THE DESIGN AND VALIDATION OF A TELEVISION BASED COGNITIVE LEVEL ASSESSMENT INSTRUMENT

JOHN DAVID FAST

THE UNIVERSITY OF MANITOBA

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF EDUCATION

DEPARTMENT OF CURRICULUM: MATHEMATICS AND NATURAL SCIENCES

1979

THE DESIGN AND VALIDATION OF A TELEVISION BASED COGNITIVE LEVEL ASSESSMENT INSTRUMENT

BY

JOHN DAVID FAST

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF EDUCATION

ABSTRACT

The use of Piagetian tasks to measure the developmental level of students, followed by the application of the results to choose and create science curriculum matched to the developmental level of the learner, creates a need for an instrument to assess the cognitive level of large groups of students.

The purpose of this study was to construct and validate a television based evaluation instrument to measure cognitive level based on the theory of Jean Piaget. The evaluation instrument was a group, paper and pencil test, using videotape recorder (VTR) as the medium for its presentation to students.

The validation study included the administration of a Rods test, designed to discriminate between concrete and formal operational students, and personal Piagetian task interviews with 30 of the 103 students in the sample. The sample was selected from grade nine students at Acadia Junior High School in Winnipeg.

The results indicated that the cognitive level of large groups of students can be tested effectively through the medium of television.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

DEFINITION AND SCOPE OF STUDY

Need for Study

In a recent study, Raven (1974) suggests that the implication of Piaget's research is that the use of logical operations by science students may enhance the acquisition and comprehension of science concepts whose structure corresponds to these logical operations. Many programs are prefaced with references to a developmental approach according to Piaget. In Manitoba, both elementary and secondary science review committees have adopted a developmental basis for creating and choosing science curriculum. This emphasis on cognitive development as defined by Piaget creates a need to measure the developmental level of students at many grade levels. It will be necessary, when using Piagetian tasks to quantify gain in cognitive development, to be wary of the effect of repitition and practice on the same or similar Piagetian tasks up through the grades. A valid and reliable series of classroom tests, such as the one developed in this study, could perhaps prove useful in the creation of an organized system of evaluation of programs, whose aims include cognitive growth.

Research based on Piaget's theory implies that a mismatch may exist between the level of cognitive development of students and the complexity of logic required by the science curriculum. About this mismatch between student and curriculum, Kaufman and Konicek (1972) suggest:

> For it seems to be more and more evident as one works with children that adult concepts are somewhat puzzling to children and not very useful in many cases. Children should be allowed

> a maximum of activity on their own, directed by means to materials which permit their activities to be cognitively useful. In the area of logico-mathematical structure, children have real understanding only of that which they invent themselves, and each time that we try to teach them something too quickly, we keep them from reinventing it themselves (p. 10).

Piaget's theory suggests that activities provided for the student should be such that they are closely matched to his cognitive level, but of a nature so as to stimulate his present mode of reasoning and cultivate it. Other forms of mismatch could include aptitude-treatment interaction and learning style (Hunt, 1973) problems. All result in a poor learning environment for the science student. An attempt to match the curriculum to the student in these respects at least requires an assessment tool to establish the cognitive level of the individual students. Since Piaget does not provide a model for the application of his theories to development of science education, we must do it ourselves. Determining the cognitive level of our students at various ages and grades on a large scale could be the first step toward the development of science programs which are truly consistent with Piaget's stage theories, not merely oblique references to them.

One must be cautious about over zealous application and misapplication of Piaget's theory to curriculum development. In this regard, Kaufman and and Konicek (1972) claim that:

> ... the Piagetian stages have become for many educators an alternative to the I.Q. or capacity for learning. The pressing question for such educators has become one of how can a given program accelerate the acquisition of a particular stage? Curricula, which purport to be Piagetian, have been used in the school, and some are primarily concerned with speeding up the intellectual growth of children (p. 6).

In conjunction with this view, Lawson and Devito (1974) make the following statement:

> With the increasing interest in the application of Piagetian

> theory for curriculum design and evaluation, the use of Piagetian tasks to quantify gain in intellectual development has and will continue to become more widespread (p. 1).

These statements describe the misinterpretation of Piaget's theory and a subsequent misapplication of it to establish goals in public education and to develop programs. No where in his theory does Piaget suggest a positive correlation between cognitive level and I.Q. or the possibility of accelerating cognitive growth (Inhelder, Sinclair, and Bovet, 1974). They are different theoretical constructs. The proper facilitation of cognitive growth through intellectual disonance in favor of acceleration is a better aim for science education. Lawson and Devito's prediction that Piaget's theory will be widely used to develop science curriculum and his tasks used to test cognitive growth has become evident over the past three years.

The preceding arguments have motivated a search for reliable and valid assessments of cognitive level. When one turns to Piaget's research for assistance in this regard, one finds he has directed his efforts toward an elucidation of two basic questions. These are 1) What is the nature of knowledge, and 2) How does man come to know?.

The verification methods used by Piaget to investigate his two epistemological questions have been: 1) a historico-critical analysis of scientific notions; and 2) the study of children as an ontogenic tool to ascertain the development of logical thought processes.

Individually administered Piagetian tasks, although found to be valid indicators of developmental level (eg. Bart, 1971; Goldschmid, 1967; Lawson and Blake, 1976; Lawson and Renner, 1975; Lovell and Shields, 1967) require an experienced interviewer, special materials and equipment, and are too time consuming and expensive for practical classroom use. This investigation sought to develop and validate a classroom assessment of cognitive development in terms of Piaget's concrete and early formal operational reasoning stages that could be administered to groups of students.

And finally, according to Ball and Sayre (1973):

> Piagetian cognitive development, a physiological as well as psychological process, appears to be a major factor in determining grades received by science students (p. v).

The extent to which science grades and developmental level correlate could be investigated with the aid of the cognitive level assessment instrument developed in this study. The instrument was a television based classroom test, with each student completing a booklet in response to tasks viewed on a television monitor. It is hoped that the educational value of the classroom test lies in this area, and not as a test of Piaget's theory.

Purpose

The purpose of this study was to construct and validate an evaluation instrument to measure cognitive level based on the developmental theory of Jean Piaget. The evaluation instrument was a group, paper and pencil test, using videotape recorder (VTR) as the medium for its presentation to students. The use of VTR for presentation is the significant difference between this and other paper and pencil assessments. The details of test item selection, scoring criteria, and administration of tests are included in chapter 3.

Problem Statment

The problem of this study was to develop and validate an instrument which would: 1) measure concrete and early formal operational reasoning;

2) be capable of administration to classes of secondary school students in a relatively short period of time; 3) be easily scored; 4) use VTR as a presentation medium with answer booklets requiring as little reading and writing as possible; and 5) include a variety of problems to assure a high degree of reliability.

Limitations of the Study

There are certain limitations inherent in a study of this type. The logistics of administering the personal interview assessment (PIA) and the time necessary to interview each student, reduced the number of students interviewed to thirty. The linguistic component of test design was considered only in an intuitive sense because of the author's lack of training in this field.

The control and effect of treatment, which students received in the classroom during the course of the study, was not considered. It was assumed that training effects were minimal and that experiences in the regular science classroom which could dramatically increase cognitive development were not taking place.

