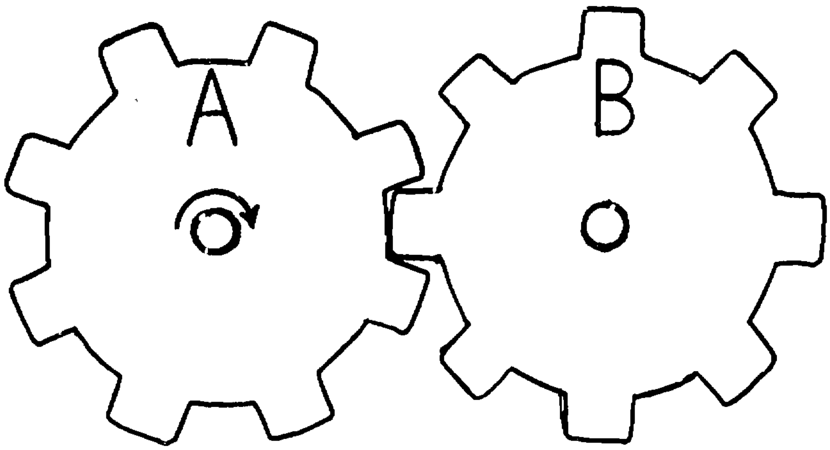
Moment = weight x lever arm
Moment = 20 pounds x 8 feet
Moment = 160 pound-feet

Work = weight x distance moved
Work = 20 pounds x 2 feet
Work = 40 foot-pounds

Moment - Work = 160 - 40 = 120
The moment has the greater numerical value.

We can arrange a number of gears or wheels in a line so that when we turn the first one, all the others in the line will also turn. If we turn the first gear in a clockwise direction, the second gear will turn in a counter-clockwise direction, and the third gear will turn in a clockwise direction. All of the even-numbered gears will turn in one direction and the odd-numbered gears will turn in the opposite direction.

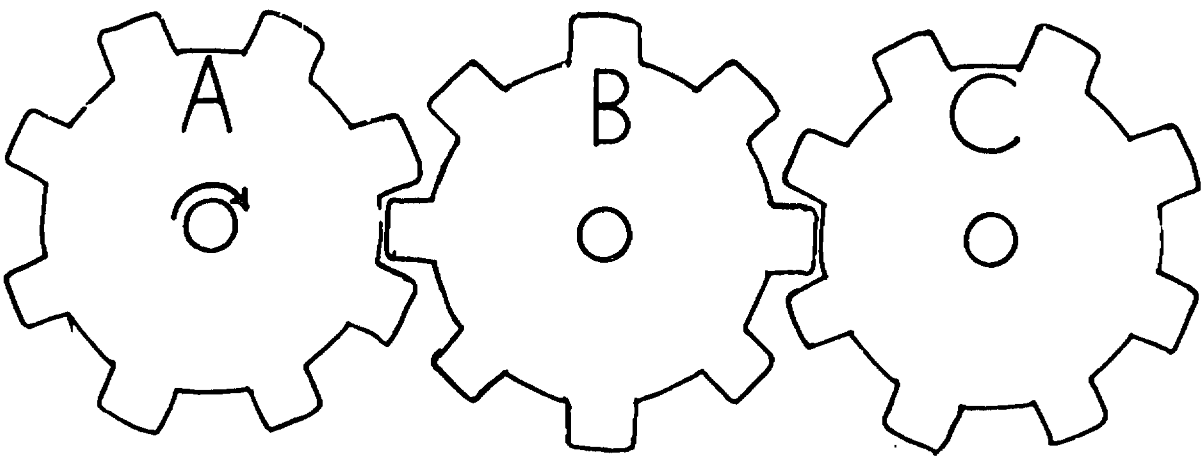
The drawing below is of two gears that are free to rotate. When one gear is rotated, the other will also rotate. If gear A is rotated in a clockwise direction, as the arrow shows, in what direction will gear B rotate?



We can arrange a number of gears or wheels in a line so that when we turn the first one, all the others in the line will also turn. If we turn the first gear in a clockwise direction, the second gear will turn in a counter-clockwise direction, and the third gear will turn in a clockwise direction. All of the even-numbered gears will turn in one direction and the odd-numbered gears will turn in the opposite direction.

In this problem, gear B will rotate in a counterclockwise direction.

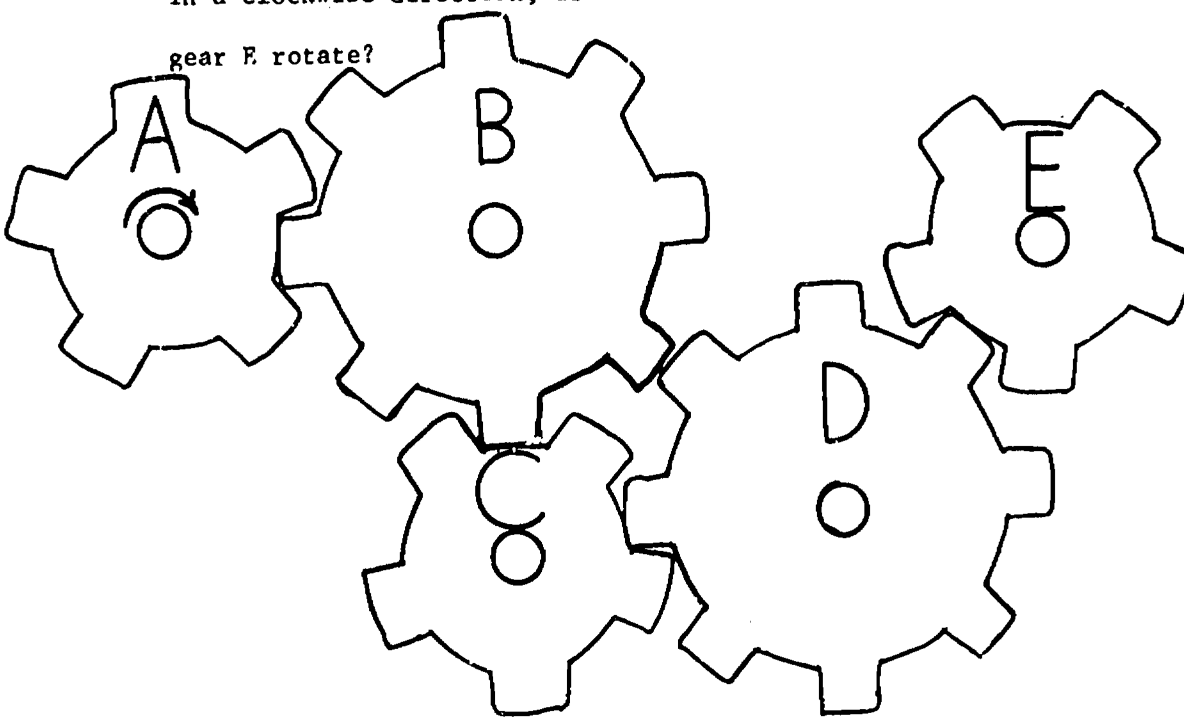
The drawing below is of three gears that are free to rotate. When one gear is rotated, the others will also rotate. If gear A is rotated in a clockwise direction, as the arrow shows, in what direction will gear C rotate?



We can arrange a number of gears or wheels in a line so that when we turn the first one, all the others in the line will also turn. If we turn the first gear in a clockwise direction, the second gear will turn in a counter-clockwise direction, and the third gear will turn in a clockwise direction. All of the even-numbered gears will turn in one direction and the odd-numbered gears will turn in the opposite direction.

In this problem, gear C will turn in a clockwise direction.

The drawing below is of five gears that are free to rotate. When one gear is rotated, the others will also rotate. If gear A is rotated in a clockwise direction, as the arrow shows, in what direction will gear E rotate?

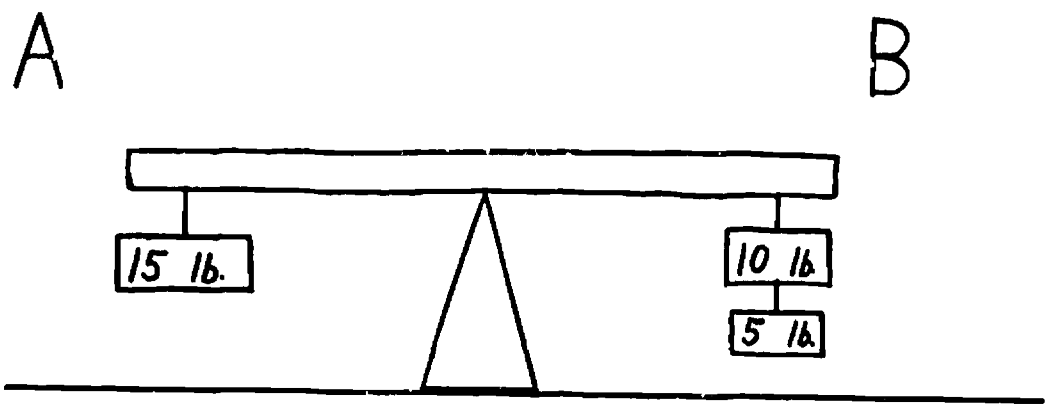


We can arrange a number of gears or wheels in a line so that when we turn the first one, all the others in the line will also turn. If we turn the first gear in a clockwise direction, the second gear will turn in a counter-clockwise direction, and the third gear will turn in a clockwise direction. All of the even-numbered gears will turn in one direction and the odd-numbered gears will turn in the opposite direction.

