THE INFLUENCE OF TEACHER CONTROL AND PUPIL CHARACTERISTICS ON PROCESS ACHIEVEMENT IN ELEMENTARY SCIENCE CLASSES

HELEN GWENLLIAN BANFIELD

THE INFLUENCE OF TEACHER CONTROL AND PUPIL CHARACTERISTICS ON PROCESS ACHIEVEMENT IN ELEMENTARY SCIENCE CLASSES

by

Helen Gwenllian Banfield, B.Sc., PGCE.

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Education

Department of Education
Memorial University of Newfoundland
September 1978

St. John's Newfoundland

Abstract

Based on three processes of science, controlling variables, interpreting data and quantifying, two 20 item science process achievement tests were constructed. After establishing the reliability and validity of these tests they were used to investigate the influence of variations in teacher control and treatment sequence upon the process achievement of grade six students engaged in process oriented science activities involving the three processes above.

Two experimental treatments referred to as high and low control, distinguished by differences in the level of teacher control, were defined in terms of a series of teacher behaviors. A repeated measures design was employed with ten grade six classes being exposed to nine weeks of each treatment. Four teachers, replacing the regular teachers, were trained to distinguish treatments and were randomly assigned to classes. Monitoring of the treatments was carried out by videotaping each class on three occasions, coding pupil and teacher behavior each time.

An achievement post test was administered after each nine week round and analysed by round and also by treatment sequence using a generalised multiple regression procedure. After ability was accounted for a number of independent variables such as dependency, submissive dominance, sex, SES, extraversion, neuroticism, self concept and responsibility for success were investigated for possible interactions with the variations in teacher control.

No significant treatment effects were found in either analysis although a significant treatment-attitude interaction was found in the analysis by round while a significant treatment-teacher interaction was found in the analysis by treatment sequence. No other significant interactions were found.

Acknowledgements

This study was carried out as part of a project of The Institute for Educational Research and Development, Memorial University of Newfoundland and funded by Canada Council. This author wishes to acknowledge with grateful thanks the patience and continual guidance of Dr. R.K. Crocker, the director of the project and supervisor of this study. This research could not have been carried out without the facilities and personnel provided by the Institute.

Thanks are also extended to Dr. R.P. Amaria for her help and moral support throughout the study.

Finally the author wishes to thank Susan Sharpe for her competent typing of the manuscript.

TABLE OF CONTENTS

Chapter		Page

LIST OF TABLES

I THE PROBLEM

Introduction to the Problem

Many studies have been carried out to discover more about the actual verbal interaction that occurs in the classroom, both between teacher and pupil and between pupils themselves. An important aspect of this type of research has been the influence of variations in this verbal interaction upon pupil achievement and attitudes.

This particular study has been designed to examine the general problem of teacher control in elementary science classes. According to Morrison (1974) the concept of control incorporates the way in which the teacher structures the classroom and imposes limits on the freedom of movement or choice of her pupils. In this project the concept was operationalized in terms of a number of teacher behaviors designed to reflect high and low teacher structuring of pupil activity in laboratory type lessons.

The lessons used in this project were based on the process approach in science teaching and were modified from the Elementary Science Curriculum Project (Crocker, 1972).

The lessons involved three of the basic processes of science namely quantifying, controlling variables, and interpreting data as defined by the American Association for the Advancement of Science, Comission on Science Education (1965) (AAAS). During each lesson period the pupils worked in small groups on practical activities, each group having its own set of simple apparatus. This lesson format thus resulted in relatively more pupil-pupil and teacher-group interaction than would be found in more conventional classrooms.

All pupils were exposed to two experimental treatments, referred to as high control (HC) and low control (LC). Operational definitions were produced by modifying the Bellack classroom observation system (Bellack, Kliebard, Hyman and Smith, 1966) to give a number of teacher behaviors on which the two treatments could be distinguished.

The basic objective of the overall project was to investigate the influence of treatment, treatment sequence, and the interactions between treatments and pupil characteristics upon:

(a) pupil performance ie. process achievement

(b) preference of pupil for treatment

(c) pupil perceptions of treatment differences.

Since the influence of both treatment and the treatment-aptitude interactions upon process achievement was the responsibility of this investigator, the following more specific objectives were identified:

1. Development of a process achievement instrument
2. Use of this instrument in
   (a) the comparison of process achievement of pupils exposed to two different treatment strategies differing in the degree of teacher control,
   (b) the exploration of possible interactions between experimental treatments and pupil characteristics.

The objectives were expressed by the following questions:

1. Is there a significant difference between the students exposed to a high control or low control teaching style as measured by an investigator-designed "Science Process Achievement Test"?
2. Is there a significant interaction between treatment and each of the following variables:
   (a) Extraversion as measured by the Junior Eysenk Personality Inventory (JEPI) (Eysenk, S.B., 1963).
   (b) Neuroticism as measured by the JEPI.
   (c) Dependence Proneness as measured by Flanders and Amidon (1961).
   (d) Sex
   (e) Canadian Test of Basic Skills (modified Iowa Test of Basic Skills).
   (f) Non Verbal Intelligence as measured by the Raven Progressive Matrices (Raven, 1960).

(g) Submissive Dominance as measured by the submissive-dominant subscale of the Children's Personality Questionaire (Porter and Catell, 1968).

(h) SES as measured by the Socio-Economic Index for occupations in Canada (Blishen, 1967)?

3. Is there a difference in the significant interactions between treatment and pupil characteristics in the two treatment sequences i.e. high to low or low to high?

Rationale for the Study

Based on social systems theory (Miller and Rice, 1967), Morrison views the classroom as a complex social system having certain boundaries which exist at various discontinuities in space, time, or behavior. Transactions across these boundaries are controlled by the teacher. The transactions involved in this concept of control are those activities in a classroom, such as interaction between teacher and pupils or pupils with pupils, over which the teacher may have a great deal of control. High control of these boundaries exists in a classroom situation where transactions across these boundaries are highly regulated. Low control of boundaries on the other hand exists where less regulation of these transactions occurs.

This study thus attempts to investigate the area of teaching over which a teacher has most influence, namely the way she conducts her lesson within the constraints of curriculum and available materials. The experimental treatments involved are thus defined in terms of variations in such teacher behavior.

Barker (1968) maintains that the variations in teacher behavior under discussion may influence pupil perceptions of goals in the classroom situation, which in turn may influence perceived roles and the behavior corresponding to these roles. Thus such variation in teacher behavior and their influence on pupil behavior is an important avenue for investigation.

One cannot however ignore the possibility that the influences of variations in teacher behavior on pupil performance are modified by characteristics of the pupils themselves. This may be more formally stated by the concept of matching pupil characteristics with teaching style (Hunt, 1971; Hunt and Sullivan, 1974). This model is based on the classic formula of Lewin (1935), B = f(P.E) (behavior is a function of the person and the environment). This concept of matching models, along with Cronbach's (1957) recommendation to "coordinate individual differences with environmental influences", provides the basis for the aptitude-treatment interaction studies in research in teaching.

As process achievement was the dependent variable of interest to this investigator, it was necessary to obtain an instrument suitable for the sixth grade which tested the three processes of science under consideration. Despite the increasing number of process oriented elementary science programs, the number of available, time-efficient, reliable, and valid instruments is small, those available being restricted to particular grade levels or pupils familiar with special terminology. Thus it was necessary to construct a test suitable for this particular study. A major part of the work reported here involves the development of such a test.

II RELATED RESEARCH

Educational research into the teaching process itself now centers on actual classroom instruction. Interaction between the teacher and class during instruction is now being carefully analyzed as a preliminary step in the specification of conditions necessary to maximise learning in all situations. This analysis is being carried out by various classroom observation systems, the number of which has increased rapidly over the past 25 years.

Classroom Observation Studies

Since the early work of Anderson (1939) the problem of teacher control has been viewed from many perspectives. The Flanders Interaction Analysis Category System (FIAC) formed the basis of many studies which viewed the issue in terms of teacher "indirectness". Other large scale studies (e.g. Soar, 1973; Brophy and Evertson, 1974) have examined the problem with respect to control of classroom organisation, grouping and handling of materials.

From a rather different line of analysis, based on the "logic" of classroom discourse, another well known classroom observation system was developed by Bellack and his associates

(Bellack et al., 1966). This system was of considerable importance to this study as a whole since it incorporated both substantive and instructional dimensions and was amenable to modification to capture the concept of variation in teacher control as well as the nature of the science process activities under investigation. The major aim of the work of Bellack and his associates was to describe the pattern processes of verbal interaction involved in classroom activity and its relationship to subsequent pupil learning. It was felt by this group that because of the reciprocal nature of verbal activities, the role of the teacher could be adequately described only in relation to the students' role. The Bellack system proposes that the pedagogical moves of students and teachers can be grouped into four major categories, classified on the basis of their pedagogical function. This classification distinguishes between structuring, soliciting, responding and reacting moves each of which can be subdivided into substantive and instructional areas. Structuring moves set the context for subsequent behavior whereas soliciting moves are those intended to elicit a verbal, cognitive, or physical response, usually in an interrogative form. Responding moves fulfill the expectation of solicitations whereas reacting moves serve to modify or rate what was said in moves that occasioned them.

Earlier stages of the project, of which this study formed a part, involved considerable modification of the original Bellack system (Crocker et al., 1972). Although the pedagogical move and the division into substantive and instructional dimensions were retained, the revised system was designed to reflect more directly the dynamics of science activities.

The Question of Teacher Control

Although the problem of teacher control has been considered from many viewpoints, perhaps the most common approach has been that based on the Flanders distinction between teacher indirectness and directness. Dunkin and Biddle (1974) have carried out an extensive review of experimental and field studies on the influence of the direct and indirect teaching styles on achievement, but there appears to be no clear consensus as to the greater effectiveness of one style over the other. In many cases results from experimental studies appear to contradict those from field studies. Results such as these may be due to the variety of existing terms used to describe different teaching styles and the lack of an operational definition of indirectness. According to Soar (1968), even when the same observational system is used, different results may be obtained when scores on indirectness are calculated using different parts of the analysis matrices. More recently however, Gage (1976) has criticized the work of Dunkin and

Biddle for their tendency to examine studies individually rather than look at the direction of results overall. Gage (1976) taking twenty studies combined, found a highly significant relationship between teacher indirectness and pupil achievement.

Soar (1968) in a study of vocabulary growth and reading, concluded that the relationship between indirectness and achievement is non-linear, suggesting that there is an optimum level of indirectness which is not the same for all subject areas. He further suggests that these results are associated with levels of stress. Carline (1970) and Rian (1969) in experimental studies found that indirectness is unrelated to the achievement of average pupils. Rainey (1965) investigating achievement differences between pupils exposed to a directed versus a non-directed approach to laboratory exercises found no significant differences on written tests but significant differences did occur on a performance based laboratory test. Shymansky and Matthews (1974) in a study involving fifth grade science comparable to this study, found significant differences in achievement under two experimental learning situations referred to as "teacher-structured" and "student-structured". The higher scores in this study were found in the "student-structured" treatment. In a smaller study, along similar lines, Crocker, Bartlett, and Elliott (1976) found that students achieved significantly higher scores in "teacher-structured" situations when process achievement was used as the dependent variable. One of the reasons for the disparity between these results and those of the Shymansky and Matthews study may be the different lengths of time under investigation in these studies.

Rosenshine (1976) has recently proposed a "direct-instruction" model based on large scale correlation studies carried out on the relationships between classroom behavior and various outcomes. The results from these studies seem to suggest that direct teacher involvement in the instructional process leads to increased achievement in elementary mathematics and reading.

Aptitude-Treatment Interactions

The relationship between teaching style and pupil characteristics has been of great interest to many researchers in recent years, although the results do not seem to be conclusive in certain areas. Herman (1967) found that higher teacher indirectness is associated with higher IQ of the pupils, whereas Fisher (1970), in an extremely careful study, found higher teacher indirectness associated with greater achievement of pupils of "low achievement orientation". Shymansky and Matthews, in the study already cited, showed insignificant interactions between student ability and treatment.

Examination of the interaction between personality variables and teaching style has received less attention than the interaction between teaching style and ability. Amidon and Flanders (1961), investigating the relationship between dependence-proneness and teaching style with eighth-grade geometry students, showed that dependence-prone students are more successful with higher teacher indirectness. Crocker, Bartlett, and Elliott (1976) found a significant interaction between treatment and neuroticism in their study with sixth grade science pupils, although insignificant interactions were found in the case of extraversion and dependence proneness.

Other pupil characteristics, such as socio-economic status have been investigated in connection with teacher control. Brophy and Evertson (1974) report significant interactions. Pupils in low SES schools benefit more from higher control whereas those in higher SES schools derive more benefit from less control.

Interactions between pupil types and classroom types have been examined by Solomon and Kendall (1976) in a most extensive study. This study involved factor analysis to group pupil characteristics and classroom characteristics into clusters. A summary of the complex results of this study indicate that relatively unmotivated, low achieving pupils tend to perform better in permissive varied classrooms and those which combined warmth with an emphasis on expressiveness. High achievement, highly motivated pupils performed better under high control and moderate to high opportunity for student initiation. Non-compliant, highly self-directed pupils give mixed results.

Good et al. (1973) have suggested that the direction of causation in studies on teacher control should be reversed to permit investigation of the influence of students on the teacher. Their study found evidence consistent with the idea that lower pupil achievement gives rise to greater directness on the part of the teacher, than higher pupil achievement.

Thus the evidence for a causal link between indirectness and pupil achievement is not a strong one. Perhaps as Soar (1968) proposes, the relationship is curvilinear; a suggestion that could account for some of the contradictory results. Other suggestions to account for the lack of consensus may be incomparable concepts and methodology. In the present study the use of the modified Bellack system to operationalize the concept of control is intended to remove the latter problem of incomparable concepts.

Science Process Tests

In recent years, although there has been an increasing number of elementary science programs designed to develop the ability of the pupils to carry out the processes of science, the number of available time efficient, reliable, and valid instruments for evaluation of this ability is small. The Individual Competency Measures of Science-A Process Approach (SAPA) (AAAS, 1967) is one of the better known tests that claim to measure the processes of science. In these tests the child must actually perform the process in question in the presence of the tester. Some of the problems associated with these tests, however, are that they are time consuming, they are restricted to students familiar with the SAPA terminology, and they require an experienced examiner.