The sample coming from an upper-middle class community could infuence the range of scores one might expect from a neighborhood of greater socio-economic heterogeneity.

Definition of Terms

Formal operational stage:	Students operating at this stage can deal with abstractions and are capable of hypo-thetico-deductive reasoning, even though unrelated to real events. They have the ability to conserve displacement volume,

effectively eliminate contradictions, analyze combinatorially, and identify, separate and exclude inoperative variables.

Concrete operational stage: Students operating at this stage do not have this freedom of thought; their thinking is restricted to real events. They have the ability to serially order, conserve amount-length-area-weight-volume, perform 2-way classification, and make one-to-one correspondences.

Hypothetico-deductive reasoning: Reasoning occurring within a combinatorial system, as defined by Piaget, making it possible to generate hypotheses based on all possible combinations.

Cognitive or developmental level: Used in reference to the stage, concrete or formal, which best describes the individual's reasoning with respect to various Piagetian based tasks.

CHAPTER II

ASSESSMENT OF COGNITIVE LEVEL-
A REVIEW OF THE LITERATURE

Introduction

The documentation of assessment of cognitive level for the purpose of investigating the nature of knowledge and how man comes to know is extensive in books and papers written by Piaget over the past 50 years. The work of Inhelder and Piaget (1958) is widely known and is also available in such secondary works as Flavell (1963). Most of Piaget's many works remain in a difficult to read French version (Mallon, 1976). Piaget's best known experiments are described in his first five books, the originals of which were published in the 1920s: *The Language and Thought of the Child* (English version); *The Child's Conception of the World* (English version); *Judgement and Reasoning in the Child* (English version); *The Child's Conception of Physical Causality* (English version); and *The Moral Judgement of the Child* (English version). The primary purpose of these studies was to provide data for a systematic and comprehensive epistemology. In the past ten years, however, they have provided education researchers with a protocol for the interview technique used in assessing cognitive development.

Traditionally, assessment of cognitive development has been based on the work of Binet, with two methodological approaches; individual and group-administered tests. These approaches have been based on psychometric rigor and convenience, with little regard to understanding why a subject performed as he did. An individual's assessment and subsequent rating has been dependent on the mastery of specific information and on his position relative to a norm group within the normal curve model of probability. Such tests are not based on any particular theory of psychological development.

Piaget used a variation of the individual testing situation (his methodé clinique) and has attempted to assess cognitive development in a manner which does not depend on specific knowledge or upon how an individual performs relative to a norm group within the normal curve model. Rather, his work has focused on assessing cognitive constructs that are necessary for competent interaction with the world, generally not teachable, and develop in all individuals at different rates, but in the same sequence. Although cognitive development is continuous, there are periods of time within which the individual's behavior is fairly stable and qualitatively different from the behavior of other periods. Because this study is concerned with concrete to early formal reasoning ability, the review of the literature was restricted to assessment studies related to the period of the adolescent's reasoning at the junior and senior high school age.

Assessment Studies

In an effort to develop a cognitive assessment instrument for the classroom, Gray (1973) created a paper and pencil test administered to 96 students between the ages of 9 and 16. He criterion-referenced three specific Piagetian tasks, exclusion, proportion and combination, on which he individually interviewed half of his subjects. The entire sample then wrote the paper and pencil test which he developed. The second half of the sample was also assessed by interview. He found that subjects taking the Piagetian tasks first and the written test second did no better than the other group, even though the logical operations tested on both tests were the same. It is noted that his multiple-choice method of test construction did not provide useful information on many test items. He suggests that an open-ended type of question on the written test would provide more reliable data. In an attempt to eliminate reading difficulty as a variable affecting his study, he excluded those students known to have reading problems from his investigation.

Raven (1973) solved the reading level problem in Raven's Test of Logical Operations (RTLO) by providing assistance to the pupils when they experienced difficulty in reading. The general format of his test involved the presentation of a problem in pictorial form, followed by a brief question posed in written form. At least three possible solutions were presented in pictorial form below. His results indicate that the items measured the logical operations for which they were designed, as well as providing the teacher with some evidence to model a concept for instruction.

The paper and pencil Rods test, which was used in this study, was designed and validated against the clinical interview technique with junior high students by Harris (1974). It has been utilized by Coulter (1976) in a student-curriculum mismatch study in Manitoba high schools. The test consists of a series of illustrations of the rod bending device described by Inhelder and Piaget (1958) with questions accompanying each illustration.

Tests which are strictly paper and pencil measures with a variety of formats, but which lack sufficient validation data, are numerous but not widely used (e.g. Burney, 1974; Tisher-Dale, 1975; Tomlinson-Keasey, 1975; Ankney and Joyce, 1974). Lawson (1977) suggests that even though these have the advantage of eliminating the need for special equipment, they lack the motivating aspects which arise from handling materials and equipment.

Conversely, Bruner (1966) argued against the need for materials or special equipment. He found that many subjects who did not conserve liquid amount when physical materials (glasses of water) were placed before them, did conserve when the liquids were partially screened. To these subjects, he claimed, the materials were distracting and evidently prevented them from reasoning logically about the problem. If one considered some of the study's possible design faults, the significance of the study was questionable.

In an attempt to retain as many aspects of the clinical interview as possible, Rowell and Hofman (1975) developed a testing format in which each student is provided a set of laboratory equipment and a test booklet of instructions and questions. This procedure again requires large quantities of equipment, is more time consuming, and therefore restricts assessment to smaller groups of students.

Shayer and Wharry (1975) attempted a clinical situation by demonstrating materials before the class and asking questions to which students record answers in test booklets. The technique suggests a good comprimise between purely clinical interviews and paper and pencil assessments. However, according to Lawson (1977), they have not developed a sufficient variety of formal level problems to make their instrument entirely satisfactory. Lawson himself adopted this method in his most recent study and expanded the number and the variety of problems posed. His research indicates that the same psychological parameters measured by classical Piagetian interview tasks were measured by his test with a fairly high degree of reliability.

The possible learning effect due to taking and retaking the same Piagetian tasks is an important concern if meaningful analysis of intellectual levels and intellectual gains are to be conducted using subjective Piagetian measures. In studies of training effects on cognitive growth, tasks involving flexible metal rods (Bredderman, 1973) and density conservation (Brainerd and Allen, 1971) were used as pre-post measures of the relative effects of training procedures on experimental subjects. Although experimental groups showed significant gains on the posttests in both of these studies, the control groups showed similar gains. It is suggested that in some cases gains attributed to training procedures may have been due to a learning effect of taking and retaking the same Piagetian tasks and not to the experimental treatment. To find out, Lawson, Nordland and Devito (1974) analyzed responses on five Piagetian formal operational tasks in a test-retest situation to determine the extent to which taking a pretest affected scores on posttests. No intervening treatment was given during the one week period between pre-post measures. Approximately 20% of the subjects are reported as having shifted from concrete to formal reasoning patterns on the basis of this analysis. Since only one week of time elapsed from pretesting to posttesting and no treatment was provided, the researchers claim that these apparent gains in intellectual ability were artificial.