In this problem, gear E will rotate in a clockwise direction.

In calculating the moment for a side of a lever, we must take into account all of the weights acting on that side. If two weights are hanging from or acting on the same point on one side of a lever and we wish to calculate the moment, we first find the sum of the weights (add them together). We then multiply this number by the length of the lever arm on which the weights are acting.

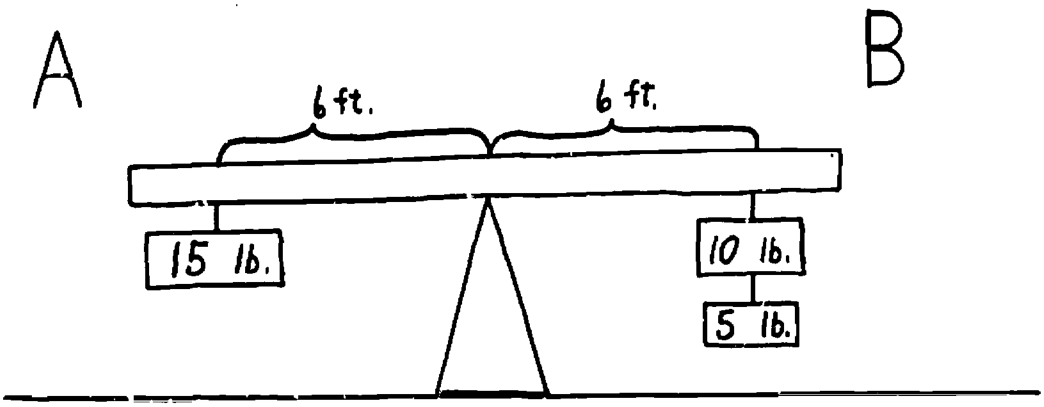
What is the total weight pulling down on side B of the lever shown below?



In calculating the moment for a side of a lever, we must take into account all of the weights acting on that side. If two weights are hanging from or acting on the same point on one side of a lever and we wish to calculate the moment, we first find the sum of the weights (add them together). We then multiply this number by the length of the lever arm on which the weights are acting.

In this problem, the total weight pulling down on Side B is 15 pounds (10 + 5 = 15).

Find (calculate) the moment for side B of the lever shown in the drawing below.

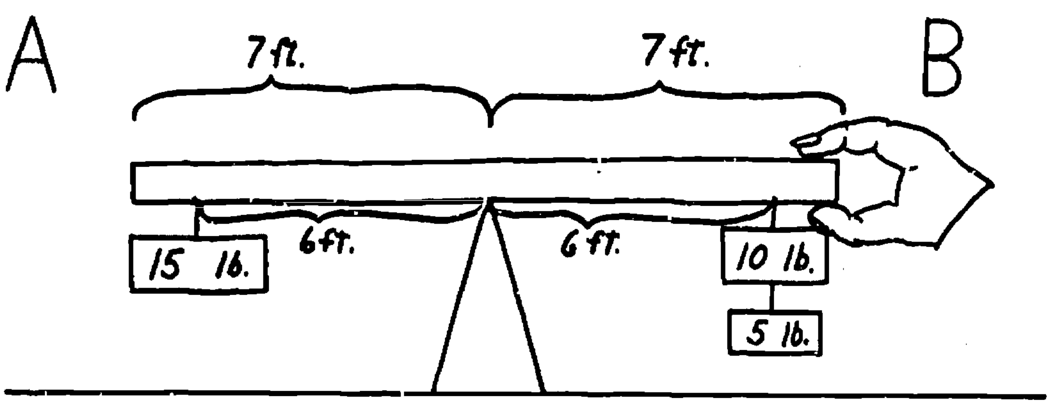


In calculating the moment for a side of a lever, we must take into account all of the weights acting on that side. If two weights are hanging from or acting on the same point on one side of a lever and we wish to calculate the moment, we first find the sum of the weights (add them together). We then multiply this number by the length of the lever arm on which the weights are acting.

In this problem, the moment for side B is:

Moment = weight x lever arm
Moment = (10 + 5) pounds x 6 feet
Moment = 15 pounds x 6 feet
Moment = 90 pound-feet

When the hand releases the lever pictured below, will it balance? If not, which end will go up and which end will go down?



In calculating the moment for a side of a lever, we must take into account all of the weights acting on that side. If two weights are hanging from or acting on the same point on one side of a lever and we wish to calculate the moment, we first find the sum of the weights (add them together). We then multiply this number by the length of the lever arm on which the weights are acting.

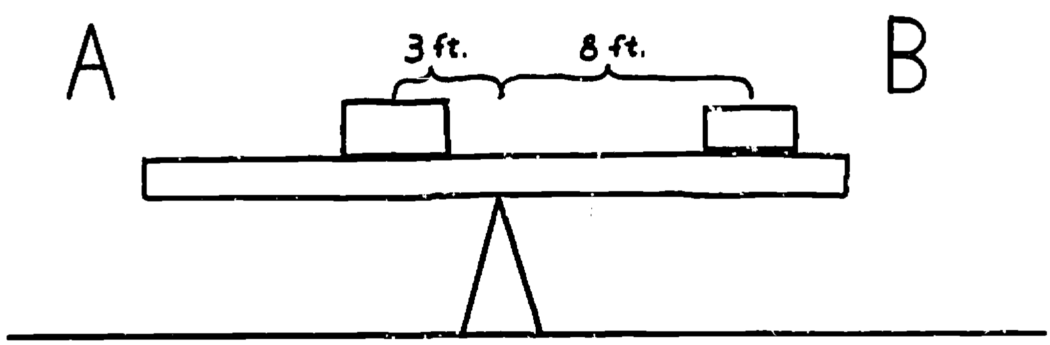
In this problem, the moments for side A and B are:

Side A:	Moment = weight x lever arm	Side B:	Moment = weight x lever arm
	Moment = 15 pounds x 6 feet		Moment = (10+5) pounds x 6 ft
	Moment = 90 pound-feet		Moment = 15 pounds x 6 feet
			Moment = 90 pound-feet

Since there is no difference in the moments of the two ends (90-90 = 0), the lever will balance and will not rotate.

To calculate the moments of a lever, we must know the weight of each object involved and the length of the lever arm on which each of the objects acts. By multiplying each object's weight by its lever arm length, we obtain the moments for the two sides of the lever. If the clockwise and counterclockwise moments of a lever are not equal, a lever will rotate. A lever will rotate in the direction of the larger moment.

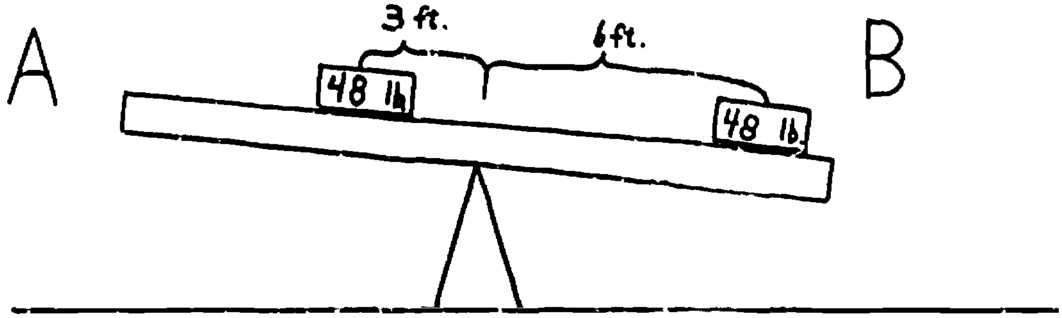
On the lever pictured below, what is the difference in the lengths of the lever arms on which the two objects push down?



To calculate the moments of a lever, we must know the weight of each object involved and the length of the lever arm on which each of the objects acts. By multiplying each object's weight by its lever arm length, we obtain the moments for the two sides of the lever. If the clockwise and counterclockwise moments of a lever are <u>not equal</u>, a lever will rotate. A lever will rotate in the direction of the larger moment.

In this problem, the difference in the lengths of the lever arms is (8 - 3 = 5) or 5 feet.

Find (calculate) the moment for end A of the lever pictured below.