Welch and Pella (1967) developed the Science Process Inventory (SPI), an instrument designed to measure secondary school pupils' knowledge of the assumptions, activities, products, and ethics of science. The test is presented to the pupils as items to which the student agrees or disagrees.

Beard (1971) developed the "Basic Science Process Test" for use with primary grade children with limited reading and writing skills. The test covers the processes of classifying and measuring to accompany parts A,B, and C of SAPA and is presented to the pupils as a series of 35mm colored slides illustrating laboratory situations involving basic science processes.

Tannenbaum (1971) developed the "Test of Science Processes", an instrument measuring process achievement of eight processes for use with junior high pupils.

The limitations of the tests considered so far, with the exception of the Individual Competency Measures, is that they lack validation against the actual behaviors that would indicate that the student can actually carry out the processes in question.

Macleod, Berkheimer, Fyffe and Robison (1975) consider that the validity of a test is of prime importance. These researchers have developed a pool of criterion-related test items for the four integrated science processes. In contrast to the previously mentioned tests, which rely on face validity only, the criterion-related items exhibit both concurrent and face validity. The criterion used to establish the latter type of validity is the child's performance on the Individual Competency Measure of SAPA.

Summary

Of the various classroom observation systems available, the Bellack system was of most relevance to this study as after modification it was well suited to the operationalization of the treatments and the activities involved. Four major categories were used to classify the pedagogical moves, each category being subdivided into substantive and instructional areas.

A large number of studies have been carried out based on the notion of teacher indirectness. When these are examined individually there appears to be no clear consensus as to the greater effectiveness of one style over the other with regard to achievement. If however the results are examined overall there appears to be a significant relationship between indirectness and pupil achievement. With regard to interactions between pupil characteristics and treatments, the results also vary.

The Science Process Achievement Test was constructed because, of those tests available, they were unsuitable either for this grade level or this material.

III METHODOLOGY

Experimental Treatments

The concept of control has already been discussed generally, but in the context of this investigation operational definitions of the two experimental treatments, high and low control, were developed using a number of teacher behaviors derived from a modified version of the Bellack system (Bellack et al., 1966).

In a high control situation the pupil is provided with detailed instructions as to how to use the apparatus, how to solve the problem, and his attention is focused upon important aspects of the activity. In low control however, the student is given the problem, the necessary apparatus, but no indication as to how to proceed or manipulate the apparatus.

According to the system devised by Crocker (Crocker et al., 1975) the treatments are distinguished by different values of the following variables:

(a) High (HC)/Low (LC) ratio of teacher-class to teacher-group talk

(b) High/Low ratio of teacher to pupil talk,

i. during teacher-class interaction

ii. during teacher-group interaction

(c) High/Low number of structuring moves

(d) High/Low number of reacting moves (ratings)

(e) High/Low ratio of command to request soliciting moves

(f) High/Low number of request-give or request-request sequences

(g) Short/Long wait time for students' response to teacher's question or teacher asking another question

(h) Low/High level of pupil reporting individual results

(i) High/Low ratio of teacher interpreting data and controlling variables to student doing same

(j) Low/High ratio of soliciting to responding moves for the teacher

(k) Clear/unclear reference to expected or desired outcomes.

Teacher implementation of high and low control treatments was monitored by coding the required behavior from videotapes of 13 sample lessons, randomly chosen, approximately 130 minutes lesson time for each treatment.

Subjects

The experimental sample was obtained from grade six classes in the Avalon Consolidated School District in St. John's, Newfoundland. Originally the sample consisted of eleven classes in six schools, with class size ranging from thirty to thirty-nine pupils. One of the classes however was found to have already studied some of the material involved in the program. As a result of this it was decided to drop this class from the analysis but keep it available for use in reliability studies on several of the instruments.

The samples were selected from a region with a population of about 250,000 half of which live in the St. John's metropolitan area and the remaining half in a large number of smaller communities with populations ranging from below 1,000 to 5,000.

A wide range of occupations exists despite the fact that the area is not heavily industrialized. The urban area is mainly a government and university center with little primary industry. Outside the urban area a large proportion of the work force is engaged in primary activity although the numbers commuting daily into the city is not insignificant. Thus a wide range of social and occupational classes exists throughout the region.

The schools in the study were generally large, ranging in type from older, central city schools to new suburban schools with open areas and resource centers. The larger classes tended to be in the suburban schools while declining enrollments had resulted in smaller classes in inner city schools.

The Curriculum

The science program used in this study was based on Science - A Process Approach (AAAS, 1968), with the processes of interpreting data and controlling variables characteristic of the grade six level. Since the pupils in the experimental sample had not been exposed to this type of program, the curriculum was modified somewhat to facilitate the treatments. Since it could not be assumed that the pupils possessed the necessary quantifying skills to carry out the two processes of controlling variables and interpreting data, a unit on measurement was included.

The program was made up of a series of activities designed to occupy one or two class sessions. Sets of apparatus were provided so that pupils were able to work in groups of two. The general pattern for each activity included an introduction by the teacher, a period of data collection in which pupils would set up apparatus and make measurements, and a class discussion.

Description of the Activities

The activities were grouped into three content units namely Batteries and Bulbs, Human Reactions and Mechanics. The first unit involved the factors influencing the brightness of bulbs in a circuit and the strength of electromagnets. The second unit included activities based on factors influencing optical illusions, pulse rates and rates of learning and forgetting. Lastly the Mechanics unit consisted of the use of a balance and factors influencing the stretching of rubber bands.

Experimental Treatments

Four experienced elementary teachers were chosen and provided with a one week training program on lesson material and treatment differences. Apart from one who acted as coordinator, the teachers were employed on a part time basis and could be assigned as needed by the experiment. All four teachers replaced the regular grade six teachers for all science classes for the duration of the experiment. The advantage of being able to assign these teachers to classes in a relatively balanced way offsets any loss of generality arising from replacing the regular classroom teachers.

A repeated measures design was chosen due to the relatively small number of classes in the experimental sample. Each class was thus subjected to both treatments. In the first nine weeks of the experimental period five classes were exposed to low control and six to high control. In the second nine weeks of the study the teachers were reassigned and the treatments reversed. In each round, each teacher was assigned to at least one class in both treatments. Table 1 summarizes the experimental design.

TABLE I

Experimental Design

	TREATMENT			
	HIGH CONTROL		LOW CONTROL	
	TEACHER	CLASS	TEACHER	CLASS
ROUND 1	4	1	4	2
WEEKS 1-9	4	4*	2	3
	1	8	2	5*
	2	10	1	6
			1	7*
			3	9*
ROUND 2	2	2	2	1
WEEKS 10-18	4	3	3	4*
	3	5*	1	8
	1	6	4	10
	4	7*		
	1	9*		

Note * Classes that were not included in the final analysis in order to balance teachers and classes in each treatment.

Removal of the eleventh class from the analysis accounts for the imbalance in the number of classes in the two treatments. This design allowed the treatment main effects and treatment/pupil characteristics interactions to be analyzed without confounding between teacher and class. In each round there remained the possibility of confounding between treatment and class and treatment-teacher interactions, although treating the two rounds as replications with re-assignment of teachers tended to minimise this problem. In addition this design allows analysis of sequence effects. Throughout the analysis shift one refers to pupils who had high control followed by low control and shift two to those who experienced low control first.

Data Sources

Table 2 summarises the independent and dependent variables of interest to this investigation. The table also includes the source of the test and its relaibility coefficient where appropriate. The information on pupil characteristics were obtained at intervals throughout the experimental period, each instrument being administered to each class at the same time. The Canadian Test of Basic Skills was obtained from the school board which was carrying out the testing during the time of the study.

Table 2

Summary of Dependent and Independent

Variables and Instruments

Variable	Source	Reliability Coefficient	Reliability Method
Treatment	Coded data from video-tapes	.50 - .99	Inter-rater
Sex	School Records	-	-
SES	Blishen Scale	-	-
Non-verbal IQ	Raven Progressive Matrices	.83 - .97	Test-retest
Basic Skills (composite)	Canadian Test of Basic Skills	.87 - .95 (on several subtests)	Split-half
Reading Achievement	Gates-Maginitie Vocab. and Comprehension	.89 - .95 Survey D	Split-half
Extraversion	Junior Eysenk Personality Inventory	.70 - .80	Split-half and test-retest
Neuroticism	Junior Eysenk Personality Inventory	.70 - .80	Split-half and test-retest

Table 2 (continued)

Summary of Dependent and Independent

Variables and Instruments

Variable	Source	Reliability Coefficient	Reliability Method
Submissive-dominance	Children's Personality Questionaire	.74 - .82	Split-half and parallel forms
Dependency-proneness	Amidon and Flanders (1961)	.68 (original) .91 (Elliott)	Test-retest
Self concept	Modified Brookover	.72	Test-retest
Locus of Control	Intellectual Achievement Responsibility	.60 - .74	Test-retest and Split-half
*Attitude to School	Semantic Differential	.83	Test-retest
Science Process Achievement	Investigator designed	.74 - .82	Parallel form test-retest
*Attitude to Science	Semantic Differential	.74	Test-retest

* Developed especially for this project

Dependence Proneness Test (Amidon and Flanders 1961). This test includes items that describe students who comply with adults and conform to group pressures. The 45 agree-disagree item scale has been derived from an initial 150 item pool on the basis of item analysis. Local test-retest reliability data (Elliott, 1973) yielded a coefficient of .93.

Blishen Socio-economic Index for Occupations in Canada. This scale developed by Blishen (1967, 1969) uses a function of the distribution of education and income to rank 320 occupations in Canada. For purposes of this study, information was obtained from school records on the father's occupation for each pupil. Occupations were then grouped into seven categories on the basis of the Blishen Scale. These seven categories were used as attribute scores in the analysis.

Junior Eysenk Personality Inventory (Eysenk, S.B., 1963). The two major personality variables of neuroticism or anxiety and extraversion-introversion in children are measured by this 60 item inventory which is an extension of the Mausley Personality Inventories. Reliability for the extraversion scale tends to increase with age more so than the Neuroticism Scale. Validation of the JEPI has been carried out against rating of extraverted or introverted symptoms of 229 guidance clinic subjects. It was found that there was a significant difference in extraversion as measured by the inventory, between children showing extraverted and introverted symptoms. Reliability coefficients for this inventory range from .70 to .80.

Canadian Test of Basic Skills. This is a modification of the Iowa Test of Basic Skills and measures general intellectual skills in vocabulary, reading, language, work habits and mathematical skills. Split half reliability data reported for this battery indicate coefficients in the range of .87 to .97.

Raven's Progressive Matrices. These matrices provide a measure of non-verbal intelligence. The reliability data available provides a range of coefficients of .83 to .93, values depending on the age of the subject. This non-verbal intelligence measure was preferred to a more comprehensive test as other measures were available for verbal intelligence and non-verbal intelligence would seem to be important in the light of the type of activities involved in the lessons.

Submissive-Dominance. This instrument is the submissive dominance subscale of the Children's Personality Questionaire (Porter and Catell, 1968). In this study the subscale of the questionaire was used to examine whether dominance of activity is related to achievement. Both test-retest and parallel form reliability data are available with coefficients ranging from .74 to .82 reported for the E factor which contains the submissive-dominant subscale.

Self Concept of Ability. The Brookover (1962) questionaire formed the basis for this scale. This scale consisted of six five-choice items on the aspect of general self concept of ability. A test-retest reliability coefficient of .72 has been established.

Intellectual Achievement Responsibility (IAR). Assessment of pupils' beliefs concerning internal versus external responsibility for the intellectual academic achievement is the aim of this test developed by Crandall, Katovsky, and Crandall (1965). This variable was included since it was anticipated that individuals who tended to take responsibility for their own success and failure would obtain higher achievement scores in the low control situation.

Attitudes to School and Science (ATSCH, ATSCI). A semantic differential with all evaluative scales was used to measure pupils' attitudes to the concepts school and science. Test-retest reliabilities for the school and science subtests were .83 and .74 respectively.

Data Analysis. In studies such as this using intact classes, a major problem to be considered is the choice of class or individual as a unit of analysis. According to Cronbach (1976) the fact that observations made on individual members of a class are not independent leads on to consider the class as a suitable unit. The use of the class as a unit however results in considerable loss of degrees of freedom. In this study the individual was chosen as the unit of analysis for several reasons. Since most of the work was done in groups, it was felt that any class effects present were not as great as usual. Furthermore the fact that project teachers, teaching both treatments, replaced the regular teachers, reduced any special class-teacher effect and balanced out any unique teacher effects over both treatments.

With a small sample of classes such as this there could be no assurance that differences in initial ability would balance out despite the random assignment to treatment. Taking this fact into consideration, adjustments for initial ability were incorporated into the analysis for achievement.

The basic mode of analysis was a generalized regression analysis applied by treatment round and treatment sequence (shift). The analysis by round was a repeated treatments design in which all classes were exposed to both treatments. Considering the two rounds as replications, with reassignment of teachers, had the advantage of reducing the confounding of treatment-class and treatment-teacher effects within a round. The second round could not be considered a complete replication of round one since there exists the possibility of carry over effects from round one. This problem was reduced by the fact that the achievement tests were 75% content specific for each round.

The analysis by shift was carried out to investigate whether there existed any differences in the significant interactions between treatment and pupil characteristics when the two shifts, high-low and low-high are considered. An advantage in this analysis is that each pupil acts as his own control.

The advantages of the generalized regression method over analysis of variance are that the pupil characteristics do not have to be blocked, thus avoiding loss of information, and the order in which the variables are to be introduced into the equation can be chosen beforehand. In this study the variables were entered in the following order:

(a) initial ability (IQ and CTBS entered simultaneously),

(b) teachers,

(c) treatment-teacher interactions,

(d) stepwise entering of ATI'S,

(e) treatment.

The reason for the position of the treatment was based upon the argument that it is more important in this type of study to examine treatment aptitude interactions than treatment effects alone. Also, pupil characteristics and teacher effects must be regarded as prior conditions in any experimental classroom study. Treatment effects, if they occur are expected to occur after these prior conditions have been accounted for.