Finally, much of the research seems to imply that language usage is a basic problem to overcome in designing evaluation instruments of this kind. On this subject one might consider the views of O'Brien (1971) and Jansson (1974), who provide some insight into the structural and linquistic variables that influence the ability of students in judging deductive arguments. The field is completely open-ended and probably beyond the range of this review. It may suffice to quote Vygotsky (1962) from one of his earlier works:

> As long as we do not understand the interrelation of thought and word, we cannot answer, or even correctly pose, any of the more specific questions in the area (p. 1).

At least careful sentence construction, with a minimum of complexity in vocabulary and phrasing structure is necessary when designing cognitive assessment instruments. Then perhaps the researcher has the right to pose the specific questions to which Vygotsky refers.

Summary

The usefulness of a reliable instrument to assess Piaget's logical operations because of the integral relationship between these operations and concept acquisition is evidenced by most of the literature. The design of instruments to assess a student's performance on Piaget's tasks is still in its infancy. His methods of experimentation are not totally standardized. However, it seems that there are three major dimensions that must be considered in the design of Piaget-type problems. They are content, the logical operation, and the context in which the problem is presented. It is hoped that this television based, cognitive level test will contribute additional knowledge toward the design of a widely acceptable logical operations test.

CHAPTER III

METHOD

Basic Design

The basic design of this validation study was diagrammed to include the steps and tasks involved in the construction of the Television Based Evaluation (TBE) and the Personal Interview Assessment (PIA).

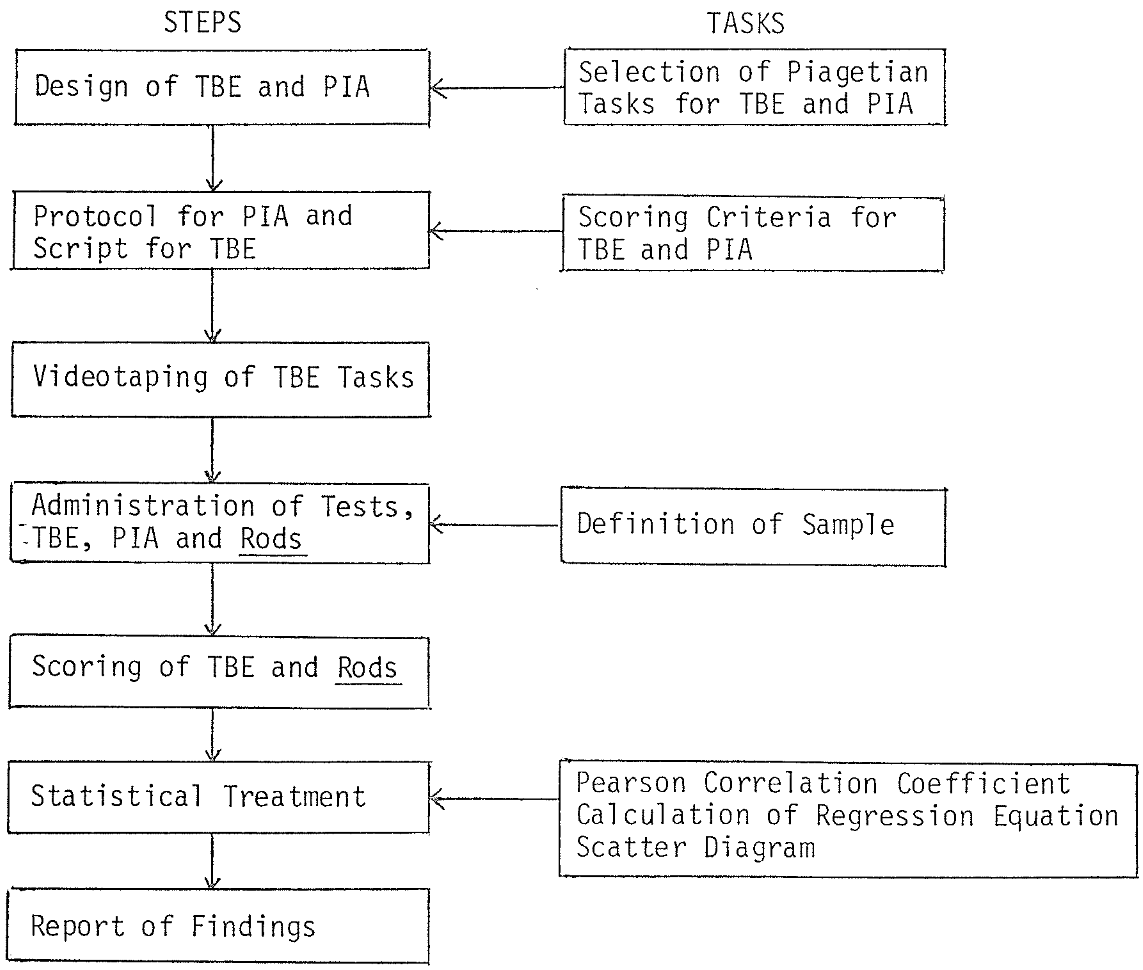


Figure 1. Diagram of Basic Design

Selection of Test Items for Television Based Evaluation Instrument (TBE)

The choice of evaluation items was consistent with the concrete operational and early formal operational stages. The concrete operational stage is characterized by the student's ability to serially order, conserve amount-length-area-weight-volume, perform 2-way classification, and make one-to-one correspondences. The early formal operational stage is characterized by the student's ability to conserve displacement volume, effectively eliminate contradictions, analyze combinatorially, and identify, separate and exclude inoperative variables.

Items selected for inclusion on the test, therefore, included Piagetian tasks with modification made to their content and context in some cases, but not the logical operation tested.

It should be noted that the two items involving the conservation of weight and the conservation of displacement volume are probably the best indicators of late concrete and early formal operational reasoning respectively (Lawson, Blake and Nordland, 1974; Nordland, Lawson and Kahle, 1974).

Construction of Television Based Evaluation (TBE)

The TBE contained 7 tasks in all. Each item was presented on the television screen, complete with questions to which the student responded in writing in a test booklet. The booklet contained the same questions heard and viewed by the student on the television monitor. The booklet also included spaces to check off answers and provide reasons for answers given. The videotape had blank spaces to provide the student with sufficient time to respond to each question, without turning off the VTR.

A brief description of the 7 task items follows:

Task 1: The Conservation of Amount

The students viewed two identical balls of clay on the television monitor. One of the balls was then rolled into a sausage shape and the students were asked about the relative amounts of clay in the two balls.

Task 2: The Conservation of Length

After seeing two cars beginning to travel on separate roadways (one straight, the other snake-like), the students were asked to compare the distance travelled by each car after a period of time if both cars travel at the same speed.

Task 3: The Conservation of Weight

The students were shown that two balls of clay weigh the same by placing them on opposite ends of an equal-arm balance. One of the balls was flattened into a "pancake" shape and the students were asked about the relative weights of the pieces.

Task 4: The Conservation of Area

Students were shown two sheets of paper (fields of grass) of equal size. A toy cow was placed on each sheet. Equal numbers of unit cubes (barns) were placed on each field; however, on one field they were grouped tightly together, while on the other field, the barns were separated and spread over the field randomly. The students were asked about the relative amounts of grass available for eating in the fields.