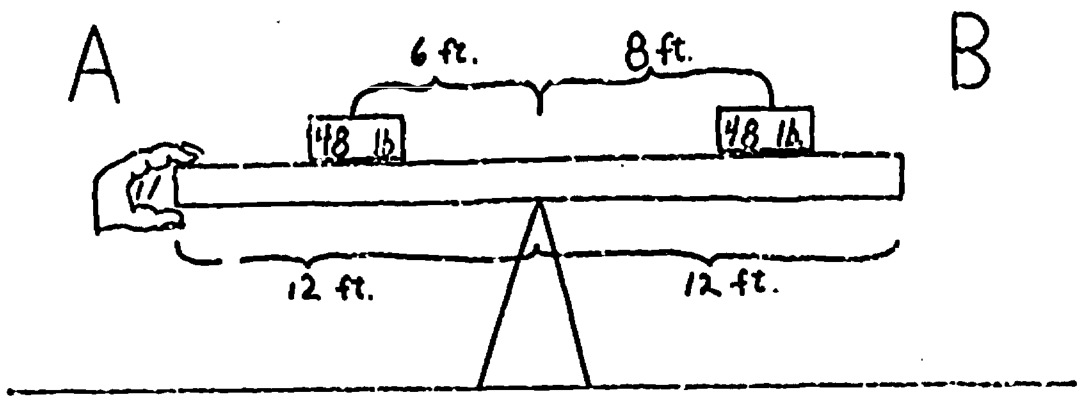


To calculate the moments of a lever, we must know the weight of each object involved and the length of the lever arm on which each of the objects acts. By multiplying each object's weight by its lever arm length, we obtain the moments for the two sides of the lever. If the clockwise and counterclockwise moments of a lever are <u>not equal</u>, a lever will rotate. A lever will rotate in the direction of the larger moment.

In this problem, the moment for side A of the lever is:

Moment = weight x lever arm
Moment = 48 pounds x 3 feet
Moment = 144 pound-feet

When the hand releases the lever pictured below, will it balance or rotate? If it rotates, in which direction will it rotate?



To calculate the moments of a lever, we must know the weight of each object involved and the length of the lever arm on which each of the objects acts. By multiplying each object's weight by its lever arm length, we obtain the moments for the two sides of the lever. If the clockwise and counterclockwise moments of a lever are not equal, a lever will rotate. A lever will rotate in the direction of the larger moment.

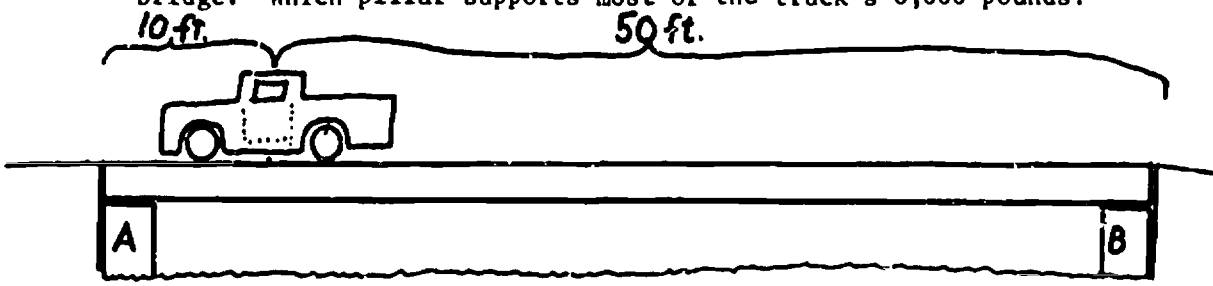
In this problem, the moments for the lever are:

Side A: Moment = weight x lever arm	Side B: Moment = weight x lever arm
Moment = 48 pounds x 6 feet	Moment = 48 pounds x 8 feet
Moment = 288 pound-feet	Moment = 384 pound-feet

Side B has a greater downward moment than side A. Side B will go down and the lever will rotate in a clockwise direction.

When an object is placed on a lever that is held up at its ends, the support on the end that is closer to the object holds up more of its weight than the support that is a greater distance from the object. To calculate the amount of weight that the support at each end holds up, we find how many times closer the object is to one end of the lever than to the other, and state these numbers as a ratio. The weights held by the two supports is in the same ratio as the ratio of distances. The larger weight of the ratio pushes down on the support that is closer to the object and the smaller weight of the ratio pushes down on the support that is at a greater distance from the object. To calculate the moment for one end of this type of lever, we multiply the amount of weight held up by the support on that end by the distance between the object and the support on that end.

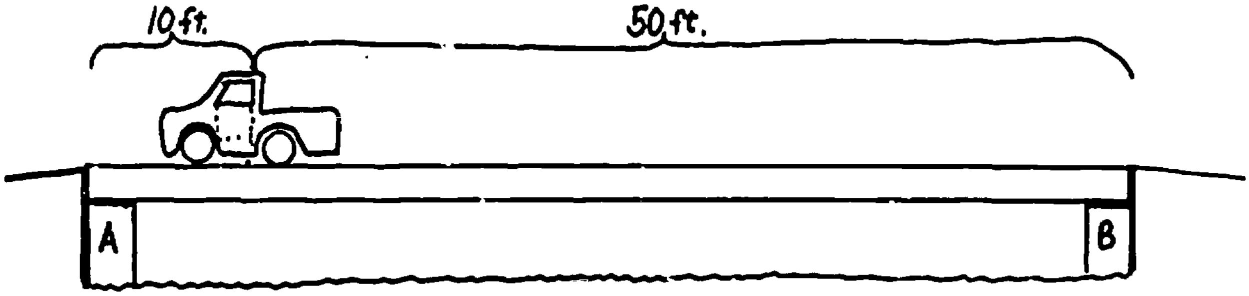
A bridge is shown in the drawing below. The bridge is supported by two pillars (A and B). A truck is stopped on the bridge. Which pillar supports most of the truck's 6,000 pounds?



When an object is placed on a lever that is held up at its ends, the support on the end that is closer to the object holds up more of its weight than the support that is a greater distance from the object. To calculate the amount of weight that the support at each end holds up, we find how many times closer the object is to one end of the lever than to the other, and state these numbers as a ratio. The weights held by the two supports is in the same ratio as the ratio of distances. The larger weight of the ratio pushes down on the support that is closer to the object and the smaller weight of the ratio pushes down on the support that is at a greater distance from the object. To calculate the moment for one end of this type of lever, we multiply the amount of weight held up by the support on that end by the distance between the object and the support on that end.

In this problem, pillar A supports more of the truck's weight than pillar B.

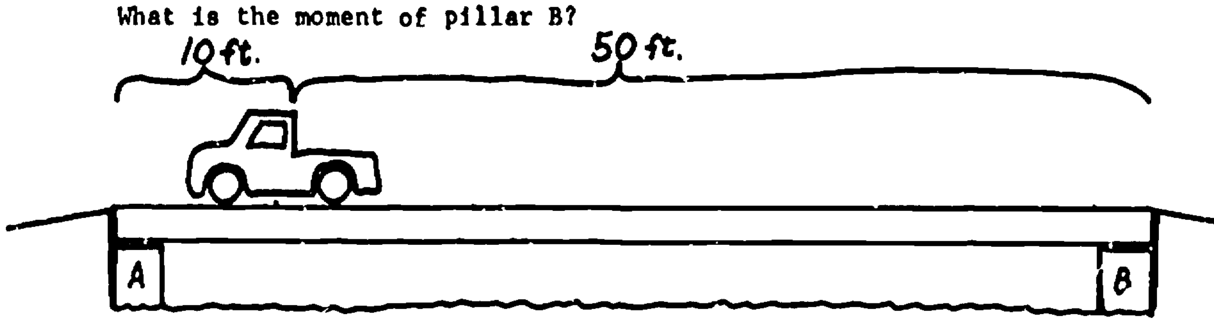
A bridge is shown in the drawing below. The bridge is supported by two pillars (A and B). A truck is stopped on the bridge. How much of the truck's 6,000 pounds is supported by pillar A?



When an object is placed on a lever that is held up at its ends, the support on the end that is closer to the object holds up more of its weight than the support that is a greater distance from the object. To calculate the amount of weight that the support at each end holds up, we find how many times closer the object is to one end of the lever than to the other, and state these numbers as a ratio. The weights held by the two supports is in the same ratio as the ratio of distances. The larger weight of the ratio pushes down on the support that is closer to the object and the smaller weight of the ratio pushes down on the support that is at a greater distance from the object. To calculate the moment for one end of this type of lever, we multiply the amount of weight held up by the support on that end by the distance between the object and the support on that end.

In this problem, the truck is 5 times closer to A than B, and A supports 5 times as much of the truck's weight as B, or 5,000 pounds.

A bridge is shown in the drawing below. The bridge is supported by two pillars (A and B). A 6000-pound truck is stopped on the bridge. What is the moment of pillar B?