IV SCIENCE PROCESS ACHIEVEMENT TEST

Processes of Science

The process achievement test, the special responsibility of this author, consisted of three 20 item multiple choice tests for use as Pretest, Post Test 1 and Post Test 2. Each test was restricted to the three processes of science considered in the study, namely controlling variables, quantifying and interpreting data as set out by the AAAS model and modified by Crocker (1977). According to this model these processes have been defined as:

1. measurement: comparison of objects or events and with agreed upon standards. This involves the use of numbers, measurement and use of spatial relationships.
2. controlling variables: the process of deciding what variables might influence the outcome of a particular experiment and of holding all variables constant except those the investigator wishes to manipulate. The reproducibility of the experimental results depends on the ability to control the proper variables.

3. interpreting data: the technique of getting the most out of data without over generalizing and without loss of information inherent in the data. Devices such as graphs, tables, maps etc. which are useful in communications are also means of interpreting data.

Construction of items

Tables 3, 4, 5, provide a detailed plan of the objectives and units which formed the basis of the eighteen week teaching period. From this detailed list the following overall plan was prepared to ensure fair representation of the objectives on the tests, keeping in mind the kinds of activities that were carried out in each subject area.

Overall Plan of each Test

Unit	Inter. Data	Contr. Var.	Meas.	Total
1				15
2				
Transfer				5
Total	8	6	6	20

The transfer items included in the tests were those whose content was not specific to the topics included in the units covered. Since interpreting data was given

Table 3

Summary of Objectives and Activities

for the Process of Measurement

Objective	Batteries/Bulbs	Mechanics	Human Reactions
1. Problems of Estimation	Bulb brightness	----	
2. Use of arbitrary units	Strength of an electromagnet	weighing objects using a balance	----
3. Metric units		----	Length of optical lines
4. Subunits	----	weighing objects using a balance	----
5. Conversion	----	----	----
6. Instruments	Brightness meter	Balances	----
7. Averaging errors	Strength of an electromagnet	----	Individual differences in learning and forgetting
8. Judgements on precision	Effect of battery condition on brightness	Stretching rubber bands	Change of pulse rates during counting

Note. Transfer items are based on the same process objectives

Table 4

Summary of Objectives and Activities

for the Process of Controlling Variables

Objective	Batteries/Bulbs	Mechanics	Human Reactions
1. Identifying relevant variables	Variables infl. bulb brightness, Variables infl. strength of e/magnet	Rolling cylinders	Variables infl. pulse rate; Variables infl. rate of learning and forgetting
2. Identifying constant and manipulated variables	Manipulating number of batteries and number of turns of e/magnet	Selection of cylinders to attain a single manipulated variable	Manipulate learning time, hold length of list constant
3. Identifying responding variables	Bulb brightness, e/magnet strength	Relative time of rolling	Interval of recall
4. Identifying variables that affect respond. variable	time of day, temp. number of batteries	Preliminary check of rolling cylinder	Variables in learning
5. Identifying uncontrolled variables which exert syst. effects on responding variable	Classroom temp. turn spacing, number of batteries	Individual variations in rubber bands	Individual differences, external history, classroom conditions
6. Conduct of invest. with one manipulated & one respond. variable with other variables constant	Brightness of bulb as a function of number of batteries or number of bulbs	-----	Pulse rate as a a function of exercise

Table 5

Summary of Objectives and Activities

for the Process of Interpreting Data

Objective	Batteries/Bulbs	Mechanics	Human Reactions
1. Calculation of means 2. Calculation of range	strength of e/magnet	strength of rubber bands	Optical illusions, pulse rates
3. Construction of data tables	strength of e/magnet, brightness of bulbs	-----	Pulse rate as function of exercise
4. Constructing graphs, (histograms)	Brightness of bulbs, strength of e/magnet	Pendulum, length/ weight of balance	-----
5. Interpolation of graphs, (prediction)	-----	Stretch of rubber bands	-----
6. Averaging errors	-----	Balance, pendulum	
7. Extrapolation prediction	-----	-----	-----
8. Limits	Limit of strength limit of brightness	Non lin-earity of length - time relationship	Limits of pulse rates
9. Law-like relationships	-----	Law of lever	-----

considerable emphasis throughout the lessons it was decided to include relatively more items reflecting this process, thus the eight, six, six proportion in both Post Test 1 and 2. Because of the process orientation of the curriculum, for each test the ratio of content specific to transfer items was set at 3:1.

Using the "AAAS Science Process Instrument", "AAAS Individual Competency Measures", Criterion Referenced Items (Robison, 1973) and the "Test of Science Processes" (Tannenbaum, 1973) as guidelines, two parallel, 20 item tests were constructed for each of Pretest, Post Test 1 and Post Test 2 and the two halves of each were recombined to give four alternative forms distributed randomly in each class to remove any bias due to the shortness of the test. The items included in the Pretest were not content specific. Post Test 1, given at the end of the first round of the study covered the "Batteries and Bulbs" unit and one third of the "Human Reactions" unit whereas Post Test 2, administered at the end of the study after completion of Round 2, covered the remainder of the "Human Reactions" unit and the "Mechanics unit.

Prior to use the investigator gave each test to ten grade six pupils outside the project. Each student was questioned as to the appropriateness of the reading level.

In addition each test was carefully scrutinized by the Director of the project, a science educator and author of an elementary process science curriculum currently in use in the province.

Responses to the parallel forms were subjected to Item Analysis. Although this procedure was carried out after the tests were administered and thus could not lead to replacement of items, it did provide some valuable information on the quality of the test which could be used in any future modifications. The results of the Item Analysis provided the following information:

(a) test score

(b) mean score for each alternative form of the test

(c) standard deviation for each alternative form of the test

(d) index of item difficulty for each question

(e) index of item discrimination for each question.

Tables 6-13 give the results of the item analysis for each of the four combinations of parallel forms A and B for each of Post Test 1 and Post Test 2. The item analysis for the Pretest are not included since it was not used in the final analysis. Table 14 gives the means of the alternative forms of both post tests. Because the means of Post Test 2 were lower than those of Post Test 1 it was felt that Post Test 2 might have been more difficult. As a result it was

Table 6

Item Analysis for Post Test 1: Form 1 (QA1-10, B11-20)

Question Number	K.	Item Difficulty	Item Responses (Quartiles)	1	2	3	4	Item Discrimination
1.	4	0.689	Top	0	1	1	23	0.19
			Bottom	2	2	2	9	
2.	2	0.340	Top	4	22	0	1	0.50
			Bottom	8	4	2	3	
3.	3	0.717	Top	0	0	26	1	0.29
			Bottom	3	1	4	6	
4.	3	0.792	Top	0	2	24	1	0.20
			Bottom	1	3	7	4	
5.	2	0.792	Top	0	24	3	0	0.18
			Bottom	3	9	4	0	
6.	2	0.198	Top	0	8	17	2	0.29
			Bottom	3	2	10	2	
7.	3	0.519	Top	2	0	18	7	0.28
			Bottom	4	3	3	6	
8	4	0.670	Top	1	1	1	24	0.29
			Bottom	6	3	3	4	
9.	2	0.434	Top	0	20	2	4	0.32
			Bottom	4	4	5	4	
10.	4	0.368	Top	1	0	12	14	0.31
			Bottom	2	3	10	2	
11.	1	0.689	Top	25	0	1	1	0.27
			Bottom	5	4	4	4	
12.	1	0.670	Top	23	1	3	0	0.20
			Bottom	9	2	4	1	
13.	3	0.434	Top	3	1	16	6	0.20
			Bottom	3	4	7	3	
14.	2	0.406	Top	3	18	0	6	0.33
			Bottom	3	4	2	8	
15.	2	0.302	Top	11	11	0	4	0.25
			Bottom	6	3	5	3	
16.	2	0.264	Top	5	11	8	2	0.29
			Bottom	5	3	4	5	
17.	2	0.472	Top	5	17	3	2	0.27
			Bottom	6	5	4	2	
18.	2	0.343	Top	0	15	11	1	0.31
			Bottom	4	4	5	3	
19.	3	0.790	Top	1	1	25	0	0.21
			Bottom	1	2	8	6	
20.	4	0.712	Top	1	0	3	23	0.23
			Bottom	6	3	3	5	

Table 7

Item Analysis for Post Test 1: Form 2 (QB1-20)

Question Number	K.	Item Difficulty	Item Responses (Quartiles)					Item Discrimination
				1	2	3	4	
1.	3	0.731	Top	1	0	19	2	0.29
			Bottom	1	2	5	0	
2.	2	0.642	Top	3	17	0	0	0.34
			Bottom	2	2	2	2	
3.	1	0.657	Top	16	0	3	1	0.32
			Bottom	2	1	4	1	
4.	3	0.642	Top	2	0	17	1	0.30
			Bottom	0	1	4	3	
5.	4	0.418	Top	1	5	2	12	0.32
			Bottom	1	1	3	3	
6.	2	0.642	Top	2	17	1	0	0.33
			Bottom	2	3	3	1	
7.	3	0.478	Top	2	5	14	4	0.38
			Bottom	3	3	2	1	
8.	4	0.433	Top	1	2	2	15	0.49
			Bottom	3	2	3	1	
9.	3	0.582	Top	0	1	18	1	0.39
			Bottom	0	3	2	4	
10.	4	0.239	Top	0	8	13	7	0.44
			Bottom	2	4	3	0	
11.	1	0.642	Top	17	2	1	0	0.32
			Bottom	3	2	2	2	
12.	1	0.866	Top	19	1	0	0	0.23
			Bottom	6	0	1	2	
13.	3	0.597	Top	1	3	15	1	0.70
			Bottom	2	2	2	3	
14.	2	0.478	Top	0	17	0	3	0.53
			Bottom	2	0	4	3	
15.	2	0.418	Top	3	14	0	3	0.43
			Bottom	4	2	1	2	
16.	2	0.284	Top	5	9	2	4	0.42
			Bottom	2	1	3	3	
17.	2	0.522	Top	1	17	1	1	0.40
			Bottom	4	3	1	1	
18.	2	0.179	Top	3	4	9	4	0.17
			Bottom	2	2	3	2	
19.	3	0.788	Top	0	4	19	1	0.27
			Bottom	0	1	5	3	
20.	4	0.485	Top	0	2	5	13	0.35
			Bottom	3	3	1	2	

Table 8

Item Analysis for Post Test 1: Form 3 (QB1-10, A11-20)

Question Number	K.	Item Difficulty	Item Responses (Quartiles)	1	2	3	4	Item Discrimination
1.	3	0.548	Top	2	0	4	0	0.05
			Bottom	3	1	3	4	
2.	2	0.694	Top	0	5	0	1	0.02
			Bottom	3	4	2	2	
3.	1	0.710	Top	6	0	0	0	0.02
			Bottom	5	1	3	2	
4.	3	0.661	Top	0	0	4	2	0.00
			Bottom	1	0	6	4	
5.	4	0.452	Top	0	0	0	6	0.19
			Bottom	3	2	5	1	
6.	2	0.597	Top	0	5	1	0	0.00
			Bottom	3	6	2	0	
7.	3	0.484	Top	0	0	6	0	0.10
			Bottom	5	1	3	2	
8.	4	0.565	Top	1	0	0	5	0.00
			Bottom	3	1	0	7	
9.	3	0.581	Top	0	0	5	1	0.00
			Bottom	0	2	5	4	
10.	4	0.484	Top	1	0	3	2	0.00
			Bottom	3	1	5	2	
11.	1	0.790	Top	6	0	0	0	0.00
			Bottom	7	0	1	1	
12.	1	0.742	Top	5	1	0	0	0.02
			Bottom	4	3	2	2	
13.	2	0.565	Top	0	6	0	0	0.12
			Bottom	4	2	2	3	
14.	1	0.661	Top	6	0	0	0	0.08
			Bottom	3	4	4	0	
15.	2	0.484	Top	1	3	0	1	0.09
			Bottom	5	1	4	4	
16.	2	0.161	Top	1	3	0	2	0.20
			Bottom	4	1	2	4	
17.	3	0.694	Top	0	0	6	0	0.03
			Bottom	3	0	5	3	
18.	2	0.177	Top	0	2	2	2	0.00
			Bottom	2	4	3	2	
19.	2	0.677	Top	0	6	0	0	0.05
			Bottom		4	5	3	
20.	2	0.177	Top	6	3	0	3	0.27
			Bottom	2	0	1	8	

Table 9

Item Analysis for Post Test 1: Form 4 (QA1-20)

Question Number	K.	Item Difficulty	Item Responses (Quartiles)	1	2	3	4	Item Discrimination
1.	4	0.682	Top	0	1	1	11	0.02
			Bottom	1	8	4	10	
2.	2	0.436	Top	2	11	0	0	0.13
			Bottom	14	5	0	4	
3.	3	0.755	Top	0	0	13	0	0.00
			Bottom	5	0	14	4	
4.	3	0.745	Top	0	1	10	2	0.00
			Bottom	2	1	16	1	
5.	2	0.835	Top	0	12	1	0	0.00
			Bottom	0	19	3	0	
6.	2	0.394	Top	0	8	15	1	0.08
			Bottom	1	5	16	0	
7.	3	0.706	Top	0	0	13	0	0.10
			Bottom	7	3	6	6	
8.	4	0.743	Top	0	0	0	13	0.05
			Bottom	4	4	5	9	
9.	2	0.565	Top	0	12	1	0	0.09
			Bottom	1	7	7	6	
10.	4	0.333	Top	0	0	7	6	0.00
			Bottom	2	2	11	6	
11.	1	0.676	Top	12	0	0	1	0.03
			Bottom	10	3	4	4	
12.	1	0.841	Top	13	0	0	0	0.0
			Bottom	13	4	2	1	
13.	2	0.682	Top	0	12	1	0	0.14
			Bottom	2	2	9	7	
14.	1	0.811	Top	13	0	0	0	0.05
			Bottom	9	7	3	0	
15.	2	0.423	Top	11	12	0	1	0.14
			Bottom	7	5	1	4	
16.	2	0.596	Top	1	11	1	0	0.12
			Bottom	5	4	5	3	
17.	3	0.703	Top	0	0	11	2	0.09
			Bottom	1	5	5	4	
18.	2	0.250	Top	1	5	6	1	0.08
			Bottom	3	3	5	3	
19.	2	0.808	Top	3	11	2	0	0.02
			Bottom	1	9	2	2	
20.	2	0.343	Top	0	9	1	3	0.21
			Bottom	2	2	2	3	

Table 10

Item Analysis for Post Test 2: Form 1 (QA1-10, B11-20)