Task 5: The Conservation of Volume

Using two containers of different height, diameter, shape and volume, water was poured from one into the other. The students were asked to compare the amount of water before and after pouring.

Task 6: The Conservation of Displacement Volume

Using two solid metal cylinders of equal size but of different density, the students were shown the level of water displaced by the lighter cylinder and asked to predict the level of water displaced by the heavier cylinder. After making their predictions, the students were shown the actual outcome of displacing the heavier cylinder and asked to comment on the results.

Task 7: Separation and Control of Variables

The students observed a series of experiments involving bouncing balls of various sizes and materials, being dropped from various heights on different surfaces. They were asked to identify, separate, and control variables affecting the height of bounce in various situations.

It should be noted that reasons for stating a given answer were asked for on all task items. The reason given for an answer frequently became the best evidence for scoring the students' performance. Appendix A contains examples of test items from the test booklet.

Scoring Criteria for Items on the Television Based Evaluation

Points on a scale of one to four were awarded in each task according to the level of reasoning exhibited by the student on a given task. Tasks requiring concrete reasoning for successful completion by the student were assigned a maximum of two points. Tasks requiring formal reasoning for successful completion by the student were assigned three and four points, depending on the nature of the task involved.

Tasks 1-4: The Conservation of Amount-Length-Weight-Area

A correct prediction with an appropriate reason was awarded two points. An incorrect prediction was awarded one point.

Task 5: The Conservation of Volume

A correct prediction with the appropriate reason was awarded three points. A correct prediction with an inappropriate reason was awarded two points. An incorrect prediction was awarded one point.

Task 6: The Conservation of Displacement Volume

A correct prediction and explanation was awarded three points. An incorrect prediction with a correct explanation, following the student's viewing of the experiment on the monitor, received two points. An incorrect prediction with no explanation received one point.

Task 7: The Separation and Control of Variables

Correct responses on questions #1 and #2 with incorrect predictions on the remaining three items was awarded three points. Correct responses on questions #1 and #2 and all remaining three items was awarded four points. Incorrect responses on questions #1 and #2 and all the remaining three items was awarded two points.

Selection of Tasks for the Personal Interview Assessment (PIA)

The Piagetian tasks chosen for use in the PIA contained the same logical operations measured in the TBE; however, the content and the context of them were changed in most cases to avoid the possible training effects of repeating the same Piagetian tasks exactly.

The Personal Interview Assessment - Tasks and Scoring Criteria

Each interview included the same number and sequence of tasks. Scores were awarded on the basis of the difficulty level of the task, the quality of the students' verbal responses and the ability to exhibit the appropriate behaviors, ie., to control variables effectively, and to make correct predictions with appropriate reasons on the conservation tasks.

A brief description of the PIA tasks follows:

Task 1: The Conservation of Volume (Lawson, Nordland and Devito, 1974)

The student was shown that two balls of clay weigh the same by placing them on opposite sides of an equal-arm balance. One of the balls of clay was flattened into a "pancake" shape. The student was then asked to verify that two beakers contain the same amount of water. The student was then asked, concerning the clay pieces: "When the pieces of clay are placed in the water, will the ball make the water level rise more? Will the pancake make the water level rise more? Or, will they both make the water level rise the same? Why?". A correct prediction was awarded two points. An incorrect prediction was awarded one point.

Task 2: The Conservation of Displacement Volume (Lawson, Nordland and Devito, 1974)

Two metal cylinders of equal volume but different weight were handed to the student. The student was asked to comment on their relative size and weight. The lighter cylinder was then carefully lowered into one of two graduated cylinders which were partially filled with equal amounts of water. The rise in water level was noted and the student was asked: "Will the other cylinder push the water up more, less, or the same as the lighter cylinder? Why?". A correct prediction and explanation received three points. An incorrect prediction with a correct explanation following the experiment received two points.

Task 3: The Elimination of Contradictions (Inhelder and Piaget, 1958, Ch. 2)

A given number of disparate objects were presented to the student who was asked to classify them according to whether or not they float on water. Then (the classification completed) the subject was asked to explain the basis of his classification in general terms. Next, the subject experimented with a container filled with water. The student was continually confronted with contradictions (should they arise) in his reasoning and encouraged in this way to generalize his statements into a universal law to explain why all materials float or sink in water.

The student was unaware of contradictions in his explanations. The contradiction was often used to explain the event. He failed to generalize his findings into a law. This behavior was awarded one point. The student was awarded two points if he showed progress in his search for a single explanation, but was unable to eliminate or incorporate terms that he verbalized such as air spaces, surface tension, or weight and volume into a general law. The subject began to verbalize a relationship between the density (he may not use the word) of water and the density of the objects being tested. This behavior was awarded three points. The student spontaneously compared weights of materials to weights of an equal volume of water and incorporated all contradictions that arose into a universal law to explain whether or not objects float on water. This behavior was awarded four points.

Task 4: The Separation of Variables (Inhelder and Piaget, 1958, Chapter 3)

Eight metal rods, weights, and the means for adjusting the length of the rods were pointed out to the student. The student was then asked to experiment with the apparatus to determine what makes some rods bend more than others. After the exploration, the student was asked to name the factors and prove the role of each factor mentioned.

The student who simply classified the rods that bend the most and/or least into thicker, thinner, longer, shorter, square, or round was awarded one point. If the student was able to classify and compare two rods using logical multiplication, two points were awarded. Logical multiplication can be thought of as (thicker) x (longer) = (thinner) x (shorter). The student who demonstrated proof that one rod will not bend as much as another by holding shape, material, length, and weight (or any combination) constant when experimenting with the diameter, but failed in some cases, was awarded three points. If the student identified all the variables (weight, length, shape, material, and diameter) and systematically set about holding all but one constant and testing that one until all were tested, he was awarded four points.

Task 5: The Exclusion of Inoperative Variables (Inhelder and Piaget, 1958, Chapter 4)

The student was shown the apparatus using the following protocol: "This is a pendulum. It consists of a string suspended from a hook and weights which can be hooked on the end. Experiment with the pendulum to see if you can find what factors affect how long it takes the pendulum to swing back and forth. The factors which can be changed are weight, length of string, height of drop, and push. You are to determine which of these factors makes a difference in how long it takes the pendulum to swing one complete swing." The student was asked questions to analyze his understanding of the situation. Possible questions included: "What factors affect the time of the swing? Can you prove that the length does have an effect? Does the height of drop affect the swing?"

The student who could serial order (eg., the longer the string the slower the swings), but could not serial order the weights was awarded one point. If the student could serial order the weights but could not prove the effect of variables by varying one factor and holding all others constant, he was awarded two points. Three points were given if the student could prove the effect of at least one variable but was unable to carry out a valid and systematic test for all variables. If the subject could demonstrate proof of the effect of each variable holding all others constant, four points were awarded.

The Rods Test

Harris (1973) designed a test to identify students who were in the transition between concrete and formal operations. She developed a scoring guide which was used in this study for the purpose of classifying students into stages IIA, IIB, IIIA, IIIB (See appendix B for sample items).