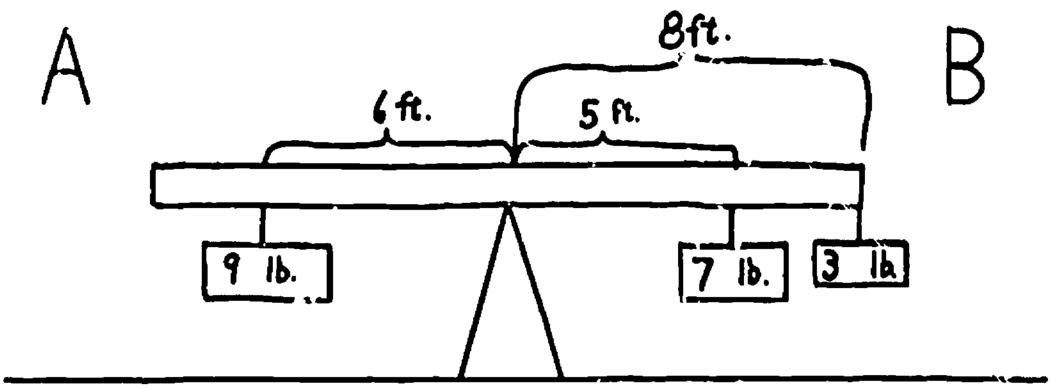


When an object is placed on a lever that is held up at its ends, the support on the end that is closer to the object holds up more of its weight than the support that is a greater distance from the object. To calculate the amount of weight that the support at each end holds up, we find how many times closer the object is to one end of the lever than to the other, and state these numbers as a ratio. The weights held by the two supports is in the same ratio as the ratio of distances. The larger weight of the ratio pushes down on the support that is closer to the object and the smaller weight of the ratio pushes down on the support that is a greater distance from the object. To calculate the moment for one end of this type of lever, we multiply the amount of weight held up by the support on that end by the distance between the object and the support on that end.

In this problem, pillar B is 50 feet from the truck and pillar A is 10 feet from the truck. Pillar B holds up 10/60 or 1/6 of the truck's weight (or 1,000 pounds). The moment of pillar B is equal to the amount of weight it holds up times its lever arm length (1000 pounds x 50 feet) or 50,000 pound-feet.

To calculate the moment of a lever arm on which one object is acting, we multiply the weight of the object times the lever arm length. To calculate the moment of a lever arm on which two objects are acting, we calculate the moment for each object separately and add the two moments together. That is, we multiply each object's weight by its distance from the fulcrum and add the two moments together. To find which end of a lever will go up or down, we calculate the moments for each of the ends. The end with the <u>larger</u> moment will go down.

Find (calculate) the downward moment for side A of the lever shown below.

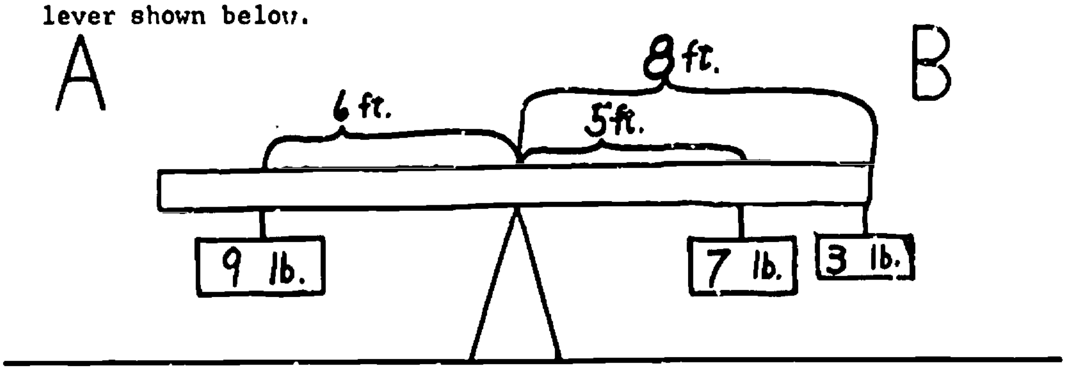


To calculate the moment of a lever arm on which one object is acting, we multiply the weight of the object times the lever arm length. To calculate the moment of a lever arm on which two objects are acting we calculate the moment for each object separately and add the two moments together. That is, we multiply each object's weight by its distance from the fulcrum and add the two moments together. To find which end of a lever will go up or down, we calculate the moments for each of the ends. The end with the <u>larger</u> moment will go down.

In this problem:

Moment = weight x lever arm
Moment = 9 pounds x 6 feet
Moment = 54 pound-feet

Find (calculate) the total downward moment for side B of the lever shown below.

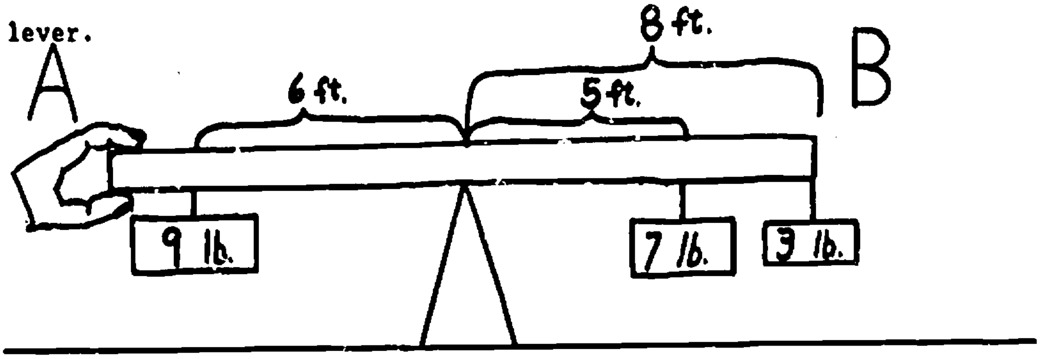


To calculate the moment of a lever arm on which one object is acting, we multiply the weight of the object times the lever arm length. To calculate the moment of a lever arm on which two objects are acting we calculate the moment for each object separately and add the two moments together. That is, we multiply each object's weight by its distance from the fulcrum and add the two moments together. To find which end of a lever will go up or down, we calculate the moments for each of the ends. The end with the <u>larger</u> moment will go down.

In this problem:

Moment = weight x lever arm + weight x lever arm
Moment = 7 pounds x 5 feet + 3 pounds x 8 feet
Moment = 35 pound-feet + 24 pound-feet
Moment = 59 pound-feet

The lever shown below is unbalanced. Perform the work (calculations) to show which end goes up and which end goes down when the hand releases the lever.



To calculate the moment of a levor arm on which one object is acting, we multiply the weight of the object times the lever arm length. To calculate the moment of a lever arm on which tw objects are acting we calculate the moment for each object separately and add the two moments together. That is, we multiply each object's eight by its distance from the fulcrum and add the two moments together. To find which end of a lever will go up or down, we calculate the moments for each of the ends. The end with the larger moment will go down.

In this problem:

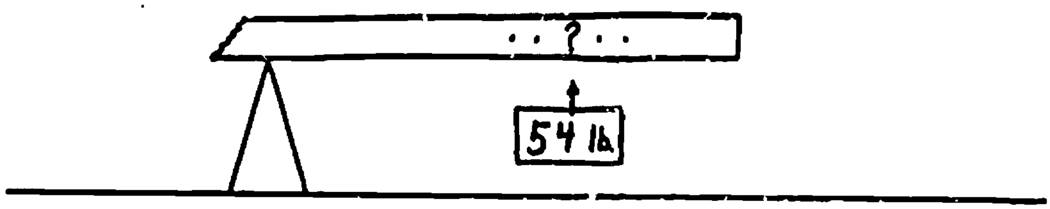
The clo kwise moments are: Moments = 7 pounds x 5 feet + 3 pounds x 8 feet
Moments = 35 pound-feet + 24 pound-feet
Moments = 59 pound-feet

Counterclockwise moments: Moments = 9 pounds x 6 feet
Moment = 54 pound-feet

Since the clockwise moments are larger than the counterclockwise moments (59-54 = 5) the lever will rotate in a clockwise direction.

To find where a given weight must be placed on an unbalanced lever in order to have it balance, we must know either the moment of the side of the lever on which an object has already been placed or the moment needed to balance the lever. If only one weight is to be placed on each side of the lever, then we divide the weight of the object whose lever arm we are going to find into the moment of the opposite side of the lever. This gives us the distance from the fulcrum that the weight must be placed in order to have the lever balance. If a weight is already attached to each side of the lever and a third weight will be used to balance it, we first calculate the moments for the two sides and find the difference between them. (Subtract the smaller moment from the larger moment.) We then divide this difference in the moments by the weight of the third object to find how far it must be placed from the fulcrum to have the lever balance.

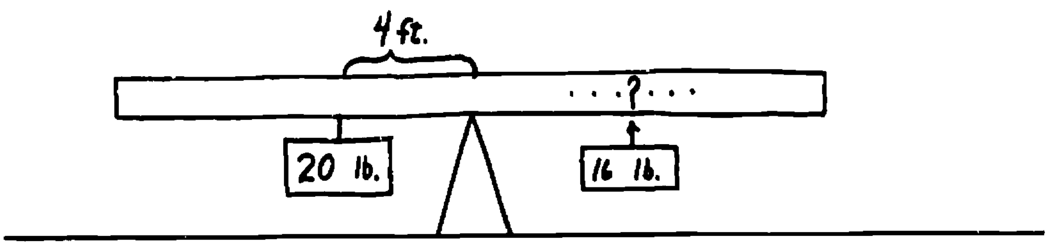
At what distance from the fulcrum must the 54 pound weight be placed to have the moment of the lever arm pictured below equal to 378 pound-feet?