Question Number	K.	Item Difficulty	Item Responses (Quartiles)	1	2	3	4	Item Discrimination
1.	2	0.545	Top	0	17	5	1	0.24
			Bottom	5	4	4	6	
2.	2	0.566	Top	1	17	1	0	0.22
			Bottom	10	5	1	4	
3.	2	0.354	Top	2	12	2	3	0.23
			Bottom	5	4	8	3	
4.	1	0.788	Top	17	1	0	1	0.06
			Bottom	13	0	4	3	
5.	3	0.606	Top	0	1	18	0	0.24
			Bottom	6	2	4	9	
6.	1	0.242	Top	4	14	0	1	0.09
			Bottom	2	10	5	3	
7.	3	0.323	Top	0	3	11	5	0.31
			Bottom	2	5	1	12	
8.	1	0.253	Top	12	1	0	6	0.48
			Bottom	0	9	4	7	
9.	2	0.414	Top	2	13	1	3	0.22
			Bottom	4	4	7	5	
10.	3	0.293	Top	3	3	8	5	0.14
			Bottom	5	3	4	7	
11.	3	0.697	Top	0	0	19	0	0.22
			Bottom	3	3	4	10	
12.	3	0.253	Top	1	0	9	8	0.28
			Bottom	3	4	2	11	
13.	4	0.566	Top	2	0	13	14	0.14
			Bottom	4	5	6	5	
14.	1	0.131	Top	2	0	0	17	0.15
			Bottom	0	4	4	12	
15.	1	0.592	Top	18	0	1	0	0.17
			Bottom	8	5	3	4	
16.	5	0.614	Top	1				
			Bottom	4				
17.	4	0.296	Top	2	5	1	11	0.28
			Bottom	7	6	4	3	
18.	3	0.361	Top	1	6	10	13	0.08
			Bottom	6	6	7	1	
19.	4	0.468	Top	1	0	0	17	0.38
			Bottom	3	5	10	0	
20.	3	0.269	Top	10	2	5	1	0.00
			Bottom	3	3	7	5	

Table 11

Item Analysis for Post Test 2: Form 2 (Q1-20)

Question Number	K.	Item Difficulty	Item Responses (Quartiles)	1	2	3	4	Item Discrimination
1.	2	0.544	Top	1	14	0	0	0.33
			Bottom	6	2	0	1	
2.	2	0.574	Top	2	13	0	0	0.21
			Bottom	2	3	2	0	
3.	2	0.206	Top	1	10	2	1	0.64
			Bottom	0	1	4	4	
4.	1	0.912	Top	15	0	0	0	0.13
			Bottom	7	0	1	0	
5.	3	0.368	Top	1	1	11	2	0.40
			Bottom	4	2	1	2	
6.	1	0.412	Top	12	2	0	1	0.40
			Bottom	1	2	5	1	
7.	3	0.368	Top	1	0	9	5	0.24
			Bottom	4	0	3	2	
8.	1	0.368	Top	8	0	1	6	0.28
			Bottom	1	5	2	1	
9.	2	0.529	Top	0	13	0	2	0.33
			Bottom	4	1	2	2	
10.	3	0.294	Top	0	3	8	4	0.30
			Bottom	2	3	2	2	
11.	3	0.721	Top	0	1	14	0	0.20
			Bottom	2	1	4	2	
12.	3	0.206	Top	0	0	10	5	0.71
			Bottom	2	0	0	7	
13.	4	0.662	Top	0	0	4	11	0.20
			Bottom	1	4	2	2	
14.	3	0.691	Top	1	1	13	0	0.15
			Bottom	0	1	6	2	
15.	4	0.691	Top	0	0	2	13	0.27
			Bottom	3	4	2	0	
16.	2	0.515	Top	1	14	0	0	0.40
			Bottom	4	0	0	5	
17.	2	0.324	Top	5	9	0	1	0.37
			Bottom	3	1	2	3	
18.	2	0.456	Top	3	11	1	0	0.23
			Bottom	3	4	2	0	
19.	3	0.397	Top	1	2	11	1	0.36
			Bottom	0	4	2	3	
20.	2	0.397	Top	2	7	6	0	0.26
			Bottom	1	0	7	1	

Table 12

Item Analysis for Post Test 2: Form 3 (QB1-10,A11-20)

Question Number	K.	Item Difficulty	Item Responses (Quartiles)	1	2	3	4	Item Discrimination
1.	1	0.495	Top	15	3	2	1	0.04
			Bottom	13	5	5	0	
2.	2	0.654	Top	0	20	1	0	0.13
			Bottom	1	11	11	1	
3.	1	0.579	Top	17	3	1	0	0.13
			Bottom	9	2	9	4	
4.	3	0.215	Top	6	1	11	3	0.40
			Bottom	17	5	2	0	
5.	3	0.570	Top	1	0	20	0	0.23
			Bottom	7	5	6	6	
6.	1	0.589	Top	17	2	1	1	0.16
			Bottom	7	8	4	5	
7.	1	0.234	Top	11	2	4	4	0.36
			Bottom	2	5	8	9	
8.	4	0.589	Top	0	3	1	17	0.13
			Bottom	4	5	6	9	
9.	3	0.393	Top	1	3	13	3	0.29
			Bottom	10	4	1	9	
10.	3	0.215	Top	0	9	5	7	0.09
			Bottom	9	5	3	7	
11.	3	0.692	Top	0	1	19	1	0.11
			Bottom	5	3	11	5	
12.	3	0.283	Top	4	0	9	6	0.20
			Bottom	8	6	1	6	
13.	4	0.585	Top	0	1	0	20	0.18
			Bottom	4	3	6	9	
14.	3	0.575	Top	4	0	17	4	0.17
			Bottom	3	9	7	3	
15.	4	0.594	Top	0	0	0	21	0.25
			Bottom	4	7	6	5	
16.	2	0.481	Top	0	18	3	0	0.21
			Bottom	9	7	5	1	
17.	2	0.324	Top	5	13	1	2	0.35
			Bottom	6	1	5	10	
18.	2	0.410	Top	5	15	0	1	0.17
			Bottom	9	8	2	5	
19.	3	0.309	Top	1	5	15	0	0.40
			Bottom	3	12	1	4	
20.	2	0.495	Top	3	14	4	0	0.13
			Bottom	3	8	9	2	

Table 13

Item Analysis for Post Test 2: Form 4 (QB1-20)

Question Number	K.	Item Difficulty	Item Responses (Quartiles)	1	2	3	4	Item Discrimination
1.	1	0.471	Top	9	1	3	0	0.13
			Bottom	5	4	4	1	
2.	2	0.735	Top	0	11	2	0	0.08
			Bottom	3	7	3	1	
3.	1	0.544	Top	13	0	0	0	0.25
			Bottom	4	4	5	1	
4.	3	0.206	Top	4	1	8	0	0.50
			Bottom	8	4	1	1	
5.	3	0.426	Top	2	0	10	1	0.31
			Bottom	5	4	1	3	
6.	1	0.632	Top	13	0	0	0	0.21
			Bottom	4	3	4	2	
7.	1	0.294	Top	7	1	5	0	0.30
			Bottom	1	3	5	5	
8.	4	0.515	Top	0	0	0	13	0.32
			Bottom	2	4	6	2	
9.	3	0.265	Top	1	0	11	2	0.61
			Bottom	3	5	0	6	
10.	3	0.265	Top	1	0	6	6	0.33
			Bottom	4	3	0	7	
11.	3	0.721	Top	0	0	13	0	0.14
			Bottom	2	1	6	5	
12.	3	0.235	Top	3	0	5	5	0.19
			Bottom	5	0	2	7	
13.	4	0.537	Top	0	1	2	10	0.17
			Bottom	2	3	5	4	
14.	1	0.642	Top	12	0	1	0	0.15
			Bottom	6	1	2	5	
15.	1	0.343	Top	5	8	0	0	0.14
			Bottom	2	3	6	3	
16.	5	0.	*Top					0.
			Bottom					
17.	4	0.387	Top	2	0	0	11	0.41
			Bottom	2	9	1	1	
18.	3	0.403	Top	0	2	11	0	0.40
			Bottom	3	4	1	5	
19.	4	0.467	Top	1	0	0	12	0.36
			Bottom	3	4	3	2	
20.	3	0.474	Top	4	2	5	1	0.07
			Bottom	2	4	2	2	

* An error was subsequently found in the question

Table 14

Means and Standard Deviations

for Post Test 1 and 2

Post Test 1

Form of Test	X	S^2	S
1	10.58	9.00	3.00
2	10.70	11.46	3.38
3	10.90	7.80	2.79
4	11.95	11.08	3.33

Post Test 2

Form of Test	X	S^2	S
1	7.96	8.71	2.95
2	9.63	13.97	3.63
3	9.17	11.75	3.43
4	8.34	13.81	3.72

decided to standardize the scores over both rounds. This process was also essential for the analysis of sequence effects.

The index of item difficulty is the percentage of the total group marking a wrong answer or omitting the item, e.g. if 60% of the students give correct answers, the index of difficulty is 100-60 = 40, thus the higher the index the more difficult was the question. The index of item discrimination is the difference between the percentage of the lower group (bottom 25%) marking the correct answer and the percentage of the upper group (top 25%) making the correct answer.

Examination of the results of the item analysis of Post Test 1 revealed that of Form 1 (Table 6) composed of questions A1-10 plus B11-20, only question 5 showed a discrimination index of less than .20. A possible reason for this was that the question was too difficult (index of difficulty .792). Form 2 (Table 7), questions B1-20, provided similar results, in this case question 18 had an index of discrimination of .17 presumably in this case the item was too easy (item difficulty .179). Form 3 (Table 8), (questions B1-10 and A11-20) provided less satisfactory results. Comparison of the item difficulty indices for questions B1-10 on this form with those for the same B1-10 questions on Form 1 shows similar levels of difficulty but comparison of the discrimination indices reveal considerable differences. Since the questions were the same the only source of variance would be pupil variation despite the fact that the tests were randomly distributed between classes. This same problem would seem to have occurred in Form 4 (Table 9) since very different discrimination indices appear with questions A1-10 on Forms 1 and 4. Questions A11-20 on Forms 1 and 4 would seem to have been less effective in discrimination than those of B11-20. Examination of the questions themselves shows that the questions were very similar in format and the complexity of the figures involved, varying only in content.

The item analysis for Post Test 2, alternative forms 1-4 show less inter-form variation in item discrimination. Form 1 (Table 10), questions A1-10 and B11-20, contained four items with very low levels of discriminability. Item 4 would appear to have been too difficult whereas questions 6, 8, and 20 may possibly have been too easy. On Form 2 (Table 11), A1-20, question 4 appeared to have been too difficult as did question 14. Form 3 (Table 12), B1-10 and A11-20) contained two questions, 1 and 10, with low levels of discrimination. In the case of question 10 the easiness of the question may be the explanation. Finally Form 4 (Table 13), B1-20, had question 2 which appeared to be too difficult whereas an explanation for the low value for question 20 would not appear to be as obvious.

Reliability

Reliability information was obtained from a mixed ability grade six class of 29 pupils which, although part of the project, was not included in the analysis. A one week test-retest design was used. On both occasions the alternative forms of the test were randomly distributed giving the advantages of a parallel form design. This method was chosen to take into consideration a variety of measurement errors. The test-retest method of estimating reliability takes into account day to day fluctuations in pupil performance, variation in administrational procedure and testing conditions. The added advantage of using parallel forms is that it reduces any practice effects associated with a test-retest design and takes into account fluctuations related to content sampling. On both occasions the same person administered the test under the same conditions. The following data were obtained.

Test	Coefficient of stability and equivalence
Pretest	.64
Post Test 1	.72
Post Test 2	.82

Validity

The validity of a test can only be assessed with reference to the proposed use of the test. Since a particular test may be used for a variety of purposes, a different type of validity is relevant each time. The "Science Process Achievement Test" was used to measure the process achievement of grade six pupils, therefore content validity is of primary importance as with any achievement test. The content validity of the test is the extent to which the test requires demonstration, by the student, of the knowledge and skills which constitute the objectives of instruction in the area.

Content Validity. In this study, since the major emphasis was on the aquisition of selected processes of science, analysis of the content validity must include an assessment of whether such processes as well as topics were sampled in a representative manner. Tables 15 and 16 provide analyses of Post Test 1 and 2 in terms of content units and science processes. Tables 17 and 18 show each of Post Test 1 and 2 in more detail, illustrating the objectives covered by each question. The parallel forms of each test covered the same objective, variation being restricted to slight changes in item stems. Since these alternative forms were randomly distributed these variations

Table 15

Table of Specifications: Content by Process

Post Test 1

	Interpreting Data	Controlling Variables	Measurement	Total
Batteries and bulbs	6	4	2	12
Human Reactions	2	1	0	3
Transfer	0	1	4	5
Total	8[a]	6	6	20

Note. This test was intended to cover the unit "Batteries and Bulbs" and approximately one third of "Human Reactions".

a. Slightly more emphasis was placed on "Interpreting Data" as this process was emphasized throughout the lessons.

Table 16

Table of Specifications: Content by Process

Post Test 2

	Interpreting Data	Controlling Variables	Measurement	Total
Human Reactions	1	4	0	5
Mechanics	4	2	4	10
Transfer	3	0	2	5
Total	8	6	6	20

Note. Since 2/3 of the "Human Reactions" unit was taught in Round 2, this was reflected in the relatively higher number of items in this test.