Sample and Procedure

The TBE and the Rods test were administered to students enrolled in the ninth grade at a junior high school located in a middle to upper-middle class south Winnipeg suburban community. The PIA was administered to 30 randomly chosen students from this sample. About half of the subjects wrote the Rods test before the TBE. The clinical interviews took place over a period of two months after the completion of the two classroom tests. Each interview lasted about 45 minutes. The science classes in the school were heterogeneously grouped, with approximately an equal ratio of males to females.

It should be noted that scores from 8 students writing the TBE were rejected because these students were not present for the Rods test. This left 103 students who wrote both tests and were included in the statistical treatment. One class writing the TBE was interrupted for approximately five minutes because of a breakdown in the VTR equipment. This may have caused some loss of concentration in the students, and thereby altered their behavior with respect to the last two tasks on the test.

The scheduling of the TBE and the Rods test was such that a class wrote both tests during the same period on the same day of the week, one week apart. With the exception of one class, which had to be rescheduled from a morning period to a late afternoon one, two weeks apart, this procedure was followed.

Expected Results and Implications

Because the PIA probably provided the best operational definition for students' behavior with respect to the various Piagetian tasks, a high correlation coefficient between the PIA and the previously validated Rods test would strengthen the confidence one could place on the PIA. The small sample size, however, could influence the results considerably and a low correlation coefficient could reflect this.

A high correlation coefficient between the PIA and TBE could imply that the TBE may adequately discriminate between concrete and formal operational students and could replace the personal interview method of the PIA. Again, the small sample size for this comparison could influence the value obtained for the coefficient and the confidence that could be placed on it.

The Rods test was validated in another study and designed to discriminate between concrete and formal operational students. Because only 30 students of the 103 in the sample participated in the PIA, it was hoped that a high correlation coefficient between the scores on the Rods and the TBE would further support whatever positive, and direct relationship was found when the results of both the TBE and Rods test were compared to the PIA.

Finally, it was expected that an individual who was successful on TBE tasks requiring formal operations would also respond correctly to tasks requiring only concrete operations. This could imply that the TBE was consistent with Piaget's theory which suggests that the sequence in cognitive development is from the concrete stage to the formal stage.

Statistical Treatment

An individual's score on the PIA (with subsequent categorization into stages IIA, IIB, IIIA, IIIB) operationally defined his reasoning and cognitive level. The extent to which his score on the TBE and Rods test each predicted his score on the PIA, was estimated by means of the Pearson product-moment correlation coefficient, r_{xy} (Glass and Stanley, 1970, pp. 109-123), for the 30 students writing all three tests. A regression equation was calculated for scores from the Rods and PIA as well as from the TBE and PIA to provide a better estimate of the degree of relationship between these tests. Confidence limits were calculated for all correlation coefficients.

A correlation coefficient was calculated and a scatter diagram drawn for results from the 103 students writing both the TBE and Rods test. This could provide supporting evidence in the validation of TBE.

CHAPTER IV

ANALYSIS OF DATA

This chapter will report the results obtained from the data collected, organized, and analyzed in accordance with the basic design.

Analysis of Data

Task responses on the TBE and the PIA were categorized and points awarded as follows:

Stage IIA - early concrete operational stage - one point

Stage IIB - fully concrete operational stage - two points

Stage IIIA - early formal operational stage - three points

Stage IIIB - fully formal operational stage - four points

The categorization of students into stages was achieved on the TBE and the PIA by considering primarily the high scores achieved on the various test items. These high scores were considered to define the developmental level of the student with respect to each test.

The Pearson Correlation Coefficients, r_{xy}, were calculated from the data collected.

Table 1

The Pearson Correlation Coefficients, r_{xy}, from Students' Scores on the Television Based Evaluation (TBE), the *Rods* Test, and the Personal Interview Assessment (PIA)

	Rods	PIA
TBE	.83[a] *	.77[b] *
Rods	--	.65[c] *

[a] $n = 103$

[b] $n = 30$

[c] $n = 30$ $\quad$ * $p < .001$

The results show that there is a strong, direct relationship ($r = .77$) between an individual's score on the PIA and the TBE. That is to say, the PIA scores operationally define the students' behavior with respect to the Piagetian tasks used, and the TBE scores are shown to be good predictors of scores on the PIA. To this extent, the TBE is a valid instrument for measuring the cognitive level of grade nine students.

A regression equation was calculated for scores on the PIA against scores on the TBE and Rods test.

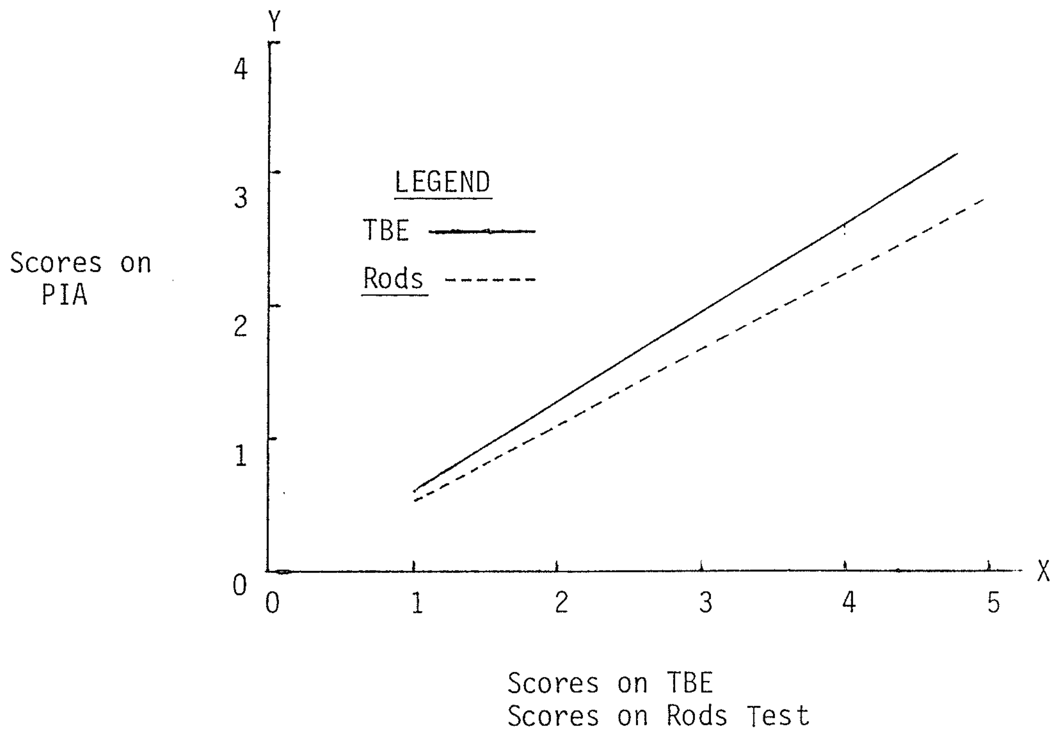


Figure 2. Regression line of Y (scores on PIA) on X (scores on TBE and scores on Rods test)

A comparison of the regression lines in Figure 2. further supports the validity of the TBE. The graph implies a similar relationship of both the TBE and Rods test to the PIA. The correlation coefficient ($r = .65$) between scores on the Rods test and the PIA, might show the extent to which the criteria for predicting cognitive level in Harris's

(1973) study and this study are the same. A higher correlation between these two tests was expected, given the fact that the Rods test was validated in a similar fashion, that is, against personal interview scores on Piagetian tasks. Perhaps the small sample size of 30 students may have been a factor.