To find where a given weight must be placed on an unbalanced lever in order to have it balance, we must know either the moment of the side of the lever on which an object has already been placed or the moment needed to balance the lever. If only one weight is to be placed on each side of the lever, then we divide the weight of the object whose lever arm we are going to find into the moment of the opposite side of the lever. This gives us the distance from the fulcrum that the weight must be placed in order to have the lever balance. If a weight is already attached to each side of the lever and a third weight will be used to balance it, we first calculate the moments for the two sides and find the difference between them. (Subtract the smaller moment from the larger moment.) We then divide this difference in the moments by the weight of the third object to find how far it must be placed from the pivot to have the lever balance.

In this problem, the weight must be placed 7 feet from the pivot. 378 pound-feet/54 pounds = 7 feet.

At what distance to the right of the center of the lever must a 16-pound weight be hung in order to have the lever balance?

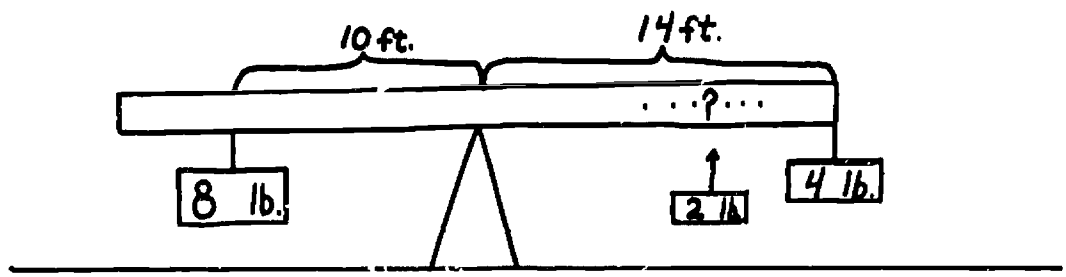


To find where a given weight must be placed on an unbalanced lever in order to have it balance, we must know either the moment of the side of the lever on which an object has already been placed or the moment needed to balance the lever. If only one weight is to be placed on each side of the lever, then we divide the weight of the object whose lever arm we are going to find into the moment of the opposite side of the lever. This gives us the distance from the fulcrum that the weight must be placed in order to have the lever balance. If a weight is already attached to each side of the lever and a third weight will be used to balance it, we first calculate the moments for the two sides and find the difference between them. (Subtract the smaller moment from the larger moment.) We then divide this difference in the moments by the weight of the third object to find how far it must be placed from the pivot to have the lever balance.

In this problem, the 16-pound weight must be hung 5 feet to the right of the center in order to have the lever balance.

$$\frac{20 \text{ pounds} \times 4 \text{ feet}}{16 \text{ pounds}} = 5 \text{ feet}$$

At what distance to the right of the center must the 2-pound weight be attached in order to have the lever balance?

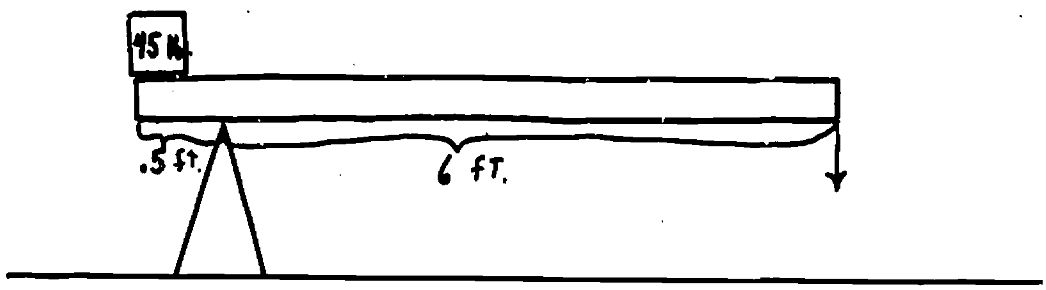


To find where a given weight must be placed on an unbalanced lever in order to have it balance, we must know either the moment of the side of the lever on which an object has already been placed or the moment needed to balance the lever. If only one weight is to be placed on each side of the lever, then we divide the weight of the object whose lever arm we are going to find into the moment of the opposite side of the lever. This gives us the distance from the fulcrum that the weight must be placed in order to have the lever balance. If a weight is already attached to each side of the lever and a third weight will be used to balance it, we first calculate the moments for the two sides and find the difference between them. (Subtract the smaller moment from the larger moment.) We then divide this difference in the moments by the weight of the third object to find how far it must be placed from the pivot to have the lever balance.

In this problem: Without the 2-pound mass the moment on the left side is 80 x 10, or 80 pound-feet and the moment on the right side is 4 x 14, or 5 pound-feet. A difference of 80-56, or 24 pound-feet, is needed to balance the lever. By dividing 2 pounds into the 24 pound-feet needed, we find that the 2-pound weight must be placed <u>at the 12-foot mark</u>.

To find the amount of weight needed on one end of a lever to lift an object on the opposite end, we first calculate the moment for the end of the lever on which the object has been placed. This moment is calculated by multiplying the object's weight by its lever arm length. We then divide this moment by the lever arm length of the opposite end to obtain the weight needed to lift the object.

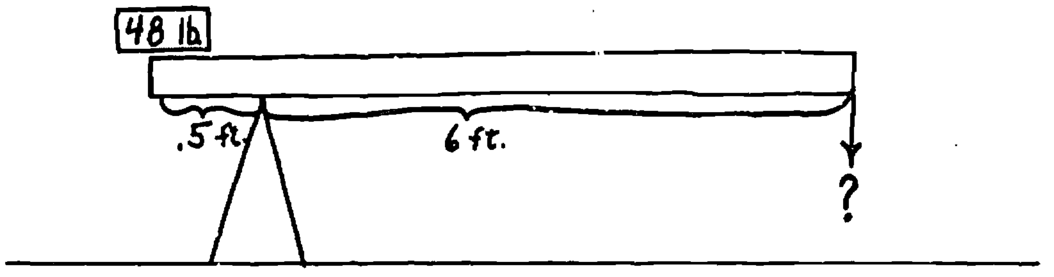
In the drawing below, what is the weight of the object that is to be moved?



To find the amount of weight needed on one end of a lever to lift an object on the opposite end, we first calculate the moment for the end of the lever on which the object has been placed. This moment is calculated by multiplying the object's weight by its lever arm length. We then divide this moment by the lever arm length of the opposite end to obtain the weight needed to lift the object.

In this problem, the object to be moved by the lever weighs 45 pounds.

Find (calculate) the moment for the end of the lever on which the object to be moved is located?

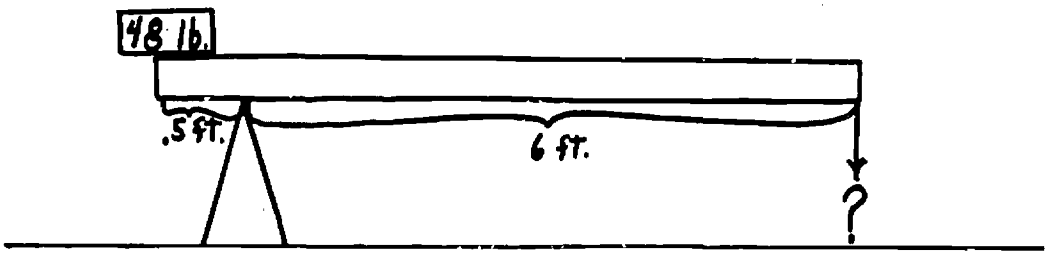


To find the amount of weight needed on one end of a lever to lift an object on the opposite end, we first calculate the moment for the end of the lever on which the object has been placed. This moment is calculated by multiplying the object's weight by its lever arm length. We then divide this moment by the lever arm length of the opposite end to obtain the weight needed to lift the object.

In this problem, the moment is calculated by:

Moment = weight x lever arm
Moment = 48 pounds x .5 feet
Moment = 24 pound-feet

How much weight must be attached to the right end of the lever in order to have the object on the opposite end balance.



To find the amount of weight needed on one end of a lever to lift an object on the opposite end, we first calculate the moment for the end of the lever on which the object has been placed. This moment is calculated by multiplying the object's weight by its lever arm length. We then divide this moment by the lever arm length of the opposite end to obtain the weight needed to lift the object.

In this problem, the moment of the left end is (48 pounds x .5 feet) or 24 pound-feet. This moment divided by the length of the other lever arm (24 pound-feet/6 feet) is equal to 4 pounds, or the weight necessary to have the object on the opposite end balance.

Appendix D
Instructions for Cooperating Teachers

In this study, as in any study, it is imperative that there be consistency in the instructions and treatments that students receive in their individual classrooms. Variation in the way in which teachers instruct or handle their classes in an experimental situation may cause significant deviations in the experimental outcome that cannot be accounted for or explained. The following statements are therefore set down as guidelines to establish a consistent set of instructions to be given students and procedures to be followed by students in each of the eight class sections.

<u>October 6th thru 22nd</u>.

During this time, students in each of the classes will be working in the programmed text in astronomy. Students will work on an individual basis, and at their own rate. Teachers should observe student progress and assist each student in reaching a stopping point in the text on Wednesday, October 22. Prior to October 22 teachers should explain to students that on this date, they will be taking a break from studying astronomy to do some work for a short time with some new learning materials. Tell students that they will be working for a short time with these new learning materials, that they should work as hard as they can in learning from these materials, and that whatever grade they earn will count towards their marking period grade.