Table 17

Content Validity Analysis

Post Test 1

Q	Subject Area	Process	Objective
1.	Pulse rates as function of exercise	Interpreting data	Construction of histogram
2.	" "	" "	Interpretation of line graph
3.	Strength of e/magnet	Interpreting data	Calculation of mean
4.	" "	" "	Calculation of range
5.	Transfer item	Controlling variables	Identifying variables infl. responding variables
6.	Pulse rates	" "	Identifying respond. var.
7.	Batteries & Bulbs	" "	Identifying relevant var.
8.	" "	" "	Identifying constant & manipulated var.
9.	" "	" "	Identifying respond. var.
10.	" "	" "	Identifying uncontrolled var.
11.	Transfer item	Measurement	Use of metric units
12.	" "	"	Interconversion of metric units
13.	" "	"	Averaging of errors—accuracy
14.	Strength of an e/magnet	Interpreting data	Prediction by interpolation
15.	" "	" "	Prediction by extrapolation
16.	" "	" "	Interpretation of line graph
17.	" "	" "	Accuracy of interpolation and extrapolation
18.	Electromagnet	Measurement	Averaging errors
19.	Transfer item	"	Use of arbitrary units
20.	Batteries & bulbs	"	Use of instruments

Table 18

Content Validity Analysis

Post Test 2

Q	Subject Area	Process	Objectives
1.	Human reactions	Interpreting data	Construction of histogram
2.	Transfer	" "	Interpretation of line graph
3.	Transfer	" "	Calculation of the mean
4.	Transfer	" "	Calculation of the range
5.	Mechanics	Controlling variables	Identifying relevant variables
6.	Mechanics	"	Identifying manipulated variables
7.	Human reactions	Controlling variables	Identifying manipulated variables
8.	Human reactions		Identifying constant variables
9.	Human reactions	"	Identifying respond. variables
10.	Human reactions	"	Identifying uncontrolled variables
11.	Transfer	Measurement	Use of metric units
12.	Transfer		Interconversion of metric units
13.	Mechanics	"	Averaging errors for accuracy
14.	Mechanics	Interpreting data	Prediction by interpolation
15.	Mechanics	" "	Prediction by extrapolation
16.	Mechanics		Interpretation of line graph
17.	Mechanics	" "	Extrapolation & interpolation
18.	Mechanics	Measurement	Use of balance as instrument
19.	Mechanics	"	Subunits
20.	Mechanics	"	Instruments

served to increase the reliability of the tests overall. Tables 19 and 20 provide data on the sampling of both process objectives and content. As can be seen from these tables each of the two post tests covered all of the objectives with the exception of B7 and 8 and C8 which were not tested directly.

Construct validity. Since the "Science Process Achievement Test" is not merely measuring recall of specific facts but also the ability to carry out the processes of science, one must also consider its construct validity in terms of some model of science process. Despite the complexity of science, it is acknowledged that in any scientific investigation certain empirical processes and methods of analysis are invariably found, e.g. scientific observations are made, classifications carried out, the effect of changing variables on other variables is investigated and interpretations of data are made. The ideal way of assessing a pupil's ability to carry out these processes is to watch him actually performing, but this is not usually possible due to time and apparatus restraints. Thus for practical purposes group paper and pencil measures are required. The construct validity of such tests is more difficult to establish. In the case of the "Science Process Achievement Test" a certain amount of evidence can be obtained from the method of construction. The model of science processes was adapted from the "AAAS Science A Process Approach"

Table 19

Comparison of Objectives and questions

set for parallel forms A and B

Post Test 1

o Objective set
* Form A
+ Form B

Process	Controlling Variables. A.						Interpreting Data B.								Measuring C.							
Objectives	1	2	3	4	5	6	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Batteries & Bulbs																						
Content Units																						
1. Measurement of bulb brightness							o	o	o				o	o	o		o	o		o		o
2. Factors influencing bulb brightness		* o +	* o +	* o +	* o +	* o +				+	+	+	+								+	
3. Measurement of strength of electromagnet							* o	* o	o	*	* o	* o	* o		o	* o						
4. Factors influencing strength of electromagnet	o	o	o	o	o	o										o					* o	

contd.

Table 19 (continued)

Comparison of Objectives and questions
set for parallel forms A and B
Post Test 1

o Objective set
* Form A
\+ Form B

Process	Controlling Variables. A.						Interpreting Data. B.								Measuring C.							
Objectives	1	2	3	4	5	6	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Human Reactions																						
1. Measurement of learning & forgetting																	o					
2. Factors influencing learning & forgetting																						
3. Measurement of pulse rate	* o									* o +	*											o
4. Factors influencing pulse rates	o	o	o	o	o	o					* +											
5. Optical illusions																	o					
Transfer	* +														* +	* +	* +	* +	* +			

Table 20

Comparison of Objectives and questions

set for parallel forms A and B

Post Test 2

o Objective set
* Form A
\+ Form B

Process	Controlling Variables. A.						Interpreting Data. B.								Measuring C.							
Objectives	1	2	3	4	5	6	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
<u>Human Reactions</u>																						
1. Measurement of pulse rate						o	o	o						o								o
2. Factors influencing pulse rate	o								o	o												
3. Measurement of learning & forgetting									* o													
4. Factors influencing 3.	o	* o +	* o +	* o +	* o +	o																o
5. Optical illusions																	o					

Contd.

Table 20 (continued)

Comparison of Objectives and questions set for parallel forms A and B

Post Test 2

o Objective set
* Form A
\+ Form B

Process	Controlling Variables. A.						Interpreting Data. B.								Measuring C.							
Objectives	1	2	3	4	5	6	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
<u>Mechanics.</u>																						
1. Measurement of rolling cylinders	o																					
2. Factors influencing No. 1.	* +	o	o	o																		
3. Measurement of stretch of rubber bands					o		o	o			+ o	+ o										
4. Factors influencing pendulum swing			o	o		o			o	* o	* o +	* o					o					
5. Weighing objects using balance																	* o +	* o +		* o +		o
Transfer							* +	* +		* +							* +	*	* +			

and "Elementary Science Curriculum Study" (Crocker, 1971) and used to prepare the objectives which formed the basis for the questions. This test thus has construct validity to the extent that one is prepared to accept the model and also to the extent that the test items actually represent the components of the model that are of interest in the study, namely quantifying, controlling variables and interpreting data. This representation has already been established in the tables of specification.

Criterion-related validity. The format for the test questions was based closely on that of the questions in Robison (1973) which had been validated against actual performances on the "Individual Competency Measures" of SAPA.

V EXPERIMENTAL RESULTS

As has been already discussed earlier, the purpose of the study was to examine the variations in teacher behavior with specific reference to "teacher control". The study was designed to examine two groups of hypotheses, one group including treatment main effects and the interactions between treatments and pupil characteristics for each of two rounds; and the other group including treatment main effects and the interactions between treatments and pupil characteristics for each of the two treatment sequences (shifts), ie. high to low and low to high.

This chapter presents the findings of these two sets of analyses and also data on the implementation of the treatments. Emphasis in the first analysis has been placed on results that are consistent over both rounds. The results presented here are confined to those involving science process achievement. Findings for other dependent variable used in the overall study are reported elsewhere (Crocker, et al., 1977).

Implementation of the Treatments

The variables on which the teachers were expected to differ in implementing the two treatments are given in Table 21. Generally speaking the results were in the expected direction. In teacher responding and reacting, no differences were found. Differences in the opposite direction from that anticipated were discovered for pupil reporting results. In this latter case however there existed large variance within lessons in each treatment which could have reduced the reliability of these results. Although in the desired direction, the results for soliciting moves showed a smaller difference than anticipated.

Analysis

The method of analysis for the achievement criterion was a generalized multiple regression. The independent variables were entered hierarchically in the following manner:

(a) initial ability (using CTBS and Raven entered simultaneously)

(b) teachers

(c) treatment-teacher interactions

(d) stepwise inclusion of ATI's

(e) treatments

Table 21

Comparison of Behavior in High and Low Control Treatments

Variable	High Control	Low Control
1. Teacher-class interaction	876	419
2. Teacher-group interaction	671	1095
3. Proportion of Pupil Talk during Teacher-Class Interaction	0.06	0.10
4. Proportion of Pupil Talk during Teacher-Group Interaction	0.24	0.29
5. Teacher Structuring	453	369
6. Teacher Soliciting	539	650
7. Teacher Responding	89	84
8. Teacher Reaching	229	227
9. Ratio of Commands to Total Solicits	0.69	0.79
10. Pupils Reporting Individual Results	76	46
11. Teacher Substantive Logical Discourse	187	230
12. Pupil Substantive Logical Discourse	34	55
13. Average Wait Time (sec)	1.5	1.9

Note. Units are total sentence uttered in a sample of approximately 130 minutes lesson time for each treatment. The sample includes 13 lessons with all teachers included at least once in each treatment. Ratios and wait time computed for same sample data.

The results of the regression analysis carried out for each round of the study are presented in Table 22. As can be seen in the table, ability measures have accounted for the largest proportion of the variance. This is consistent with the findings of most such studies. The interaction effects account for six and eight percent of the variance in Round 1 and Round 2 respectively. In both rounds the largest single interaction was with the attitude measures ie. attitude to school and attitude to science. Interaction between treatment and attitude would seem to be the only one that is stable across the two rounds. Despite the significant interactions between treatment and teacher, locus of control and neuroticism in Round 2 these are not replicated across the two rounds. Since there appear to be more significant interactions in Round 2 it may be possible to conclude that the treatments were better differentiated in the second round. It is however impossible to assess this directly since implementation of treatments was not analyzed for the two rounds separately.

The results of the regression analysis carried out on the two possible shifts or treatment sequences are given in Table 23. Shift one is the sequence high to low treatment; shift two is the sequence low to high treatment. As in the previous regression by rounds, the ability measures account for the largest proportion of the variance. In the high to low shift ie. shift one, ability accounts for

Table 22

Summary of Regression Analysis by Round

Science Process Achievement

Source of Variance	Round 1 (N=164)		Round 2 (N=232)	
	R^2	F	R^2	F
Ability (IQ, CTBS)	0.364	46.12***	0.500	114.34***
Teachers	0.007	0.89	0.037	6.02*
Treatment x Teachers	0.005	0.63	0.040	7.03*
x Attitude to School	0.038	10.11**	0.002	1.11
x Attitude to Science	0.001	---	0.010	5.40*
x Locus of Control	0.002	0.54	0.008	4.50*
x Neuroticism	0.001	---	0.011	6.08*
x Extraversion	0.001	---	0.004	2.12
Treatment	0.004	1.02	0.001	---

*P < .05 **P < .01 ***P < .001

Table 23

Summary of Regression Analysis by Shift
Science Process Achievement

Source of Variance	Shift 1 (N=194) R^2	F	Shift 2 (N=198) R^2	F
Ability	0.252	30.910***	0.231	35.277***
Teachers	0.017	0.837	0.026	1.566
Treat. x Teachers	0.061	2.010*	0.044	1.631
x Compre.	0.001	0.013	0.002	0.051
x Att. Sch.	0.002	0.034	0.003	0.070
x Extra.	---	---	0.008	0.262
x Vocab.	0.001	0.011	0.007	0.115
x CTBS	0.001	0.004	0.008	0.191
x DPS	0.001	0.001	---	---
x Subdom	0.001	0.015	0.005	0.0069
x SES	---	---	0.003	0.076
x Scona.	0.002	0.023	0.008	0.013
x Sex	0.001	0.711	0.003	0.061
x IAR	0.001	0.006	0.015	0.487
x Neuro.	0.002	0.070	0.010	0.259
x Att. Sci.	---	---	0.001	0.110
Treatment	0.001	0.001	0.001	0.000

Note. Shift analysis: Shift 1 = High to Low Treatment Sequence
Shift 2 = Low to High Treatment Sequence

*p < .05 **p < .01 **p < .001

25% of the variance whereas in the opposite shift, low to high, ability accounts for 23% of the variance.

The treatment teacher interaction accounts for the largest single interaction in this shift analysis. In the low to high shift the percentage of variance accounted for is 4% whereas the opposite shift this interaction accounts for 6% and is significant at the .05 level. All other interactions were non-significant. Table 24 gives the cell means for the teacher-treatment interaction for shift one, high to low treatment sequence. This interaction was significant at the .05 level. From this table it can be seen that teacher four was associated with higher scores in high control whereas higher scores were found in low control with teachers one and two.

Table 24

Cell Means for Teacher-Treatment Interaction

High To Low Sequence

Teacher	High	Low
4	0.2512	-0.5400
1	-0.0754	0.6568
2	-0.3648	0.2373

Low To High Sequence

Teacher	High	Low
4	0.1519	-0.3697
1	0.9197	0.2112
2	-0.1234	0.0291

VI SUMMARY AND CONCLUSIONS

The notion that pupil classroom behavior and achievement are a combined function of the person and the environment is not a new one, but one that would seem particularly appropriate in the context of elementary science programs. Most of the newer elementary science programs emphasize less control on the part of the teacher, a requirement that might contribute to problems in the implementation of such programs. Thus the aim of this study was to examine the effects of variation in teacher control upon process achievement.

This study has included several features designed to overcome some of the limitations of earlier studies. Based on a number of teacher behaviors, operational definitions of the two extremes of the continuum of teacher control were prepared. The treatments were monitored by the coding and analysis of videotapes of a sample of class activities.

All classes received both treatments although not in the same sequence. Since the regular grade six teachers were replaced by project teachers, it was possible to train the teachers in the implementation of the treatments and randomly assign them to class.

Summary of findings

1. Generally speaking, treatment differences were implemented as defined.
2. In the analysis by round, pupil ability accounted for the largest percentage of the variance in process achievement. Teacher effects and treatment-teacher interactions were found in round two, but results consistent with these were not found in round one. Overall, ATI effects only accounted for 6% of the variance, statistically significant effects being restricted to pupil attitude variables. After all, other effects had been considered in the regression equation, treatment effects did not account for a significant portion of the variance. The results of the analysis for the second round do seem to indicate that teacher effects and interactions were more significant than those for round one. A possible explanation for this difference could be the clearer differentiation of treatments in round two.
3. In the treatment sequence analysis, no significant ATI's were found. In the low to high shift however a significant treatment teacher interaction was found.

Conclusions

Comparing these results with those of the other two similar studies, (Shymansky and Matthews, 1974 and Crocker, Bartlett and Elliott, 1976) it appears that all three yield different results for treatment differences. Shymansky and Matthews showed that pupil achievement was greater in low control; Crocker, Bartlett and Elliott showed the opposite results, and no significant results were found for treatment differences in the present study. A possible explanation of these differences lies in the varying length of the studies. Shymansky and Matthews conducted a longer study than the present study whereas that of Crocker, Bartlett and Elliott was shorter. It would seem that an interesting avenue for further investigation would be whether the length of the study period influences the effectiveness of the treatment.

Neither of these two studies examined sequence effects so in this respect are not comparable with the shift analysis. In this analysis it would seem that treatment sequence has no significant influence on achievement. Ability appears to account for about the same percentage of the variance in both treatment sequences. The significant treatment-teacher interaction in the high to low shift suggests that the influence of treatment sequence upon teachers is an important factor to be taken into consideration.