Final evidence to support the claim that the TBE is a valid test of cognitive level appears in the correlation coefficient ($r = .83$) between scores on the TBE and the Rods test. The scatter diagram of these scores indicates a strong, positive correlation (Glass and Stanley, 1970, p. 117).

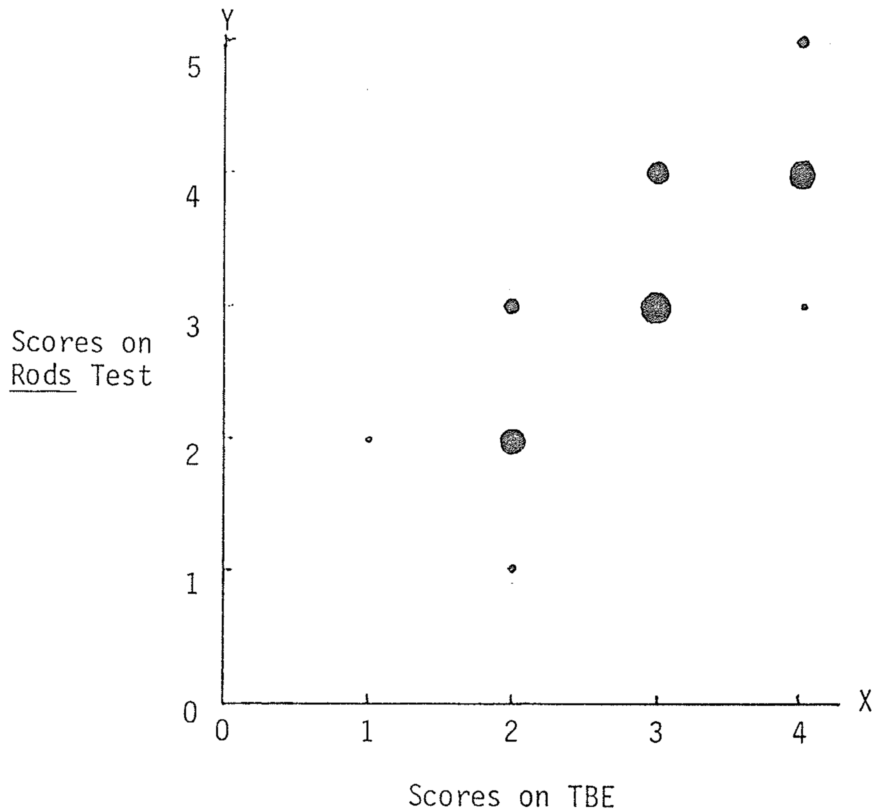


Figure 3. Scatter diagram of X (scores on TBE) and Y (scores on Rods test). Note: Size of points are proportional to the number of students with that score.

Categorization of Students into Developmental Levels

As a point of interest to other researchers in this area, the results from the PIA, the TBE, and the *Rods* test were categorized into their developmental levels as a percentage of the total number of students completing the tests (Table 2).

Table 2

Categorization of Students into Developmental Levels

	% Concrete	% Early Formal	% Formal
PIA[a]	33	47	20
TBE[b]	35	41	24
Rods[c]	30	35	35

[a] n = 30

[b] n = 103

[c] n = 103

Note: The developmental level "early formal" was named *transitional* on the *Rods* test.

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE STUDY

Summary and Conclusions

The science curriculum in our schools today contains a large variety of topics, most of which have never been critically analyzed with respect to their appropriateness to the developmental level of the receiving student. Furthermore, the developmental level of the student has not been measured at all grades. Even if we wanted to measure the developmental level of students immediately, a satisfactory testing instrument probably does not exist. The only information relating directly to potential for academic achievement that now exists about the individual student is his I.Q. and achievement level on aptitude tests. These may have little, if any value for matching science curriculum to students. Including information on cognitive growth levels for individual students could help science teachers plan programs more effectively, particularily the teaching of science concepts which require the students to perform various logical operations.

Piaget's theory of cognitive development was used as a basis for the construction and validation of a television based evaluation of cognitive development in this study. The focus of this evaluation instrument was on the concrete and formal stages of Piaget's theory. Individuals at the formal operational stage can deal with abstractions and are capable of hypothetico-deductive reasoning, even though unrelated to real events. The concrete operational thinker does not have the freedom of thought but rather, as the term implies, is much more restricted in his thinking.

The purpose of this study was to develop a tool for cognitive assessment that might assist the development of science curriculum in the schools throughout the province; it was not a test of Piaget's theory.

The results show that the cognitive level of large groups of students can be tested effectively through the medium of television. The television based test was validated against personal interviews; the results implied a strong positive relationship, evidenced by the correlation coefficients, regression lines and the scatter diagram.

The value of the television based test to education lies in its usefulness as a convenient, packaged, "teacher-proof" method for measuring cognitive level. The instrument could be expanded to include a greater number of test items in order to distinguish with greater certainty between the categories of concrete, transitional, and formal reasoning.

Recommendations for Future Study

1. The small sample involved in the personal interview assessment could be increased in a replication study involving a second group of grade nine students to increase the generalizability of results.
2. The television based assessment could be used to investigate the extent of the possible mismatch between the cognitive level of students and that inherent in the science curriculum content.
3. The television based test could be used to measure the effect of newly designed and implemented science curriculum.
4. A study to expand the television based test into two tests could prove useful for measuring developmental levels of students in studies where gains in cognitive development are being measured. For example, a study could use the television based test to determine whether a new teaching technique, designed to develop a

greater interest in science, facilitated cognitive growth.

5. A replication study could be carried out at the elementary or high school grades using Piagetian tasks appropriate to the grade level or ages of the students.

REFERENCES

References

Ankney, P. & Joyce L. The development of a Piagetian paper-and-pencil test for assessing concrete-operational reasoning (Doctoral Dissertation, University of Northern Colorado, 1974).

Ball, D. W. & Sayre, S. A. Relationships between student piagetian cognitive development and achievement in science (Doctoral dissertation, University of Northern Colorado, 1973).

Bart, W. M. The factor structure of formal operations. British Journal of Educational Psychology, 1971, 41, 70-79.

Brainerd, C. J. & Allen, T.W. Training and generalization of density conservation: effects of feedback and consecutive similar stimuli. Child Development, 1971, 42, 693-704.

Bredderman, T.A. The effects of training of the development of ability to control variables. Journal of Research in Science Teaching, 1973, 10(1), 189-200.

Bruner, J. On the conservation of liquids. In J. Bruner, R.R. Oliver and P. M. Greenfield, et al. Studies in Cognitive Growth, New York: Wiley, 1966.

Burney, G. M. The construction and validation of an objective formal reasoning instrument (Unpublished doctoral dissertation, University of Northern Colorado, 1974).

Coulter, D. M. Extent of discrimination between concrete and formal operational students as reflected by achievement on teacher-made tests of high school science. (Unpublished Masters Thesis, University of Manitoba, 1977).