Thursday, October 23

On this date all students in the eight class sections will take the Ability to Interpret Reading Materials in the Natural Sciences Test. Students should be told that you are trying to determine how well they can read science materials before they work on some new subject material. Please do not refer to the new materials as an experiment.

1. Distribute pre-labeled answered sheets and test booklets to students as quickly as possible. Students should be given as much of the class period as possible to work on the test.
2. Read the sample question aloud to the class and give them a brief explanation.
3. Instruct classes that they may guess at answers if they feel reasonably sure that they know the answer. That is, guessing will not count against them.
4. Allow students to work up to, but not past, the passing bell for that period. Collect papers as students finish and return them to Mr. Yost.
5. Make a list of students who are absent so that they may take the test at a later date. Students not taking this test may begin work with the remainder of the class on the next day. Arrangements to give the test to students who were absent will be made at a later date.

Friday, October 24

1. Take roll - note absences.
2. Send the control group students to the library. Tell them that they will be going to the library to read during science period for the next few days (Each teacher will be given a list of the names of students in the control group).
3. Repeat the cover story if you have not done so as yet (This is not an experiment).
4. Distribute learning materials booklets - instruct students not to open them.
5. Read the instructions on page 1 of the booklet aloud as students follow along in their own booklets. Answer student questions.
6. Give students about 30 minutes to work in their booklets. Please make note of exactly how long (in minutes) you give students to work.
7. Collect all booklets - tell students how and where they will pick up their own booklets as they enter class on Monday. Explain that booklets will be given out at the beginning and collected at the end of each class period.

Monday, Tuesday, Wednesday, October 27, 28, 29

1. Each day students enter class, they should pick up their booklets and begin working.
2. No talking should be allowed.

Monday, Tuesday, Wednesday, October 27, 28, 29 Cont'd.

3. Teachers should circulate through class to check student progress and to see if they are following directions. If a student is not following directions, please attempt to make the correction process a personal (one to one), somewhat pleasurable experience. A strenuous individual or class reprimand may cause serious harm to the outcome of the study. Please try to avoid this type of activity.
4. Teachers should not answer student questions dealing with subject matter content.
5. If a student is absent, mark the day that he is absent on the front of his booklet (please circle and initial).
6. If a student, for some reason, leaves the room or stops working, mark the number of minutes that he loses on the front of his booklet (please circle and initial the number).
7. When a student completes his booklet, collect it, write the day he finished and the number of minutes that day that he worked on it on the cover. If more than 15 minutes remains in the period, allow him to begin on the final evaluation.

Thursday, Friday, October 30th, 31st. (approximately)

1. As students complete their booklets, they should begin taking their final evaluations.
2. Normal testing procedures should be adhered to - no talking or working together. Students may guess at answers if they feel that they might know the answer.

Thursday, Friday, October 30th, 31st. (approximately) Cont'd.

3. To avoid any effect on the experimental outcome of students taking one of the tests first, each teacher will have one of his classes take Part I of the test first and the other of his classes take Part II of the test first. It makes little difference which of your classes takes which part first. Please inform me of your choice. Needless to say, it would be catastrophic if students answered Part II of the evaluation in the Part I answer blanks. PLEASE CHECK EACH STUDENTS WORK TO MAKE SURE THAT THIS DOES NOT HAPPEN.
4. When a student completes one quiz, he may begin immediately on the next. The questions that a student does not complete in one day, he may complete the next day. Please collect tests and answer sheets at the conclusion of each period and redistribute them the following day.
5. When a student completes both parts of the evaluation he may sit quietly and read or work on some other material.

Appendix E

Judges Evaluations of Content Validity for Test II

Judges / Question No.	1	2	3	4	5
1.	x	x	x	x	x
2.	x	x	x	x	x
3.	o	x	x	x	x
4.	x	x	x	x	x
5.	x	x	x	x	x
6.	x	o	x	x	x
7.	x	x	x	x	x
8.	x	x	x	x	x
9.	x	x	x	x	x
10.	x	x	x	x	o
11.	x	x	x	x	x
12.	x	x	x	x	x
13.	x	x	x	x	x
14.	x	x	x	x	x
15.	o	x	x	x	x
16.	o	x	x	x	x
17.	x	x	x	x	x
18.	x	x	x	x	x
19.	x	x	o	x	x
20.	x	x	o	x	x
21.	x	x	o	x	x
22.	x	x	o	x	x
23.	x	x	x	x	x
24.	x	x	x	x	x
25.	x	o	x	x	x

x = Acceptable.
o = Not Acceptable.

Appendix F

Evaluation Instruments

FINAL EVALUATION

PART I

Name: ______________________
Last First

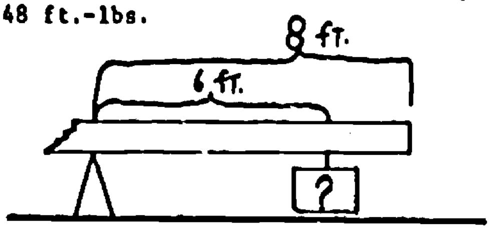
Teacher: ______________________

DIRECTIONS: Read each question and decide which word, phrase, or number is the correct answer. Look at the answer spaces on your answer sheet. Fill in the space which has the same letter as the answer you have chosen. There is only one correct answer for each question. Should you change an answer, you should completely erase the incorrect answer.

SAMPLE: The person who invented the lever was probably a(n)________________.

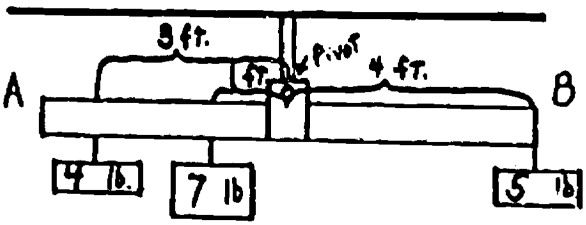
a. auto mechanic c. sailor
b. caveman d. airplane pilot

1. A weight of __________must be placed on the lever to have the moment equal 48 ft.-lbs.



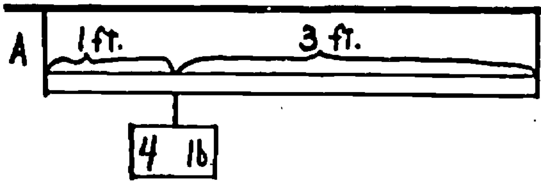
a. 6 lbs. c. 2 lbs.
b. 14 lbs. d. 8 lbs.

2. The moment for side A of the lever is ____________.



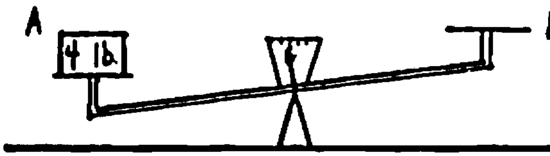
a. 28 lb.-ft. c. 20 lb.-ft.
b. 84 lb.-ft. d. 19 lb.-ft.

3. The bar pictured below is supported fro ceiling by two strings and has an objec hanging from it.



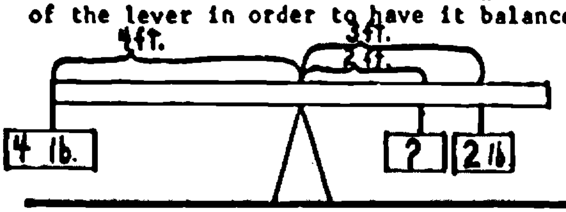
a. The strings support equal amounts o the object's weight.
b. String A supports more of the objec weight than string B.
c. String B supports more of the objec weight than string A.

4. The picture below is of an equal-arm pa balance. If a 4-pound weight is placed the pan of side A, how much weight must placed on the pan of side B in order to have the pans balance?



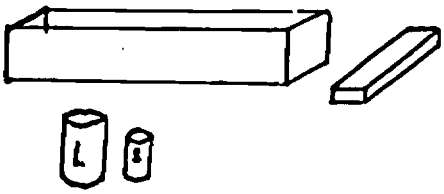
a. A weight larger than four pounds.
b. A weight smaller than four pounds.
c. A weight of four pounds.

5. A weight of __________must be placed 2 feet from the fulcrum on the right si of the lever in order to have it balanc



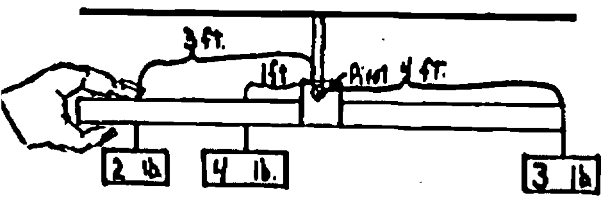
a. 16 lbs. c. 6 lbs.
b. 5 lbs. d. 10 lbs.

6. A boy is given an empty box, a rectangular stick, and two cans of water as shown in the drawing below. Both cans are filled with water and one is much larger than the other. If the boy places the larger can of water inside the box on the left edge, and the smaller can of water inside the box on the right edge, approximately where should the stick be placed to have the box balance on it and not touch the top of the table?