In all four of the analyses, ability accounted for by far the greatest percentage of the variance. It would seem that pupil ability remains the dominant influence on achievement. In all cases interactions accounted for more variance than the treatments. With the exception of treatment attitude interactions in the analysis by round, the ATI's were insignificant over one or both rounds. The non significance of the interactions between treatment and neuroticism in the analysis by round would seem to be inconsistent with that of the Crocker, Bartlett and Elliott (1976) study. The lack of a significant interaction between treatment and extraversion, and treatment and dependence proneness in this study would however correspond to a similar situation in the Crocker, Bartlett and Elliott (1976) study.

It is noted that in another part of the study it was shown the pupils tended to prefer the low control treatment. Inspite of this the differences in treatment do not lead to differences in achievement. It would appear however that the interaction between teacher and treatment in both preference (Crocker, et al., 1977) and in some cases achievement are important factors to be considered. It also appears that teachers have greater impact on pupil preference than achievement.

REFERENCES

American Association for the Advancement of Science. Commission on Science Education. The Psychological Bases of Science - A Process Approach. AAAS Miscellaneous Publication, 1965.

American Association for the Advancement of Science. Individual Competency Measures of Science - A Process Approach, 1967.

American Association for the Advancement of Science. An Evaluation Model and its Application. Washington: AAAS, 1968.

Anderson, H.H., The Measurement of Domination and Socially Integrative Behaviour in Teachers' Contacts with Children. Child Development, 10:73-89, 1939. Reprinted in E.J. Amidon and J.B. Hough (Eds.), Interaction Analysis: Theory, Research and Application. Reading, Mass.: Addison Wesley, 1967.

Barker, R.G., Ecological Psychology. Stanford, California, Stanford University Press, 1963.

Beard, J., The Development of Group Achievement Tests for Two Basic Processes of AAAS Science - A - Process Approach. Journal of Research in Science Teaching, 8:179-183, 1971.

Bellack, A.A., et al. The Language of the Classroom, New York: Teachers College Press (Columbia University) 1966.

Biological Science Curriculum Study. Processes of Science Test. New York: The Psychological Corporation, 1962.

Blishen, B.R. A Socio-economic Index for Occupations in Canada. Canadian Review of Sociology and Anthropology, 40: 627-646, 1970.

Brookover, W.B., Peterson, A., Thomas, S. Self Concept of Ability and School Achievement, Office of Research and Publications, College of Education, Michigan State University, Michigan, 1962.

Brophy, J. and Evertson, C. Process-product correlations in the Texas Teacher Effectiveness Study: Student sex, grade, and socio-economic status differences in classroom process measures. Report No. 74-4. Austin, Texas: Research and Development Centre for Teacher Education, University of Texas, 1974.

Carline, J.L. An Investigation of the Relationship between Various Verbal Strategies of Teaching Behaviour and Achievement of Elementary School Children. Unpublished Doctoral Dissertation, Syracuse University, Syracuse, New York, 1969.

Crandell, V.C., Katovsky, W., Crandell, V.J. Children's Beliefs in their Own Control of Reinforcements in Intellectual Academic Achievement Situations. Child Development, 36: 91-109, 1965.

Crocker, R.K. Elementary Science Curriculum Study. (Teaching Guides Volumes I & II) Toronto: McGraw Hill Ryerson, 1973.

Crocker, R.K., et al. A Category System for the Analysis of Interact-Patterns in Elementary Science Classes. Unpublished Paper, 1975.

Crocker, R.K., Bartlett, K.R. and Elliott H.G. A Comparison of Structured and Unstructured Modes of Teaching Elementary Science Process Activities. Journal of Research in Science Teaching, 13: 267-274, 1976.

Crocker, R.K., et al., Teacher Control, Classroom Behaviors and Pupil Performance in Elementary Science Classes. Institute for Educational Research and Development, Memorial University of Newfoundland, 1977.

Dunkin, M.J., Biddle, B.J. The Study of Teaching. New York. Holt Rhinehart and Winston Inc., 1974.

Eysenk, S.B., The Junior Eysenk Personality Inventory. California Educational and Industrial Testing Service, 1963.

Flanders, N.A. Interaction Analysis in the Classroom: A Manual for Observers. University of Michigan, Ann Arbor, 1960.

Flanders, N.A. Teacher Influence in the Classroom. in E.J. Amidon and J.B. Hough (Eds.), Interaction Analysis: Theory, Research and Application. Reading, Mass., 1967.

Flanders, N.A., Amidon, E.J. The Effects of Direct and Indirect Teacher influence on Dependence proness of Students Learning Geometry. Journal Of Educational Psychology, 52: 286-291, 1961.

Flanders, N.A. Anderson, J.P., Amidon, E.J. Measuring Dependence Proness in the Classroom. Educational and Psychological Measurement, Vol. XXI, No. 3, 1961.

Gates, A.I., MacGinitie, W.H. Gates MacGinitie Reading Tests. Teachers College, Columbia University, New York, 1965.

Good, T.L., Brophy, J.E., Sykes, J.N. Effects of Teacher Sex, Student Sex and Student Achievement on Classroom Interaction. Journal Of Educational Psychology, 65: 74-87, 1973.

Herman, W.L. An Analysis of the Activities and Verbal Behaviour of Selected Fifth Grade Social Studies Classes. Classroom Interaction Newsletter, 2 (2): 27-29, 1967.

Hunt, D.E. Matching Models in Education. Monograph Series No. 10, Ontario Institute for Studies in Education, 1970.

Hunt, D.E. and Sullivan, E.V. Between Psychology and Education. Hinsdall Illinois, Dryden Press, 1974.

Kerlinger, F.N. and Pedhauzer, E.J. Multiple Regression in Behavioural Research. New York, Holt Rhinehart and Winston Inc, 1973.

McLoed, R.J. et al. The Development of Criterion Validated Test Items for Four Integrated Science Processes. Journal of Research in Science Teaching, 12: 415-421, 1975.

Morrison, T.L., Control as an Aspect of Group Leadership in Classrooms: A Review of Research. Journal of Education, 156: 38-64, 1975.

Porter, R.B. and Catell, R.B. Handbook for the Children's Personality Questionaire. Institute for Personality and Ability Testing, Illinois, 1968.

Rainey, R.G. The Effects of Directed versus Non-directed Laboratory on High School Chemistry Achievement. Journal of Research in Science Teaching, 12: 415-421, 1965.

Raven, J.C. Guide to the Standard Progressive Matrices Sets A,B,C,D,E. H.K. Lewis and Co. Ltd., London, 1960.

Rian, H. Teacher Leadership and Pupil Reaction: The Authoritarian-Democratic Dimension Revisited. Scandinavian Journal of Educational Research, 13: 1-15, 1969, in Dunkin, M.J. and Biddle, B.J. (Eds.) The Study of Teaching. New York, Holt Rhinehart and Winston Inc. 1975.

Robison, R.W. The Development of Items which Assess the Processes of controlling Variables and Interpreting Data. Unpublished Doctoral Thesis, Michigan State University, 1970.

Rosenshine, B. Classroom Instruction. in N.L. Gage (Ed.) The Psychology of Teaching Methods. 77th yearbook of the National Society for the Study of Education, 1976.

Shymansky, J.A. and Matthews, S.C. A Comparative Laboratory Study of the Effects of Two Teaching Patterns on Certain Aspects of the Behaviour of Students in Fifth Grade Science. Journal of Research in Science Teaching, 11: 157-168, 1974.

Soar, R.S. Optimum Teacher-Pupil Interaction for Pupil Growth. Educational Leadership, 26: 275-280, 1968.

Solomon, D. and Kendall, A.J. Final Report on Individual Characteristics and Children's Performance in Varied Educational Settings. Psychological Services Section, Montgomery County Public Schools, Rockville, Maryland, 1976.

Tannenbaum, R.S. The Development of the Test of Science Processes. Journal of Research in Science Teaching, 8: 179-183, 1971.

Tisher, R.P. Verbal Interaction in Science Classes. Journal of Research in Science Teaching, 8: 1-8, 1971.

Welch, W. The Development of an Instrument for Inventorying Knowledge of the Processes of Science. Journal of Research in Science Teaching, 5: 64-68, 1967.

APPENDIX

Table 1

Mean CTBS Scores for each Class

Class	Mean CTBS (comp.)
1	61.5
2	66.5
3	62.5
4	67.3
5	66.4
6	70.2
7	73.5
8	72.9
9	76.5
10	67.7

Table 2

Summary Table for Multiple Regression

High to Low Shift Classes 1, 8, 10

Variable	Multiple R	R^2	ΔR^2
Ability	0.5025	0.2525	0.2525
Teachers	0.5192	0.2696	0.0171
Tretch.	0.5747	0.3303	0.0607
TreSex.	0.6046	0.3655	0.0352
Treneur.	0.6082	0.3699	0.0044
TreatSch.	0.6110	0.3721	0.0022
Tresco.	0.6110	0.3737	0.0016
TreSub.	0.6123	0.3749	0.0012
TreGmCom.	0.6133	0.3761	0.0011
TreIar.	0.6137	0.3766	0.0005
TreRav.	0.6141	0.3771	0.0010
TreDPS	0.6142	0.3773	0.0001
TreGmVoc.	0.6143	0.3773	0.0001
TreCTBS	0.6144	0.3774	0.0005
Treat.	0.6143	0.3774	0.0000

Table 3

Summary Table for Multiple Regression

Low to High Shift Classes 2, 3 and 6

Variable	Multiple R	R^2	ΔR^2
Ability	0.48107	0.23142	0.23142
Teachers	0.50752	0.25757	0.02615
Tretch.	0.54935	0.30179	0.04422
TreIar.	0.56280	0.31674	0.01495
TreExt.	0.57072	0.32572	0.00898
TreNeur.	0.57926	0.33554	0.00982
TreCTBS.	0.58608	0.34348	0.00794
TreSES.	0.58902	0.34694	0.00346
TreATSCH.	0.59195	0.35040	0.00346
TreSex	0.59467	0.35364	0.0032
TreGmCom.	0.59709	0.35651	0.00288
TreGMVoc.	0.60290	0.36349	0.00697
TreScona.	0.60361	0.36435	0.00086
TreatSci.	0.60424	0.36511	0.00076
TreSub.	0.60467	0.36562	0.00051
TreRav.	0.60494	0.36595	0.00033
Treat.	0.60494	0.36595	0.00000

Table 4

Correlation Matrix

Variables	Raven	CTBS comp.	Pretest	Post test 1	Post test 2	Neuro- ticism	Extra.	DPS	Submiss. domin.	S.E.S.
Raven	1.00	0.553								
C.T.B.S.	0.553	1.000								
Pretest	0.332	0.559	1.000							
Post test 1	0.429	0.624	0.389	1.000						
Post test 2	0.443	0.604	0.442	0.498	1.000					
Neuroticism	-0.116	-0.131	-0.114	-0.195	-0.108	1.000				
Extra.	0.194	0.257	0.161	0.090	0.249	-0.172	1.000			
D.P.S.	0.010	0.024	0.017	0.080	-0.092	-0.227	0.045	1.000		
Submis.Dom.	0.191	0.174	0.144	0.099	0.208	0.008	0.213	-0.444	1.000	
S.E.S.	-0.239	-0.309	-0.203	-0.159	-0.290	0.105	-0.152	0.041	-0.141	1.000

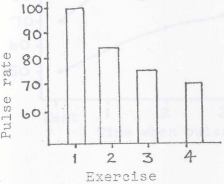
POST TEST 1 FORM A

1. An experiment was carried out to compare the pulse rates of two persons A and B immediately after they finished three minutes of an exercise. The pulse rates were taken after lying down, typing, walking, running up and down stairs. Here are the results.

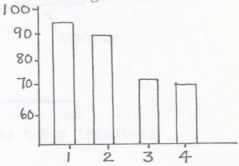
Exercise	Pulse rate of Person A	Pulse rate of Person B
1. Lying down	70	72
2. Typing	75	75
3. Walking	85	90
4. Running up and down stairs	100	95

Look at the histograms below.

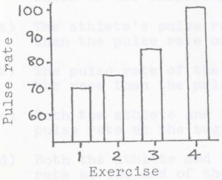
Histogram I



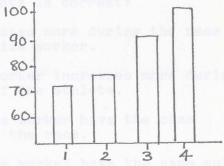
Histogram II



Histogram III



Histogram IV



Which histogram represents the results for person A?

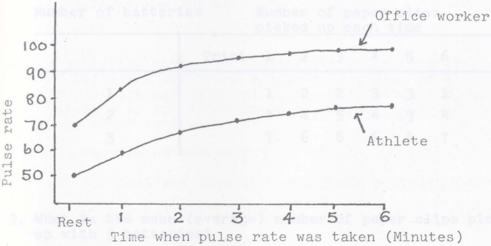
a) Histogram I

b) Histogram II

c) Histogram III

d) Histogram IV

2. The graph below shows the change in pulse rate of an athlete and an office worker while they were both running a mile in 6 minutes. The pulse rate was taken every minute.



Look at the graph.

Which of the following statements is correct?

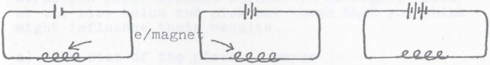
a) The athlete's pulse rate increases more during the race than the pulse rate of the office worker.

b) The pulse rate of the office worker increases more during the race than the pulse rate of the athlete.

c) Both the athlete and the office worker have the same pulse rate at the beginning of the race.

d) Both the athlete and the office worker have the same pulse rate at the end of the race.

An experiment was done to see if an electromagnet could be made stronger by adding more batteries as in the diagram below.



1 battery 2 batteries 3 batteries

Each electromagnet in the diagram was used to pick up paper clips. In each case several trials were made and these are the results:

Number of batteries	Number of paper clips picked up each time					
Trial	1	2	3	4	5	6
1	1	2	2	3	3	1
2	4	4	5	4	3	4
3	7	6	6	6	4	7

3. What is the mean (average) number of paper clips picked up with 3 batteries?

a) 6

b) 5

c) 7

d) 4

4. Which number of batteries picks up the largest range of paper clips?

a) 1

b) 2

c) 3

d) 1 and 3

5. A group of students wanted to see which color of construction paper becomes warmest in sunlight. Look at the list below and pick out those that you think might influence their results.

    a) the size of the piece of paper

    b) the color of the paper

    c) the type of paper

    d) all of the above.

6. A scientist wanted to compare the pulse rate of people of different ages from a baby to a person of 80 years. If you were carrying out such an investigation, which variable do you think would it be most important to keep under control.