Flavell, J. H. The developmental psychology of Jean Piaget. New York: D. Van Nostrand Company, 1963.

Glass, G. V. & Stanley, J. C. Statistical methods in education and psychology. New Jersey: Prentice-Hall, 1970.

Goldschmid, M. L. Different types of conservation and non-conservation and their relation to age, sex, IQ, MA, and vocabulary. Child Development, 1967, 38, 1229-1246.

Gray, W. M. Development of Piagetian-based written test: a criterion-referenced approach. A paper presented at the annual meeting of the American Educational Research Association, New Orleans, Feb., 1973.

Harris, A. A problem solving technique to facilitate concrete formal transition (Doctoral Dissertation, University of Calgary at Calgary, Alberta, 1973).

Hunt, D. E. Matching models in education. Toronto: Ontario Institute for Studies in Education, 1971.

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

Inhelder, B., Sinclair, H. & Bovet, M. Learning and cognitive development. Cambridge, Massachusettes: Harvard University Press, 1974.

Jansson, L. C. Structural and linguistic variables that contribute to difficulty in the judgement of simple verbal deductive arguments. Educational Studies in Mathematics, 1974, 5, 493-505.

Kaufman, B. A. & Konicek, R. D. The applicability of Piaget to contemporary curriculum reform? Revision of a paper presented at the Piaget symposium at William James College at Grand Valley State College, Allendale, Michigan, May, 1972.

Lawson, A. E. The development and validation of a classroom test of formal reasoning. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Cincinnati, March, 1977.

Lawson, A. E. & Blake, A. J. D. The factor structure of some Piagetian tasks. Journal of Research in Science Teaching, 1976, in press.

Lawson, A. E., Blake, A. J. D. & Nordland, F. H. Piagetian tasks clarified: the use of metal cylinders. The American Biology Teacher, 1974, 36(4), 209-211.

Lawson, A. E. & Devito, A. Piagetian formal operational tasks: a crossover study of learning effect and reliability. Science Education, 1974, 58(2), 267-276.

Lawson, A. E. & Renner, J. W. Relationship of science subject matter and developmental levels of learners. Journal of Research in Science Teaching, 1975, 12(4), 347-358.

Lovell, K. & Shields, J. B. Some aspects of a study of the gifted child. British Journal of Educational Psychology, 1967, 37, 201-208.

Mallon, E. J. Cognitive development and processes: review of the philosophy of Jean Piaget. The American Biology Teacher, January, 1976, 28-33.

Nordland, F. H., Lawson, A. E. & Kahle, J. B. A study of concrete and formal reasoning ability in disadvantaged junior and senior high school science students. Science Education, 1974, 58(4), 569-575.

O'Brien, T. C., Shapiro, B. J. & Reali, N. C. Logical thinking - language and context. Educational Studies in Mathematics, 1971, 4, 201-219.

Raven, R. J. Programming Piaget's logical operations for science inquiry and concept attainment. Journal of Research in Science Teaching, 1974, 11(3), 251-261.

Raven, R. J. The development of a test of Piaget's logical operations. Science Education, 1973, 57(3), 377-385.

Rowell, J. A. & Hofman, P. J. Group tests for distinguishing formal from concrete thinkers. Journal of Research in Science Teaching, 1975, 12(2), 157-164.

Shayer, M. & Wharry, D. Piaget in the classroom I: testing a whole class at the same time. Unpublished manuscript, Chelsia College, University of London, 1975.

Tisher, R. P. & Dale, L. G. Understanding in science test. Victoria: Australian Council for Educational Research, 1975.

Tomlinson-Keasey, C. Can we develop abstract thought in college freshman? Paper presented at the 84th Annual American Psychological Association Convention, Washington, D. C., September, 1975.

Vygotsky, L. S. Thought and language. Cambridge, Massachusetts: The M.I.T. Press, 1962.

Appendix A

Sample Test Items From The Television Based Evaluation Instrument

TASK 6:

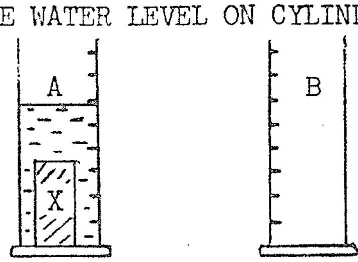
IS WEIGHT X (SMALLER THAN, LARGER THAN, OR THE SAME SIZE AS) WEIGHT Y?

_____Smaller _____Larger _____Same

DOES WEIGHT X WEIGH (LESS THAN, MORE THAN, OR THE SAME AS) WEIGHT Y?

_____Less _____More _____Same

MARK THE WATER LEVEL ON CYLINDER B.



GIVE A REASON FOR YOUR PREDICTION . . .

____

____

____

. . . (TASK 6 CONTINUED)

IS THE WATER LEVEL IN CYLINDER B (LOWER THAN, HIGHER THAN, OR THE SAME AS) THE WATER LEVEL IN CYLINDER A?

____Lower ____Higher ____Same

IS THE RESULT (THE SAME AS, OR DIFFERENT FROM) WHAT YOU PREDICTED IT WOULD BE?

____Same ____Different

HOW DO YOU EXPLAIN THIS RESULT?

____

____

____

____

____

TASK 7:

QUESTION 1: WHAT DO YOU THINK MADE ONE BALL BOUNCE HIGHER THAN THE OTHER?

QUESTION 2: THIS TIME, WHAT DO YOU THINK MADE ONE BALL BOUNCE HIGHER THAN THE OTHER?

. . . (TASK 7 CONTINUED)

PUT AN "X" BESIDE ALL THE ITEMS YOU WOULD USE TO PROVE THAT THE MATERIAL A BALL IS MADE OF MIGHT CAUSE ONE BALL TO BOUNCE HIGHER THAN ANOTHER.

SIZE OF BALL	SURFACE TO BOUNCE BALL ON	HEIGHT OF DROP	MATERIAL BALL IS MADE OF
____	wood ____	6 inches ____	rubber ____
____	cardboard ____	12 inches ____	wood ____
____	rug ____	24 inches ____	steel ____
____	sponge ____	36 inches ____	brass ____
____			styro-foam ____
____			