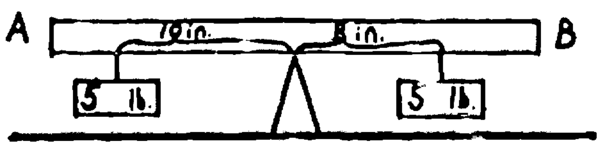
a. Under the center of the box.
b. To the right of the center of the box.
c. Under the left edge of the box.
d. To the left of the center of the box.

7. When the hand releases the lever, the lever will ____________________.



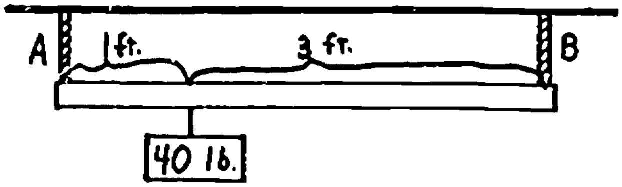
a. Remain balanced.
b. Rotate in a clockwise direction.
c. Rotate in a counterclockwise direction.

8. The drawing below is of a lever with equal weights hanging at different distances from the fulcrum. What is the difference in the distances between weight A and the fulcrum and weight B and the fulcrum?



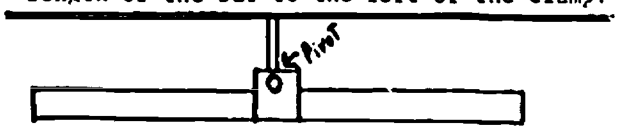
a. 18 inches. c. 8 inches.
b. 10 inches. d. 2 inches.

9. The bar pictured below is supported from a ceiling by two ropes and has an object hanging from it. How much of the object's weight is supported by rope A?



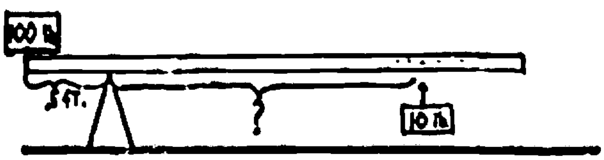
a. 10 lbs. c. 30 lbs.
b. 20 lbs. d. 12 lbs.

10. The bar picture below is hanging from a ceiling by a clamp that will allow it to pivot or turn. The bar is of the same thickness all along its length, and is balanced. How does the length of the bar to the right of the clamp compare with the length of the bar to the left of the clamp?



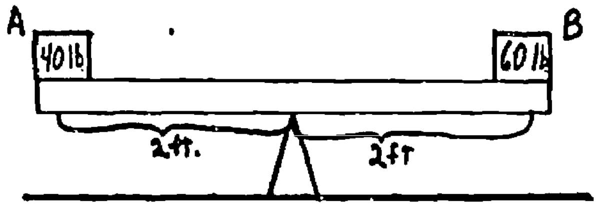
a. They are of the same length.
b. The length of the bar on the left is longer than the length of the bar on the right.
c. The length of the bar on the right is longer than the length of the bar on the left.
d. The length of the bar on the left is shorter than the length of the bar on the right.

11. A pull downward of 10 pounds must be exerted at a point __________ to the right of the fulcrum in order to balance the object on the left.



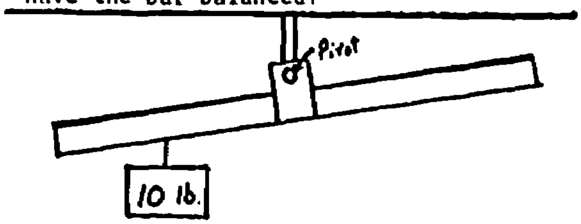
a. 4 ft. c. 6 ft.
b. 5 ft. d. 7 ft.

12. The weight on end __________ exerts a push downward of ________ pounds more than the weight on the opposite end.



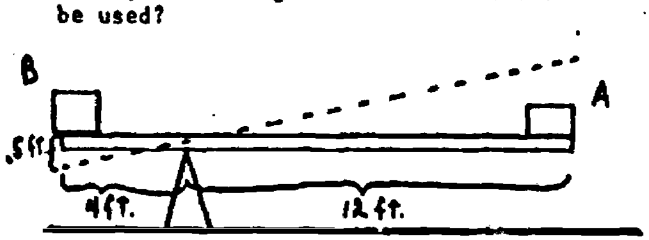
a. A, 40
b. B, 20
c. A, 20
d. B, 60

13. The bar pictured below is hanging from a ceiling by a clamp that will allow it to pivot or turn. The bar is of the same thickness all along its length and is balanced. If a 10-pound weight is hung half way out on the left side, where can a second 10-pound weight be hung to again have the bar balanced?



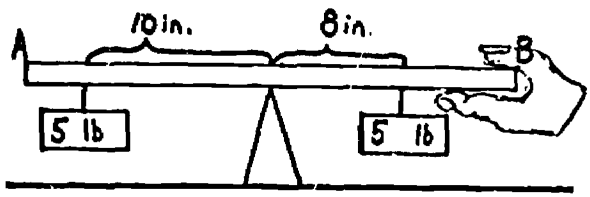
a. Just to the right of the pivot.
b. Half way out on the right side.
c. Half way out on the left side.
d. On the extreme right end of the bar.

14. In calculating the amount of work done in moving the object on end B of the lever, which length or distance would be used?



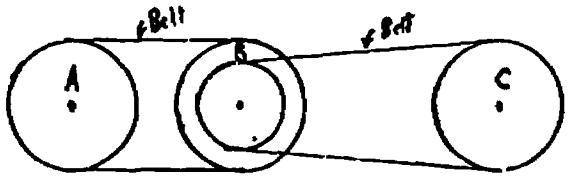
a. 48 ft.
b. .5 ft.
c. 2. ft.
d. 4. ft.

15. When the hand releases the lever pictured below-



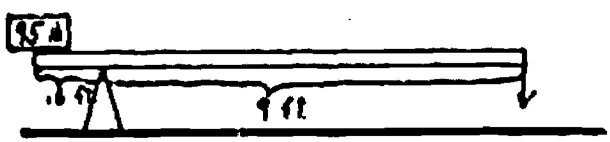
a. It will remain in a balanced position.
b. End A will go down.
c. End B will go down.
d. It will rotate in a clockwise direction.

16. The drawing below is of three wheels that are free to rotate. A and C are single wheels and wheel B is two wheels firmly fastened together. When one wheel turns, the belts will cause all of the other wheels to turn. If wheel A rotates in a clockwise direction, in what direction will wheel C rotate?



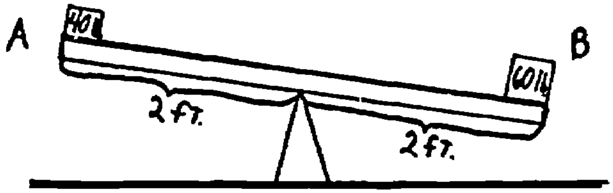
a. Clockwise
b. Counterclockwise

17. ____________ of work will be done in lifting the 95-pound object a vertical distance of .2 feet with the lever.



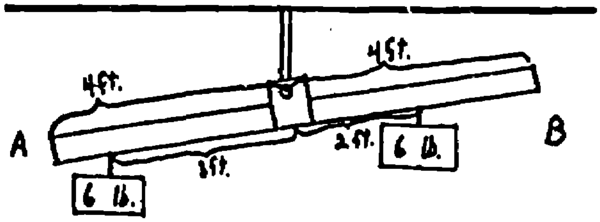
a. 19 ft.-lbs.
b. 855 ft.-lbs.
c. 54 ft.-lbs.
d. None of the above amounts

18. The moment for side A is ______________.



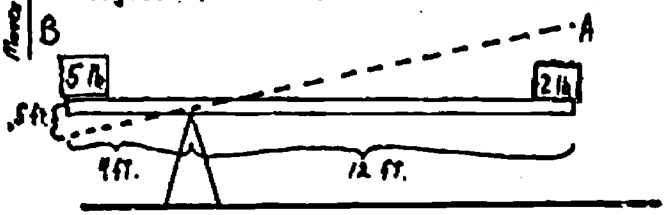
a. 80 lb.-ft. c. 40 ft.-lbs.
b. 120 ft.-lbs. d. 80 ft.-lbs.

19. The difference in the moments for ends A and B of the lever are ______________.



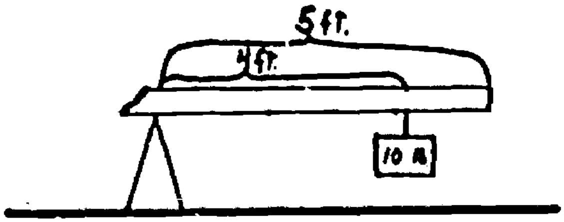
a. 12 ft.-lbs. c. 30 ft.-lbs.
b. 18 ft.-lbs. d. 6 ft.-lbs.