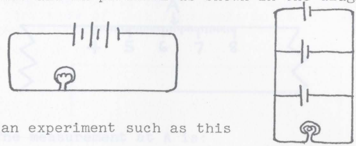
    a) the age of the people

    b) the activity that the people are carrying out

    c) the sex of the people

    d) the air temperature.

An experiment was carried out to compare the brightness of a bulb when the batteries in the circuit were connected in series and in parallel as shown in the diagram below.



In an experiment such as this

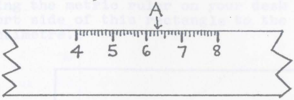
7. What variable was changed by the experimenter?
   a) The number of batteries
   b) The number of bulbs
   c) Type of circuit (series or parallel)
   d) Brightness of the bulbs

8. What variables were held constant in the experiment above?
   a) The number of batteries
   b) The number of bulbs
   c) Arrangement of the bulbs
   d) Both a) and b).

9. Which variable in the above experiment would you measure to see if it had changed?
   a) The number of bulbs
   b) The number of batteries
   c) The brightness of the bulbs
   d) Both a) and b).

10. Which variables were not under control but which might have altered the results?
   a) The temperature of the air
   b) The distance apart of the batteries
   c) The strength of the batteries
   d) Both b) and c).

11. Look at this portion of a metric ruler.



The measurement at A is:

a) 6 centimetres 3 millimetres

b) 6 centimetres 4 millimetres

c) 5 centimetres 3 millimetres

d) 1 centimetre 3 millimetres

12. Using the metric ruler on your desk measure the short side of this rectangle to the nearest centimetre.

The short side is:

a) 3 centimetres

b) 4 centimetres

c) 2 centimetres

d) 1 centimetre

13. In the metric system:

1 metre = 100 centimetres

1 centimetre = 10 millimetres

Convert

35 millimetres into centimetres and millimetres

The answer is:

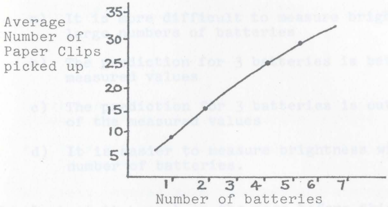
a) 35 centimetres

b) 3 centimetres 5 millimetres

c) 5 centimetres 3 millimetres

d) 3 metres 5 centimetres

The following graph shows how increasing the number of batteries in a circuit, will increase the brightness of a bulb.



14. Predict the brightness of the bulb when there are 3 batteries in the circuit.

   a) 20

   b) 15

   c) 25

   d) 10

15. Predict the brightness of the bulb when there are seven (7) batteries in the circuit.

   a) 30

   b) 35

   c) 25

   d) 27

16. Why do you think the prediction for the brightness for 3 batteries is better than the prediction for 7 batteries?

a) It is more difficult to measure brightness with large numbers of batteries

b) The prediction for 3 batteries is between 2 measured values

c) The prediction for 3 batteries is outside any of the measured values

d) It is easier to measure brightness with a large number of batteries.

17. Look at the graph on the page before this one. If you got a brightness reading of 20 sheets of paper, how many batteries were being used in the circuit.

a) 5

b) 2

c) 3

d) 4

18. A student carried out an experiment to see how many paper clips an electromagnet with 15 coils would pick up. He repeated the experiment 3 times and here are his results.

Trial	No. of paper clips picked up
1	9
2	7
3	5

Since these results vary quite a bit, how could the student best get a more accurate result?

a) take one more reading and see which of the three it is nearest to.

b) take several more readings and average all the results

c) average the three results he already has

d) none of the above

19. You need to compare the time it takes 5 children to complete a jigsaw puzzle but you don't have a clock to time them. Instead you use a dripping tap and count the number of drips.

Which of the following is most important if you use a dripping tap as a clock?

a) the tap must be of a certain size

b) there must be equal time between drips

c) you must time the drips with a clock first

d) the water must be clear

20. If you were asked to compare the brightness of a bulb in your circuit with a bulb in your partner's circuit, which of the following statements do you think is most important?

a) the type of paper in your brightness meter

b) the same meter is used to test both bulbs

c) all the pieces of paper must be of the same kind

d) both a and c

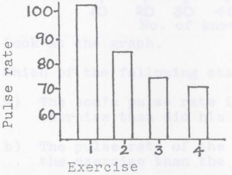
POST TEST 1 FORM B

1. An experiment was carried out to compare the pulse rates of two persons A and B immediately after they finished three minutes of an exercise. The pulse rates were taken after lying down, typing, walking, running up and down stairs. Here are the results.

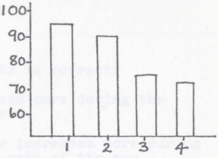
Exercise	Pulse rate of Person A	Pulse rate of Person B
1. Lying down	70	72
2. Typing	75	75
3. Walking	85	90
4. Running up and down stairs	100	95

Look at the histograms below.

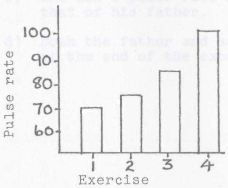
Histogram I



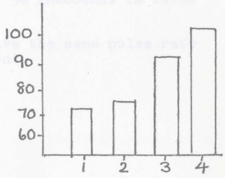
Histogram II



Histogram III



Histogram IV



Which histogram represents the results for person B?

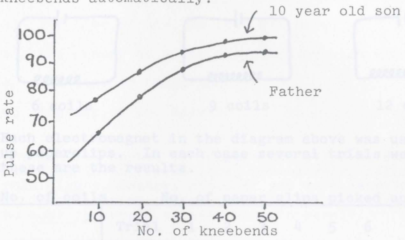
a) Histogram I

b) Histogram II

c) Histogram III

d) Histogram IV

2. The graph below shows the change in pulse rate of a father and his 10 year old son while they were doing 50 kneebends. The pulse rate was measured every ten kneebends automatically.



Look at the graph.

Which of the following statements is correct?

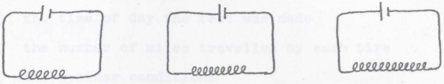
a) The son's pulse rate increased more during the exercise than did his father's.

b) The pulse rate of the father increases more during the exercise than the pulse rate of the son.

c) The son's pulse rate after 50 kneebends is twice that of his father.

d) Both the father and son have the same pulse rate at the end of the experiment.

An experiment was done to see if an electromagnet could be made stronger by putting more coils in the wire as in the diagram below.



6 coils   9 coils   12 coils

Each electromagnet in the diagram above was used to pick up paperclips. In each case several trials were made and these are the results.

No. of coils	No. of paper clips picked up each time					
Trial	1	2	3	4	5	6
6	1	2	2	3	3	1
9	4	4	5	5	2	4
12	7	6	6	6	4	7

3. What is the mean (average) number of paper clips picked up with 6 coils?

   a) 6

   b) 5

   c) 2

   d) 4

4. Which number of coils picks up the greatest range of paper clip

   a) 6

   b) 9

   c) 12

   d) 6 or 9

5. A car tire company wanted to know if their new tires would last as long as their old ones. If you were carrying out such a test which variables do you think it would be most important to keep under control?

   a) the time of day the test was made

   b) the number of miles travelled by each tire

   c) the weather conditions

   d) the age of the driver

6. A scientist wanted to compare the pulse rates of people involved in different jobs, for example, office worker, road digger, doctor, etc. If you were carrying out such an investigation, which variable do you think it would be most important to keep under control?

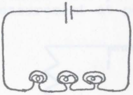
   a) the air temperature

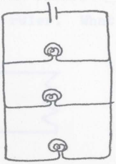
   b) the age of the people

   c) the kind of job that they do

   d) the sex of the people

An experiment was set up to compare the brightness of three bulbs when they were connected up in series and in parallel as shown in the diagram below.





In an experiment such as this

7. Which variable was changed by the experimenter?

   a) number of batteries

   b) number of bulbs

   c) Type of circuit (series or parallel)

   d) brightness of the bulbs

8. In the above experiment which variables were held constant?

   a) number of batteries

   b) number of bulbs

   c) Type of circuit (series or parallel)

   d) both a) and b).

9. In the above experiment which variable would you measure to see if it had changed?

   a) number of bulbs

   b) brightness of the bulbs

   c) both a) and b)

   d) none of the above

10. Which variables were not under control but which may have altered the results?

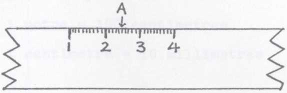
    a) The temperature of the air

    b) The distance apart of the bulbs

    c) The strength of the battery

    d) Both b) and c)

11. Look at this portion of a metre ruler. What is the measurement at point A?



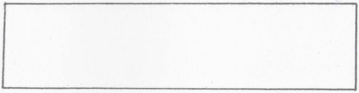
a) 2 centimetres 5 millimetres

b) 2 millimetres 5 centimetres

c) 2 centimetres 6 millimetres

d) 3 centimetres 4 millimetres

12. Using the metric ruler on your desk measure the long side of this rectangle to the nearest centimetre.



The length is:

a) 10 centimetres

b) 9 centimetres

c) 8 centimetres

d) 11 centimetres

13. In the metric system

1 metre = 100 centimetres

1 centimetre = 10 millimetres

Convert

144 centimetres into metres and centimetres

The correct answer is:

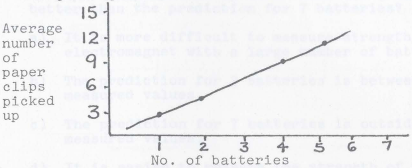
a) 1 metre 14 centimetres

b) 14 metres 4 centimetres

c) 1 metre 44 centimetres

d) 1 metre 44 millimetres

The following graph shows how increasing the number of batteries in a circuit increases the strength of an electromagnet. The strength of the electromagnet is measured by the average number of paper clips that it can pick up.



14. Predict from the graph the average number of paper clips picked up when there are 3 batteries in the circuit.

a) 5

b) 7

c) 3

d) none of the above.

15. Predict from the graph the average number of paper clips picked up when there are 7 batteries in the circuit.

a) 12

b) 15

c) 10

d) none of the above

16. Why do you think the prediction for the average number of paper clips picked up when there are 3 batteries is better than the prediction for 7 batteries?

    a) It is more difficult to measure strength of an electromagnet with a large number of batteries.

    b) The prediction for 3 batteries is between two measured values.

    c) The prediction for 7 batteries is outside any measured values.

    d) It is easier to measure the strength of an electromagnet with a large number of batteries.

17. Look at the graph on the page before this one. If the electromagnet picks up 7 paper clips how many batteries must there be in the circuit.

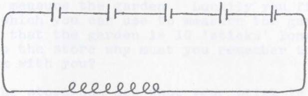
    a) 4

    b) 3

    c) 5

    d) 2

18. A student did an experiment to see how many paper clips an electromagnet would pick up when there were 5 batteries in the circuit as shown in the diagram below. He repeated the experiment three times and these are his results.



Results

Trial	No. of paper clips picked up
1	9
2	7
3	5

Since these results vary quite a bit, how could the student best get a more accurate result?

a) take one more reading and see which of the three it is nearest to

b) take several more readings and average the results

c) average the three results he already has

d) none of the above

19. Your mother sends you to the store to buy enough clothes line to stretch from one end of the garden to the other. The problem is you don't have a ruler to measure the garden. Luckily you find a 'stick' which you can use to measure the garden. You find that the garden is 10 'sticks' long. When you go to the store why must you remember to take the stick with you?

a) so the store can see what your stick is made of

b) so the store can see what shape your stick is

c) so the store knows how long your stick is

d) all of the above

20. If you were to construct a brightness meter, which of the following would be <u>most</u> important?

a) the length of the brightness meter

b) the width of the brightness meter

c) the type of paper

d) all of the pieces of paper must be the same type.

POST TEST 2 FORM A

1. An experiment was carried out with two groups of students to compare how many words they could remember out of a list of 25 words, when the time between learning the words and being asked to remember them was changed.

Here are the results of the experiments.

Group A

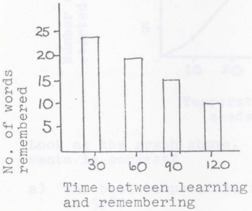
Time between learning and remembering	No. of words remembered
30 minutes	23
60 minutes	18
90 minutes	12
120 minutes	7

Group B

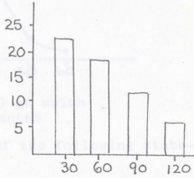
Time between learning and remembering	No. of words remembered
30 minutes	23
60 minutes	19
90 minutes	15
120 minutes	9

Look at these four histograms and answer the question below.

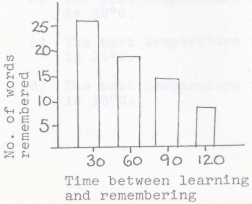
Histogram I



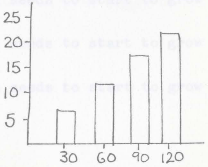
Histogram II



Histogram III



Histogram IV



Which histogram represents the data for group B?

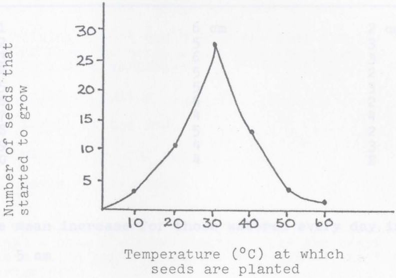
a) Histogram I

b) Histogram II

c) Histogram III

d) Histogram IV

2. Thirty bean seeds were planted at each of the following temperatures 10°C, 20°C, 30°C, 40°C, 50°C, 60°C. The following graph shows how many of the 30 seeds started to grow at each temperature.



Look at the graph above. Which of the following statements is correct?

a) The best temperature for bean seeds to start to grow is 59°C.

b) The best temperature for bean seeds to start to grow is 30°C.

c) The best temperature for bean seeds to start to grow is 26°C.

d) The best temperature for bean seeds to start to grow is 10°C.

A scientist read that tomato plants watered every day grow taller than tomato plants watered only every three days. To test this he planted 10 seedlings in each of two boxes labelled A and B. He treated boxes A and B the same except that he watered seedlings A every day and seedlings B every three days. He measured the length of all the seedlings after 10 days.