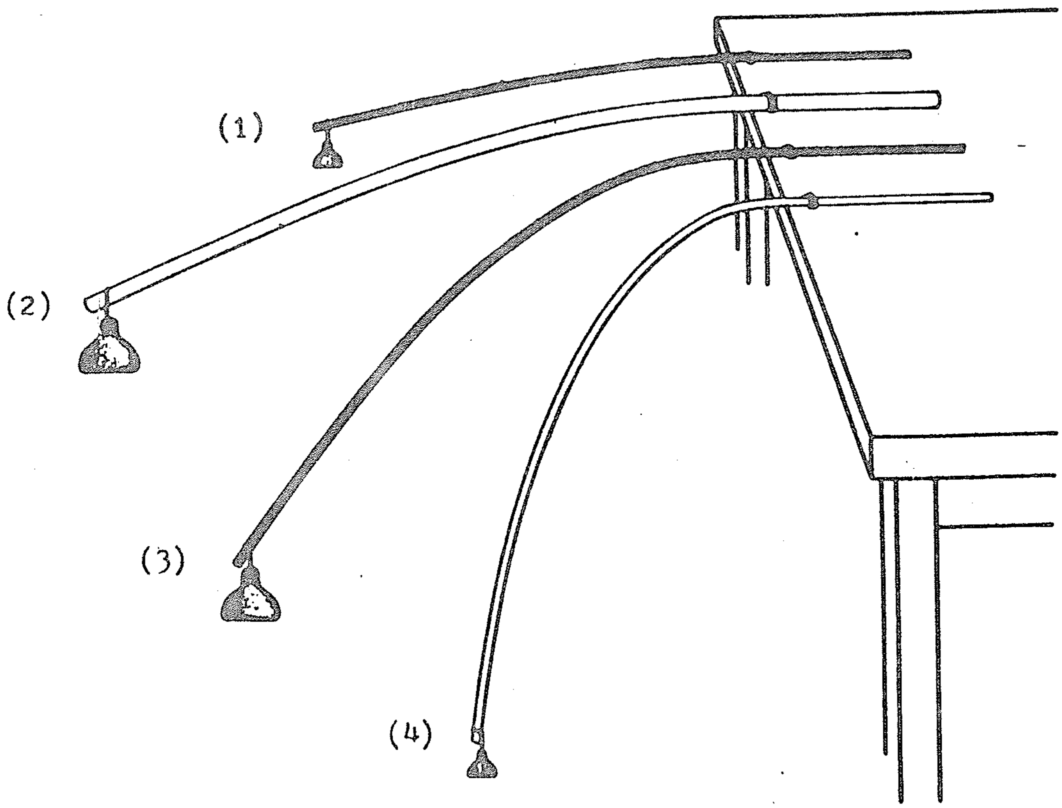
PUT AN "X" BESIDE ALL THE ITEMS YOU WOULD USE TO PROVE THAT THE HEIGHT FROM WHICH A BALL IS DROPPED MIGHT CAUSE ONE BALL TO BOUNCE HIGHER THAN ANOTHER.

SIZE OF BALL	SURFACE TO BOUNCE BALL ON	HEIGHT OF DROP	MATERIAL BALL IS MADE OF
____	wood ____	6 inches ____	rubber ____
____	card-board ____	12 inches ____	wood ____
____	rug ____	24 inches ____	steel ____
____	sponge ____	36 inches ____	brass ____
____			styro-foam ____
____			

Appendix B

1. Sample Questions From _Rods_ Test

2. Sample Of Scoring Guide For _Rods_ Test



QUESTION 2 Probably (3) bent more than (2) because

QUESTION 3 Probably (4) bent more than (1) because

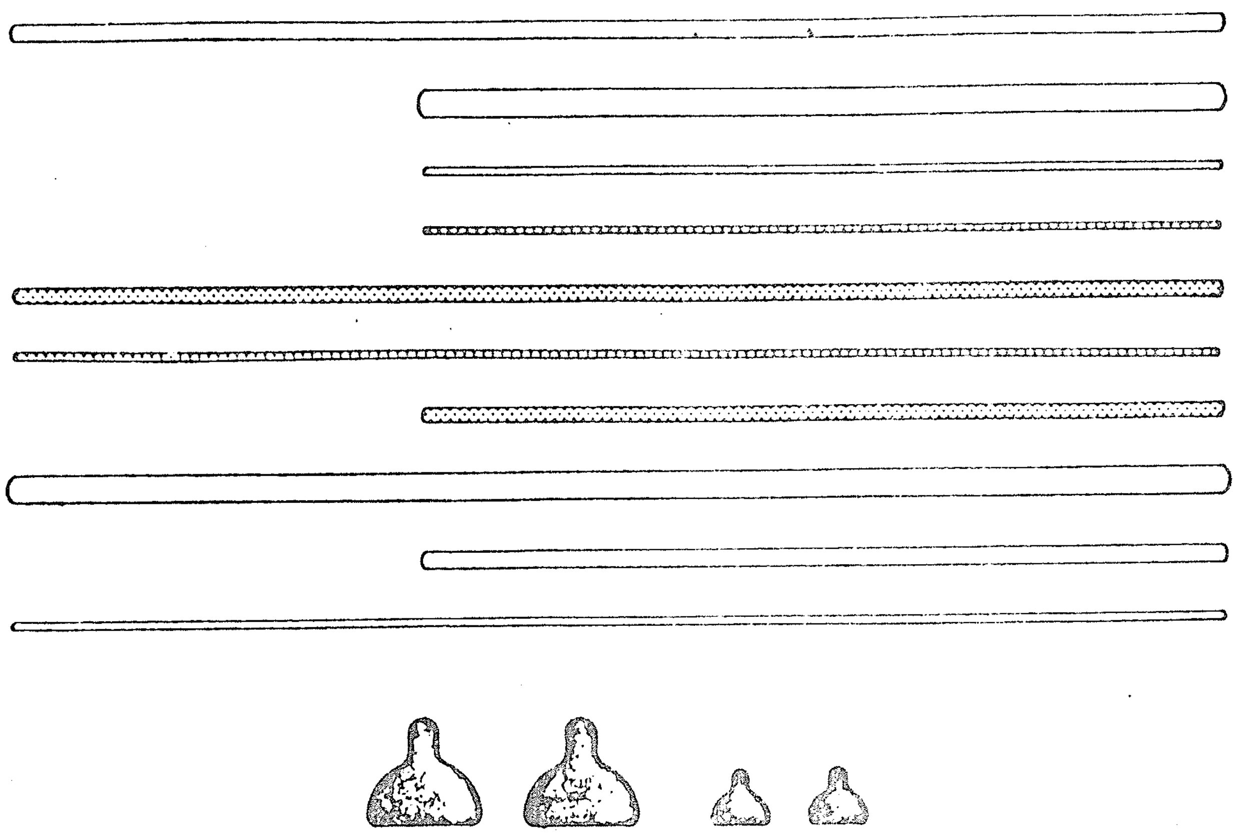
QUESTION 4 Probably (3) bent more than (1) because

QUESTION 5 Suppose you think that a difference in material (that is, metal or wood) might cause one rod to bend more than another.

Mark X's on the two rods you would use to prove that bending depends on the kind of material.

Mark X's under the two weights you would use.

(You should use four X's, two for rods and two for weights.)



Scoring Criteria for Rods

General Instructions

1. The scoring is related to the variables named in the response.
2. Ignore irrelevant terms in the answer as long as the correct variable is also included in the answer.

   Irrelevant terms include: lightness, heaviness, size, weaker, bigger, smaller, softer
3. For all questions width, diameter, thickness, radius, circumference are considered as synonymous terms

---

		Score
#1	weight of weights (weights, size of weights) width length (where the weight is placed) material rod is made of (type of rod, whether the rod is metal or wood)	
	response includes all four variables listed above	5
	response includes all four variables, but breaks them into attributes, for example "Whether the rod is thin or thick" or response includes irrelevant terms such as gravity, flexibility, air pressure	4
	response includes three variables listed above	3
	response includes one or two variables listed above	2
	no response or one which includes only irrelevant terms	1
#2	difference in material difference in thickness	
	response includes the two differences given above	3
	response includes only one of the differences	2
	any other answer	1