20. What is the difference between the amount of work done in moving the 5-pound object and the moment of the 5-pound object? (Hint: Moment - Work = Answer)



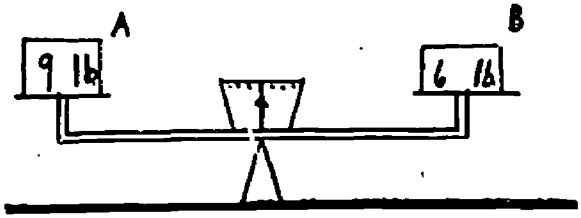
a. 2.5 c. 2.
b. 17.5 d. 18.5

21. The drawing below shows half of a lever. What is the length of the lever arm that the 10-pound weight acts on?



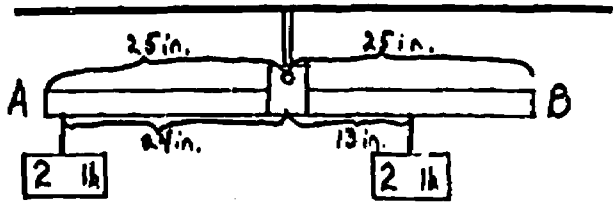
a. 5. ft. c. 9. ft.
b. 4. ft. d. 20. ft.

22. The picture below is of an equal-arm pan balance. An additional weight of ________ pounds must be placed on side ________ to have it balance.



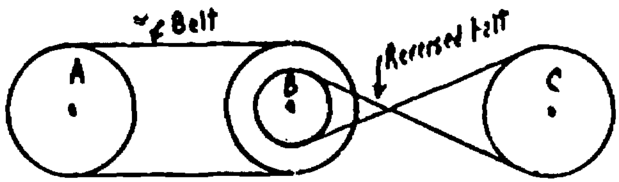
a. 6 ; A c. 9 ; B
b. 3 ; B d. 3 ; A

23. In the lever pictured below, the weight on side ____________ acts through a lever arm ________ inches longer than the weight on the opposite side.



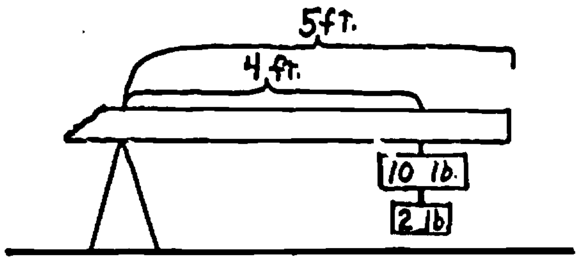
a. A ; 1 c. B ; 12
b. A ; 11 d. B ; 11

24. The drawing below is of three wheels that are free to rotate. Wheels A and C are single wheels and wheel B is two wheels firmly fastened together. The belt between wheels A and B is straight and the belt between wheels B and C is reversed. If wheel A rotates in a clockwise direction, in what direction will wheel C rotate?



a. Clockwise b. Counterclockwise

25. The drawing below shows half a lever. The moment for the attached weights is ______________.



a. 50 ft.-lbs.
b. 60 ft.-lbs.
c. 40 ft.-lbs.
d. None of the above.

FINAL EVALUATION
PART II

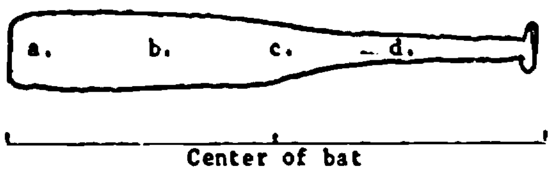
Name: ______________________
Last First

Teacher: ______________________

1. All of an object's weight seems to be concentrated at its __________.

   a. moment c. fulcrum
   b. center of gravity d. center

2. The center of gravity of the baseball bat is located nearest to point ________.



3. Which usually would be the best shears for cutting hard metal?

   a. b. c. d.

4. The support or point about which a lever rotates is called the __________.

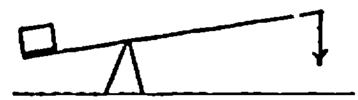
   a. fulcrum c. Center of gravity
   b. moment d. force

5. __________ opposes motion and causes the conversion of mechanical energy to heat energy.

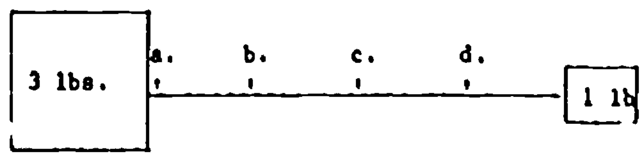
   a. friction c. inertia
   b. weight d. mass

6. The lever pictured below increases the __________ that acts on the box.

   a. weight
   b. energy
   c. force
   d. moment



7. A three-pound object and a one-pound object are connected by a rod. The center of gravity of the two connected objects is at point ________.

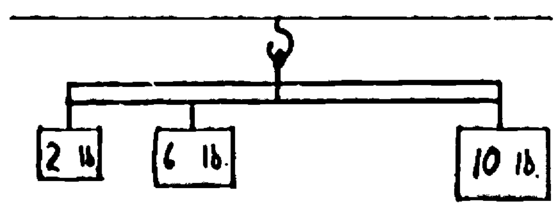


8. __________ is a push, pull, or lift that is exerted on an object.

   a. mass c. work
   b. moment d. force

9. The downward pull on the hook is ________.

   a. 10 lb. c. 18 lb.
   b. 16 lb. d. 8 lb.



10. When using a lever to move an object, a man places the fulcrum closer to the object to reduce:

a. The weight of the object being moved.
b. The amount of effort force that he must apply.
c. The force that is applied to the object.
d. The number of times that his force is increased.

11. Friction causes loss of mechanical energy as ___________.

a. mass
b. temperature
c. weight
d. heat

12. A man does 45 ft-lbs. of work in pulling on a lever that does 40 ft-lbs. of work in lifting a rock. The amount of work that was used in overcoming friction was ________.

a. 45 ft-lbs.
b. 40 ft-lbs.
c. 5 ft-lbs.
d. 85 ft-lbs.

13. A lever having 50 ft-lbs. of work input and 40 ft-lbs. of work output has ____ efficiency.

a. 10%
b. 80%
c. 90%
d. 100%

14. ___________ reduces the amount of useful work that we get out of machines.

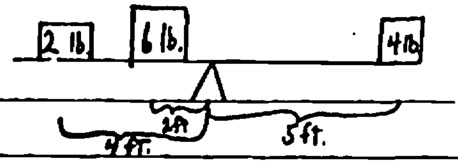
a. the size of the resistance
b. friction
c. efficiency
d. the weight of the object being moved

15. Car is to automobile as moment is to ___________.

a. torque
b. weight
c. lever arm
d. mass

16. The downward force on the fulcrum is ___________.

a. 40 lbs.
b. 8 lbs.
c. 4 lbs.
d. 12 lbs.

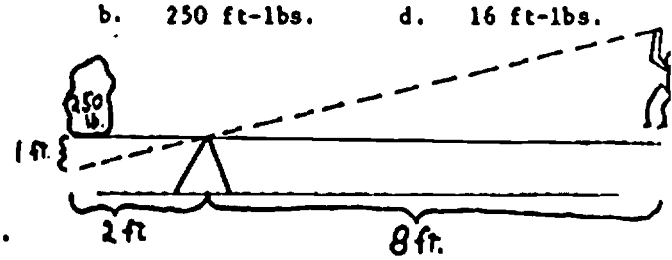


17. When calculating a moment, we multiply a force by a ________.

a. torque
b. length
c. weight
d. mass

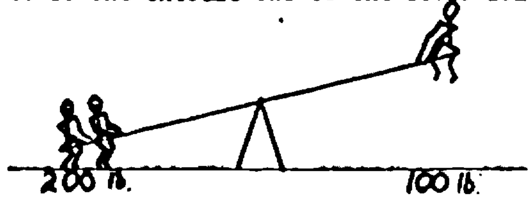
18. How much work does the boy do in pulling down on the lever to move the rock?

a. 1000 lb-ft.
b. 250 ft-lbs.
c. 500 ft-lbs.
d. 16 ft-lbs.



19. One boy on one end of a see-saw can lift 2 boys on the other end if the 2 boys move-

a. Closer to the fulcrum, making a shorter lever arm.
b. Farther away from the fulcrum, making a longer lever arm.
c. To the extreme end of the lever arm.



20. A man can lift a 25-lb. box with a lever by applying a force of 5 lbs. If he moves the fulcrum farther away from the box, he would then

a. use less force
b. use more force
c. use the same amount of force.

21. Its easier for a man to raise a heavy object with a lever if he moves the fulcrum closer to the object because

a. The object's moment becomes larger than the man's moment.
b. The man's force is multiplied a larger number of times.
c. The weight of the object is reduced.
d. A great deal less force is needed.

22. When using a lever, the closer the fulcrum is to the object being moved, the

a. shorter the effort arm of the object being moved.
b. larger the effort force has to be.
c. larger an object that can be lifted with a given force.
d. smaller the object we can lift.

23. A boy pulls on one end of a lever and moves a rock. If we ignore friction, the amount of work that the boy does in pulling on the lever is

a. greater than the amount of work done on the rock.
b. equal to the amount of work done on the rock.
c. less than the amount of work done on the rock.

24. __________ reduces the work output of a machine.

a. moment c. inertia
b. friction d. weight

25. Work is to the force parallel to distance moved as __________ is to the force at right angles to lever arm length.

a. fulcrum c. weight
b. movement d. moment