Seedlings#	Watered every day	Watered every 3 days
1	6 cm	2 cm
2	5	3
3	6	3
4	5	2
5	5	3
6	6	2
7	4	4
8	5	2
9	4	3
10	4	2

3. The mean increase for those watered every day is:

   a) 5 cm

   b) 5 1/2 cm

   c) 6 cm

   d) 6 1/2 cm

4. Which type of watering produced the largest <u>range</u> of results?

   a) every day

   b) every three days

   c) both the same

   d) impossible to say

5. A class was carrying out an experiment to find out what variables might alter the time it takes a rod to roll down a sloping table. The class was given four rods, A, B, C, D.

A was 10 cm. long, hollow and made of plastic
B was 5 cm. long, hollow and made of plastic
C was 10 cm. long, hollow and made of metal
D was 10 cm. long, solid and made of metal

By rolling rods A and C down the same sloping table and timing them the class could test the effect of which of the following variables?

a) solid or hollow

b) length of the rod

c) material of the rod

d) angle of the slope

6. In the same experiment as question 5 if you rolled rods A and B down the same sloping table and timed them, the class could test the effect of which of the following variables?

a) length of rod

b) material of rod

c) solid or hollow

d) angle of the slope

A class was interested to find out whether boys remembered more words than girls after studying a group of words for 3 minutes.

7. In an experiment such as this which variable would be changed by the experimenter?

   a) whether the learner is a boy or girl

   b) number of words to be learned

   c) number of words learned

   d) time allowed to learn the words

8. In the above experiment which variable must be kept constant?

   a) whether the learner is a boy or girl

   b) number of words to be learned

   c) time allowed to learn the words

   d) both b and c

9. In the above experiment which variable would you measure to see if it had changed as a result of the experimenter?

   a) whether the learner is a boy or girl

   b) number of words to be learned

   c) number of words learned

   d) time allowed to learn the words

10. Which variable was not under control but which may have altered the results?

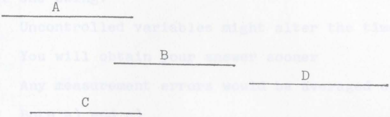
    a) the number of words to be learned

    b) whether the learner is boy or girl

    c) whether the learners read a lot of books or not

    d) both a and c

11. Using your metric ruler measure these four lines to the nearest millimetre.



Which of the following shows the lines A,B,C,D in order of increasing length from shortest to longest?

a) A B C D

b) B C A D

c) C B A D

d) D B C A

12. If you measured your height and found that you were 1 metre 20 centimetres tall, what would be your height in millimetres?

a) 130 millimetres

b) 13 millimetres

c) 1300 millimetres

d) none of the above

13. In an experiment to measure the time of swing of a pendulum why do you think it would be more accurate to time five swings and then find the average time for one swing?

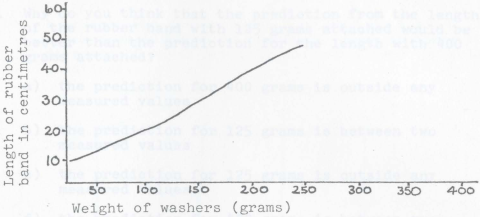
a) Uncontrolled variables might alter the time of swing

b) You will obtain your answer sooner

c) Any measurement errors would be averaged out

d) Both a) and c)

A group of students were looking at what happens to the length of a rubber band when washers of different weights were attached to it. Here is a graph of their results.



14. Predict the length of the rubber band when a 125 gram weight is attached.

 a) 20 centimetres

 b) 30 centimetres

 c) 25 centimetres

 d) 100 centimetres

15. Predict the length of the rubber band when a 400 gram weight is attached.

 a) 45 centimetres

 b) 50 centimetres

 c) 55 centimetres

 d) 60 centimetres

16. How many grams are needed to change the length of the rubber band from 30 to 40 centimetres?

 a) 5 grams

 b) 50 grams

 c) 200 grams

 d) 150 grams

17. Why do you think that the prediction from the length of the rubber band with 125 grams attached would be better than the prediction for the length with 400 grams attached?

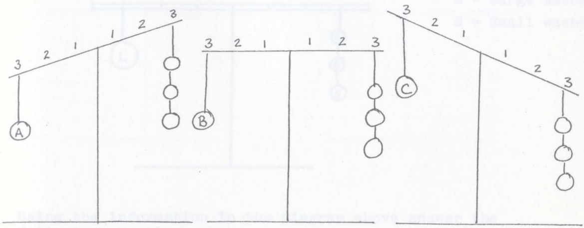
a) the prediction for 400 grams is outside any measured values

b) the prediction for 125 grams is between two measured values

c) the prediction for 125 grams is outside any measured values

d) the prediction for 400 grams is between two measured values

18. A student was asked to put 3 objects A, B and C, in order from lightest to heaviest. Look at these three diagrams of equal arm balances and answer the following questions.



Which of the following shows the 3 objects in increasing order from lightest to heaviest?

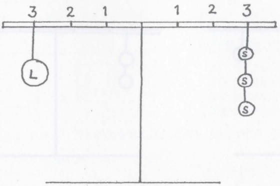
a) A B C

b) C B A

c) B A C

d) B C A

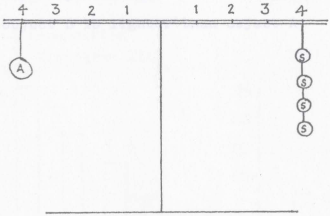
19. Look carefully at the diagram below which shows how many small washers are equal in weight to one large washer.



L = Large washer
S = Small washer

Using the information in the diagram above answer the following question.

If you were given 2 large washers and 2 small washers, which two or three of them would you use at A in the diagram below to balance the washers on the right?



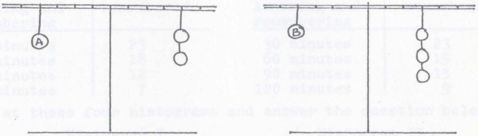
a) 2 large washers

b) 1 large washer and 2 small washers

c) 1 large washer and 1 small washer

d) 2 large and 1 small washer

20. Look carefully at the diagrams below in which two objects A and B are balanced with weights.



Which of the following statements is correct?

a) Object A is heavier than object B

b) Object B is heavier than object A

c) Both objects have the same weight

d) Object B is lighter than object A

POST TEST 2 FORM B

1. An experiment was carried out with two groups of students to compare how many words they could remember out of a list of 25 words, when the time between learning the words and being asked to remember them was changed.

Here are the results of the experiments.

Group A

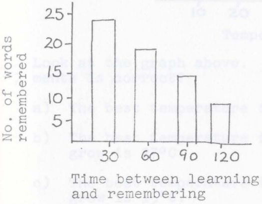
Time between learning and remembering	No. of words remembered
30 minutes	23
60 minutes	18
90 minutes	12
120 minutes	7

Group B

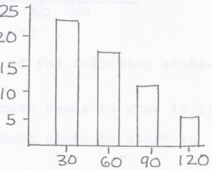
Time between learning and remembering	No. of words remembered
30 minutes	23
60 minutes	19
90 minutes	15
120 minutes	9

Look at these four histograms and answer the question below.

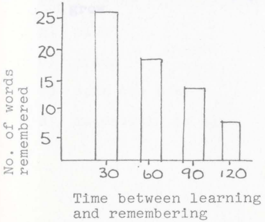
Histogram I



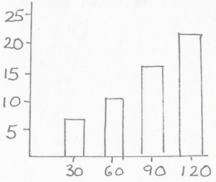
Histogram II



Histogram III



Histogram IV



Which histogram represents the data for group A?

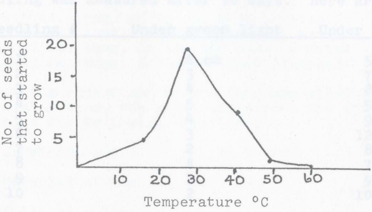
a) Histogram I

b) Histogram II

c) Histogram III

d) Histogram IV

2. Twenty tomato seeds were planted at each of the following temperatures 10°C, 20°C, 30°C, 40°C, 50°C, and 60°C. The following graph shows how many of the 20 seeds started to grow at each temperature.



Look at the graph above. Which of the following statements is correct?

a) The best temperature for tomato seeds to grow is 19°C.

b) The best temperature for tomato seeds to start to grow is 27°C.

c) The best temperature for tomato seeds to start to grow is 15°C.

d) 60°C is the best temperature for seeds to start to grow.

A student was asked to find out whether seedlings growing under a blue light grew taller than seedlings growing under a green light. 10 seeds were planted under a blue light and 10 seeds were planted under green light. The length of each seedling was measured after 10 days. Here are the results:

Seedling #	Under green light	Under blue light
1	6 cm	5 cm
2	3	7
3	4	8
4	5	5
5	4	9
6	3	12
7	2	8
8	4	7
9	3	9
10	2	10

3. The mean increase for seedings grown under blue light is:

   a) 7 cm

   b) 8 cm

   c) 6 cm

   d) 9 cm

4. Which color light produces the largest range of results?

   a) blue

   b) green

   c) both the same

   d) impossible to say

5. A class was carrying out an experiment to find out what variable might alter the time it takes a rod to roll down a sloping table. The class was given four rods A, B, C, D.

   A was 5 cm. long, solid and made of plastic
   B was 15 cm. long, solid and made of metal
   C was 10 cm. long, hollow and made of metal
   D was 10 cm. long, solid and made of metal

   By rolling rods C and D down the same sloping table and timing them, the class could test the effect of which of the following variables?

   a) length of rod

   b) material of rod

   c) solid or hollow

   d) angle of slope

6. In the same experiment as question 5, if you rolled rods A and B down the same sloping table and timed them, the class could test the effect of which of the following variables?

   a) length of rod
   b) material of rod
   c) solid or hollow
   d) angle of slope

A class was interested in finding out whether the age of a person had any effect on the number of words that could be remembered after studying for three minutes.

7. In an experiment such as this which variable would be changed by the experimenter?

   a) the number of words to be learned

   b) the number of words remembered

   c) the age of the person learning the words

   d) the time allowed to learn the words

8. In the above experiment which variable must be kept constant?

   a) the number of words to be learned

   b) the number of words remembered

   c) the age of the person learning the words

   d) both a and c

9. In the above experiment which variable would you measure to see if it had changed as a result of the experiment?

   a) the number of words to be learned

   b) the number of words remembered

   c) the age of the person learning the words

   d) both a and c

10. Which variable was not under control but which may have altered the results?

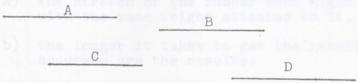
   a) the number of words to be learned

   b) the age of the person

   c) whether the learners read a lot of books or not

   d) both b and c

11. Using your metric ruler measure these four lines to the nearest millimetre.



Which of the following shows the lines A,B,C,D in order of increasing length from shortest to longest?

a) A B C D

b) B C A D

c) C B A D

d) D B C A

12. If you found that your height was 1 metre 50 centimetres, what would be your height in millimetres?

a) 150 millimetres

b) 15 millimetres

c) 1500 millimetres

d) none of the above

13. In an experiment to measure the stretch of a rubber band, why would it be more accurate to repeat each measurement several times and average the results?

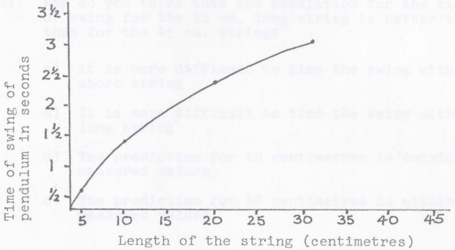
    a) the stretch of the rubber band might vary even with the same weight attached to it.

    b) the longer it takes to get the results the more accurate are the results.

    c) any measurement errors would be averaged out.

    d) both a) and c).

A group of students were investigating the effect of changing the length of the string on the time of swing of a pendulum. Here is a graph showing their results.



14. Predict the time of swing of the pendulum when the string is 10 cm.

a) 1 1/2 seconds

b) 2 seconds

c) 1 second

d) 2 1/2 seconds

15. Predict the time of the pendulum when the string is 45 cms.

a) 3 1/2 seconds

b) 3 1/4 seconds

c) 3 seconds

d) 2 1/2 seconds

16. What increase in length of the string would change the time of the swing from 3-4 seconds?

a) 10 centimetres

b) 15 centimetres

c) 5 centimetres

d) 25 centimetres

17. Why do you think that the prediction for the time of swing for the 10 cm. long string is better than that for the 45 cm. string?

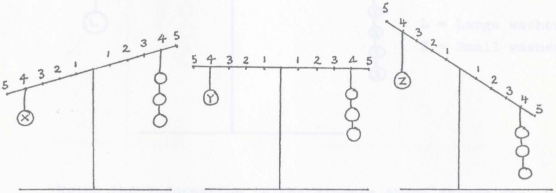
    a) It is more difficult to time the swing with a short string

    b) It is more difficult to time the swing with a long string

    c) The prediction for 10 centimetres is outside any measured values

    d) The prediction for 10 centimetres is within two measured values.

18. A student was asked to put 3 objects X, Y and Z in order from heaviest to lightest. Look at the following diagram and answer the question.



Which of the following shows the 3 objects in increasing order from lightest to heaviest?

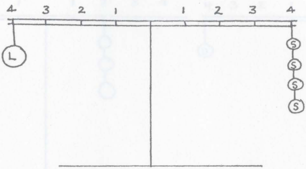
a) Y X Z

b) X Y Z

c) Z Y X

d) Z X Y

19. Look carefully at the diagram below which shows how many small washers are equal in weight to one large washer.

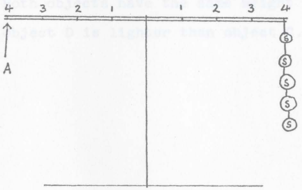


L = Large washer

S = Small washer

Using the information in the diagram above answer the following question.

If you were given 1 large washer and 4 small washers, which two or three of them would you use at A in the diagram below to balance the washers on the right?



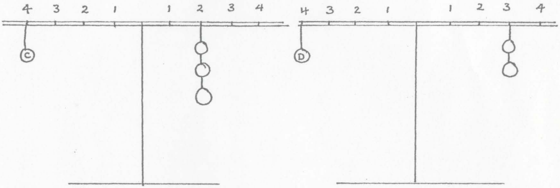
a) 2 large washers

b) 1 large washer and 2 small washers

c) 1 large washer and 3 small washers

d) 1 large washer and 1 small washer

20. Look carefully at the diagrams below in which two objects C and D are balanced with weights.



Which of the following statements is correct?

a) Object C is heavier than object D

b) Object D is heavier than object C

c) Both objects have the same weight

d) Object D is lighter than object C.